Horizontal Educational Inequalities and Civil Conflict:

The nexus of ethnicity, inequality, and violent conflict

Katharine Lindquist

Carleton College

Abstract

Prominent quantitative studies on civil conflict have brought into question the relationship between inequality, ethnic heterogeneity, and conflict. This paper argues that much of the current literature focuses on only one dimension of inequality – that between individuals – and fails to account for the multidimensional aspect of conflict. This paper thus analyzes the relationship of horizontal inequality – that between groups sharing a common identity – with civil conflict, focusing on horizontal educational inequality (HEI) and its effect on propensity for civil conflict and ethnic-based civil conflict. This study includes 44 countries for which national household survey data was collected through Demographic and Health Surveys from 1986 to 2005. Results of logistic regressions show that measures of female and male HEI are marginally significant or not significant in predicting the onset of civil conflict, but both female and male HEI are strongly significant in predicting the onset of ethnic civil conflict. Furthermore, male HEI is not shown to affect propensity for conflict at a higher level than female HEI.

1. Introduction

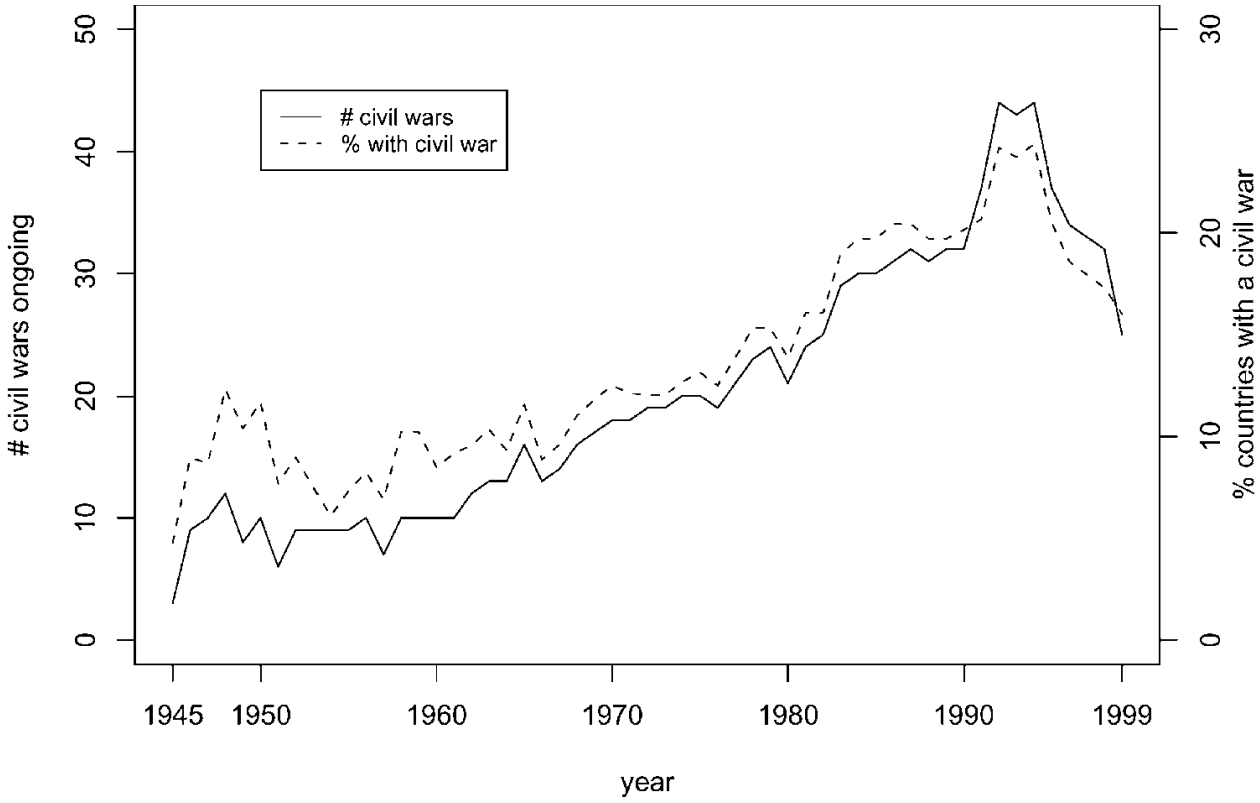
"Simple inequality between rich and poor is not enough to cause violent conflict. What is highly explosive is […] 'horizontal' inequality: when power and resources are unequally distributed between groups that are also differentiated in other ways – for instance by race, religion or language."

- Kofi Annan, 1999

Over the past half-century, civil conflict has been prevalent in a majority of the world's developing countries. Civil conflicts, measured by a threshold of 25 battle-related deaths per year, have occurred in over half of all countries since 1960. Civil wars, measured by a threshold of 1000 battle-related deaths per year, have occurred in a third. Certain regions of the world, such as sub-Saharan Africa, have consistently been embroiled in intrastate conflicts; at one point during the mid-1990s, a third of sub-Saharan African countries were involved in civil conflict at the same time (Blattman & Miguel, 2008). Beyond its growing prevalence, civil conflict is also linked to devastating social and economic impacts. For these reasons, there is strong consensus among economists on the need to increase our understanding of the causes of civil conflict, as

can be seen in the sharp rise in conflict research since the 1990s (Dixon, 2009). A graph of the number and percentage of countries with ongoing civil wars over the period 1945 – 1999 can be seen in Figure 1 (Fearon & Laitin, 2003).

FIGURE 1: Number and Percentage of Countries with Ongoing Civil Wars, 1945-1999



Yet among the research and resulting publications, there is still little consensus on many of the most commonly proposed indicators of conflict. In an aggregation of results from 46 quantitative studies on civil conflict, Dixon (2009) found clear consensus on no more than seven of over 200 tested variables, and even among those seven, he found contention over definition and measurement. Two of the most commonly included indicators are measures of inequality and ethnic heterogeneity. However, while theoretically supported by much of the qualitative literature, measures of both inequality and ethnic heterogeneity were found to be insignificant in

two of the most prominent quantitative analyses of civil war onset published in the last decade – studies by Fearon & Laitin (2003) and Collier & Hoeffler (2004).

In reaction to these results, scholars are now focusing on measures of inequality and ethnic division that better capture the group aspect of conflict (Østby, 2008). As stated by Kofi Annan in a 1999 address to the World Bank, inequality between groups is likely to lead to conflict, not simply inequality between individuals. Furthermore, inequality that coincides with ethnic divides may be particularly dangerous (Annan, 1999; Stewart, 2000; Østby, 2008). This type of inequality – formed between groups sharing a common identity – has been referred to as "horizontal inequality" in recent literature (Stewart, 2000; Østby, 2008). Different from "vertical inequality" that captures differences between individuals (often measured by Gini coefficients), horizontal inequality (HI) captures aspects of both group and socio-economic polarization. Thus, a measure of HI may better reflect the multidimensional nature of inequity that leads to conflict.

Through extensive case study work, Stewart has added significantly to the literature on HI by examining the relationship between inequality and conflict in developing countries; however, due to the micro-level nature of her work, she typically selects on the dependent or independent variables, limiting the broader application of her results (Stewart, 2000; 2001; Stewart, Brown, & Mancini, 2005). For this reason, Østby has sought to conduct the first large-N analyses of HI in developing countries, using survey data from Demographic and Health Surveys (DHS) to build measures of HI for socio-economic factors (Østby, 2005; 2006; 2008; Østby, Nordås, & Rød, 2006).

This paper builds upon Østby's work examining HIs in education between ethnic groups in developing countries (Østby, 2008). While Østby found horizontal educational inequality (in this paper referred to as HEI) to be marginally significant in predicting the onset of civil conflict,

my analysis seeks to strengthen or refute her results by extending her methodology to a broader sample of countries and years and to data on both females and males.[1] Furthermore, I will broaden my independent variable to include measures of both civil conflict and ethnic civil conflict at two different violence thresholds. Much of the literature on ethnicity and violence indicates that the causes of ethnic and non-ethnic conflicts may be significantly different, and I will examine if this holds true for HEI across my sample countries and years (Sambanis, 2001).

The structure of my paper is as follows: I first discuss the literature on ethnic heterogeneity and inequality in relation to civil conflict, focusing on the differences between vertical inequality, polarization, and horizontal inequality. I then present a theoretical framework for studying horizontal educational inequality, followed by my hypotheses and research design. After discussion of variable selection and my model, I present the results of my analysis. My main findings are that while female and male HEI are marginally significant or not significant in predicting civil conflict across both violence thresholds (in line with Østby's results), both female and male HEI are strongly significant in predicting the onset of ethnic civil conflict, and female HEI is significant in predicting ethnic civil war. Furthermore, male HEI interacts with conflict in different ways than female HEI. I end my paper with a discussion of the limitations and further extensions of my study and some concluding remarks.

2. Literature Review

Over the past decade, two of the most prominent and comprehensive studies of civil conflict onset have been published by economists seeking to establish the rationale behind group mobilization. Fearon & Laitin (2003) and Collier & Hoeffler (2004) both drew on theory that conceives the decision to engage in civil conflict as an expected utility calculation, weighing the

[1] For reasons discussed in the following section, Østby restricts her analysis to females.

expected gains from rebellion, given a set of grievances, with the expected losses, including the opportunity cost of forgone production.[2] Theories regarding income and asset inequality and ethnic heterogeneity played into both analyses.

Collier & Hoeffler (2004) modeled the demand for rebel labor as a function of collective grievances – including income inequality – and the supply of rebel labor as a function of the economic costs and benefits of rebellion – including social fractionalization. Income inequality was measured by a Gini coefficient and by the ratio of the top-to-bottom quintiles of income, and asset inequality was measured by the Gini coefficient of land ownership. Despite their hypothesis that the poor may rebel to induce wealth redistribution or that the rich may rebel to further secure their economic status, the measures of inequality were found to be insignificant. Collier & Hoeffler also hypothesized that rebel armies would prefer to recruit along ethnic lines, since ethnic diversity within groups may cause frictions; thus, a society with a high degree of ethnic fractionalization would limit rebels to small recruitment pools, perhaps increasing the costs and decreasing the propensity of conflict. Conversely, they hypothesized that a society that is ethnically polarized, or divided into a few larger groups, would have a higher risk of conflict, as there would be a higher possibility of disputes developing between distinct ethnic groups. They measured ethnic diversity using an index of ethno-linguistic fractionalization[3] and adopted a measure of ethnic polarization[4] from Esteban & Ray. However, once again, they found both measures to be insignificant in their models.

[2] An elaboration of this theory in relation to my study can be found in the subsequent section on the theoretical framework for group mobilization (2.3).
[3] Interpreted as the probability that two randomly drawn people within a country will be from different ethnic groups.
[4] Measures of fractionalization increase when the number of ethnic groups in a society increases; measures of polarization increase when there are few (equally) large ethnic groups exhibiting within-group similarities.

Fearon & Laitin (2003) also included income inequality in their analyses to test for possible grievances between groups, but like Collier & Hoeffler (2004), they found a measure of the Gini coefficient for income inequality to be insignificant. Concerning ethnicity, however, they took an alternative approach and hypothesized that there may be a negative relationship between the level of ethnic diversity and propensity for civil conflict, as more diversity provides grounds for more disputes. They also tested for a possible inverted U-shaped relationship, with increased risk for either very high or very low levels of fractionalization. However, they found these measures of ethnic fractionalization to be insignificant in their models. Theorizing that ethnic diversity may lead specifically to ethnic conflict, Fearon & Laitin reran their analyses with ethnic conflict as the independent variable. However, despite the theoretical relationship, they found no significant empirical relationship between ethnic fractionalization and ethnic civil conflict.[5]

2.1 Issues with the Inequality–Conflict Literature

While the findings of Collier & Hoeffler (2004) and Fearon & Laitin (2003) seem to suggest that there exists no significant relationship between inequality, ethnicity, and civil conflict, I argue that their conclusions may be misleading, as both studies failed to account for the multidimensional aspect of conflict. Both studies examined grievances and opportunities for conflict, but they only looked at one dimension of each. Measures of income or asset inequality calculate relative deprivation between individuals in a country, referred to as "vertical inequality." Thus, these measures fail to incorporate the fact that rebellions are staged by groups, not individuals, and that groups need a common identity behind which to unite.

[5] It should be noted that while Collier & Hoeffler and Fearon & Laitin have published the most prominent studies on civil conflict onset, their results concerning inequality and ethnic heterogeneity have been supported throughout much of the conflict literature (Dixon, 2009; Sambanis, 2002; Østby, 2008).

Measures of ethnic heterogeneity or polarization, on the other hand, provide evidence of unified groups that could possibly engage in conflict, but they fail to provide any compelling reason for conflict. In fact, many countries are characterized by high levels of ethnic diversity but experience no significant conflict, such as Tanzania and Ghana, and it may be due to the lack of socio-economic inequalities (Østby, 2008).

Due to this disconnect in the literature on inequality and conflict, scholars have more recently turned to measures that capture the multidimensional aspect of conflict – the need for clearly identified groups (e.g. high ethnic polarization) and also for inter-group grievances (e.g. inequality). Inequalities that form between culturally formed groups have become known as "horizontal inequalities" (HIs). The most comprehensive work on HIs comes from extensive case studies by Stewart across developing countries (Stewart, 2000; 2001; Stewart, Brown, & Mancini, 2005). However, while contributing substantially to our understanding of HIs on a micro-level, her work is restricted to case studies and she typically selects on the dependent or independent variables, looking at countries where HIs have been found to lead to conflict. This limits the application of her work to a broader range of countries or a more systematic approach.

Another weakness of the inequality-conflict literature is that most studies, including the influential works of Collier & Hoeffler (2004) and Fearon & Laitin (2003), look only at aspects of economic inequality. While many economic factors have been shown to play a significant role in predicting civil conflict, income and asset inequality are not the only factors that may lead to grievances or that signal potential costs and benefits of rebellion. Stewart's work has covered many aspects of inequality other than income, such as political and social welfare factors (Stewart, 2002). One aspect of social welfare that may be a particularly important factor in terms of group grievances is education, because inequality in education is more likely to be the result

of visible, systematic discrimination than income or even asset inequality is. In many documented cases, the ruling power in a country has used state resources to limit educational access to minorities, such as in South Africa and Sri Lanka (Gurr, 2000; Murshed & Gates, 2005; Stewart, 2002). Inequality stemming from such blatantly discriminatory actions may provide an impetus for group mobilization, and it is this aspect of inequality that I hope to capture in my study.

2.2 Horizontal Educational Inequality

The first large-N studies of horizontal inequality done in relation to conflict have been conducted by Østby, and these analyses include both economic and social factors of HI (Østby, 2005; 2006; 2008; Østby, Nordås, & Rød, 2006). While measures of economic HI for income and asset distribution were consistently found to be insignificant in her work, Østby found some support for a measure of social HI, proxied by years of schooling. In my study, I will refer to this measure as horizontal educational inequality (HEI). To test for its significance, Østby created an index of HEI from survey data and regressed it on civil conflict onset, controlling for three commonly included variables in the conflict literature – GDP per capita, population, and conflict history. She found HEI to be significant at the 10% level for predicting the onset of civil conflict, a result she claimed to be noteworthy given the small sample size and non-significance of the other variables included in the model (Østby, 2008).

However, there are a few discrepancies in Østby's theory that may undermine her results. First, Østby creates her indicators for HI based solely on data gathered from females, but she does not acknowledge this in her work. She most likely focuses on females because of the nature of Demographic and Health Surveys (DHS), the source for her explanatory HI variables, which focus on females and households rather than males. However, I would expect female and male

HEI to interact differently with conflict, perhaps at least with different magnitudes, so I believe the assumption that they perform the same may be misleading. Another critique is that Østby does not examine ethnic conflict as an independent variable in her models, despite the inherent definition of HI as promoting group mobilization along ethnic lines. Much of the conflict literature on ethnic mobilization suggests decreased transaction costs and increased intra-group trust when groups act along ethnic lines, decreasing the costs and therefore increasing the likelihood of conflict (Fearon & Laitin, 2003; Sambanis, 2001; Østby, 2008). I would expect a stronger relationship between HEI and ethnic civil conflict than I would between HEI and overall civil conflict. Finally, Østby is limited in her analysis by the availability of DHS surveys. Yet since the time of her study, several more DHS surveys on a wider range of countries have become available. Testing her methodology on a more representative sample would serve as a desirable robustness check to her analysis.

2.3 Theoretical Framework for Group Mobilization

In developing a theoretical framework for my own study of HEI and the propensity for civil conflict, I adopt the intuition behind a model by Esteban & Ray (2008) which seeks to explain the behavior of participants in civil conflict.[6] Esteban & Ray describe conflict as a rent-seeking game, where in deciding whether to rebel, participants weigh the expected payoffs of participating in conflict against the payoffs of accepting the status quo distribution of a public good. For my study, I will assume that the public good at stake is education (proxied by years of schooling) and that there are two groups per country vying for access.

[6] The model presented in the 2008 paper is one application of a broader model developed earlier (Esteban & Ray, 1999). However, for the purposes of this paper, the model presented in the 2008 paper includes all relative aspects for examining the distribution of a public good between two distinct groups.

To measure the expected payoffs of conflict, Esteban & Ray assume that each individual in a group that decides to rebel incurs some cost of resources used in rebellion, measured by $c(r_i)$, where $i = 1, 2$ indicates membership in one of the two groups. Each individual that participates in rebellion also gains some expected benefit in increased access to education. This is measured by the difference in utility between a preferred educational outcome and the current status quo outcome, b_i, multiplied by the probability of attaining the preferred outcome, p_i.[7] The expected utility of an individual who rebels is thus given by

$$u_i = p_i b_i - c(r_i). \quad (1)$$

The expected utility for participating in conflict given in (1) is then compared with the utility derived from retaining the status quo distribution of education. Let this status quo utility be measured by γ_i for each individual in a group. By assuming rationality in maximizing utility, it can be seen that members of a group will rebel if and only if

$$u_i \geq \gamma_i \text{ for } i = 1, 2. \quad (2)$$

Thus, the decision to rebel is a function of the expected gains from rebellion, the probability of staging a successful rebellion, the cost of staging a rebellion, and the level of satisfaction with the status quo distribution. If the expected gains in utility from fighting never outweigh the utility derived from the status quo, then an equilibrium will be reached at which both groups are at peace.

The simple intuition used to derive (2) provides a basis for predicting the relationship between HEIs and the propensity for civil conflict. Since for conflict to take place, $u_i \geq \gamma_i$ must occur for one of the groups, I would expect a higher propensity for conflict when γ_i is low for one of the groups than when γ_i is high, *ceteris paribus*. Since γ_i is the utility derived from the

[7] In Esteban & Ray (2008), this probability is defined as a function of the group size, but it could also be seen as a measure of wealth, power, motivation, etc.

status quo distribution of education in a country, this is likely to be low for one group when there are large disparities in years of schooling between the two groups. As noted in the earlier discussion of horizontal inequalities, this disparity is captured by a measure of HEI, thus leading to my first hypothesis:

H_1: The higher the level of HEI in a country, the higher the risk of civil conflict.

Since HEI is a measure of inequality between ethnic groups, I would also expect a strong relationship between the level of HEI in a country and the risk of ethnic civil conflict. Ethnic conflict is typically defined as civil conflict in which armed organizations both explicitly pursue ethno-nationalist aims and also recruit fighters and forge alliances along ethnic lines. Scholars have theorized that groups with shared identities have lower costs of rebellion, since they can more easily recruit from within the identity group, they are less burdened by collective action problems due to suspicions between group members, and often have political or cultural symbols and ideals to rally behind (Cederman, Wimmer, & Min, 2010; Fearon, 2004; Gurr, 2000; Sambanis, 2001). I theorize that this decrease in costs for ethnic mobilization, measured in (2) by a drop in $c(r_i)$, leads to an increase in the expected utility from rebellion, measured by u_i, thus leading to an increase in the likelihood that $u_i \geq \gamma_i$, inducing an ethnic group into rebellion. This theory provides me with my second hypothesis:

H_2: The higher the level of HEI in a country, the higher the risk of ethnic *civil conflict.*

A final application of (2) in my study is that a group that places a higher value on education would experience a higher value of b_i, since it takes into account the utility derived from any given level of education. Since males are the primary income earners in many developing countries, they stand to gain more economically from an increase in education than

do females. This may suggest that given the same level of HEI for males and females, males may perceive more value in the potential to increase their access to education. A higher level of b_i in (2) would correspond with a higher level of u_i, again increasing the likelihood of rebellion. This provides me with my third hypothesis regarding gender differences within measures of HEI:

> *H_3: Given equal levels of HEI, there will be a higher risk for civil conflict and/or ethnic civil conflict associated with male HEI than with female HEI.*

This hypothesis may gain additional support from conclusions drawn by Collier & Hoeffler (2004) that males with low levels of secondary education are more likely to join rebellions, as their opportunity cost of fighting is much lower. This would play into the utility maximizing function for males as a lower $c(r_i)$, leading to a higher u_i, but would not play into the function for females, as they typically do not supply rebel labor. This difference may further increase the likelihood of my third hypothesis holding true.

3. Data & Methodology

This study seeks to test the hypotheses presented above through an empirical analysis of all countries in which at least one Demographic and Health Survey was conducted over the period 1986 to 2005. DHS surveys are part of an ongoing research project by the US Agency for International Development (USAID) and are intended to collect reliable information on the welfare of women and households in developing countries. To be qualified for my analysis, the surveys needed to include data on ethnicity and years of schooling attained for each participant. As some surveys excluded data on ethnicity due to the sensitivity of ethnicity information in many developing countries, the scope of my study was limited to 93 surveys on 44 countries for females and 57 surveys on 32 countries for males. The number of valid surveys is fewer for

males because DHS data focuses primarily on females; however, many countries have begun to collect information on males more consistently in recent years, allowing for the inclusion of male indicators in my study.

While relying on DHS surveys limits the scope of my study, it also allows for the creation of highly reliable group-level indicators of horizontal inequalities, as done by Østby in several studies (Østby, 2006; 2008; Østby, Nordås, & Rød, 2006). DHS surveys are characterized by large sample sizes, random sampling, and standardized questionnaires, leading to the creation of reliable indicators across countries and across time. Any validity problems are also minimized by the simplicity of the questions used for this study, since questions on ethnicity and years of schooling should require straightforward answers. The list of all DHS surveys and countries used in this study can be found in Appendix A. Furthermore, a map of the geographic distribution of countries can be found in Appendix B.

3.1 Dependent Variables: Civil and Ethnic Civil Conflict

To test my hypotheses, I will employ four separate variables of civil conflict onset in my analysis – civil conflict, civil war, ethnic civil conflict, and ethnic civil war. The data come from the Ethnic Armed Conflict dataset compiled by Cederman, Min, & Wimmer and are based on the UCDP/PRIO Armed Conflicts Data Set (ACD), Version 3-2005b (Gleditsch et at. 2002). Each variable is binary and takes the value of 1 for years in which a conflict begins and 0 otherwise. *Civil conflict* is defined as any armed and organized confrontation between government troops and rebel organizations or between army factions that reaches an annual threshold of 25 battle-related deaths. *Civil war* is similarly defined, but with an annual threshold of 1000 battle-related deaths. *Ethnic civil conflict* is defined as civil conflict in which armed organizations both explicitly pursue ethno-nationalist aims and interests and also recruit fighters

and forge alliances along ethnic lines. Ethnic civil conflict refers to an annual threshold of 25 battle-related deaths, whereas *ethnic civil war* is similarly defined but refers to an annual threshold of 1000 battle-related deaths. It should be noted that by these definitions, *civil conflict/ethnic civil conflict* include those conflicts also identified as wars, and *civil conflict/civil war* include those conflicts also identified as ethnic. The frequency of each type of conflict in my dataset can be seen in Table 1 below.

TABLE 1: Frequency of conflict onset in the data (female and male data sets)

FEMALE		**Identity**	
		Ethnic and Non-ethnic	Ethnic only
Intensity	Civil Conflict	39	28
	Civil War	20	13

MALE		**Identity**	
		Ethnic and Non-ethnic	Ethnic only
Intensity	Civil Conflict	28	22
	Civil War	16	12

3.2 Independent Variable: Horizontal Educational Inequality

My measure of horizontal educational inequality is adopted from Østby (2008) and is based on information from DHS survey data on ethnic affinity and years of schooling. Following the approach of Brockerhoff & Hewett (2000) that was applied by Østby, I focus on the two largest ethnic groups in each country. In doing so, I assume that the level of inequality between the two largest ethnic groups within a country is most important in predicting civil conflict. My indicator of HEI is generated using two variables from the DHS surveys for females/males: v131/mv131 (ethnicity) and v133/mv133 (total years of education completed). In

order to compare HEIs across countries, a measure is needed that reflects levels of inequality on a standardized scale. To do so, Østby uses the following formula:

$$\text{HEI} = 1 - exp\left(-\left|ln\left(E_1/E_2\right)\right|\right) \quad (3)$$

where E_1 is the average educational attainment in years of schooling for the first ethnic group and E_2 is the average educational attainment for the second ethnic group. This measure is scaled from 0 to 1, with 0 indicating perfect equality in educational attainment (both groups have same average years of schooling) and 1 indicating that one group has all of the educational attainment (measured in years of schooling) and the other group has none. In my dataset, the measure of female HEI ranges from .000 to .800 and the measure of male HEI ranges from .014 to .749 (the highest actual level of observed inequality).

The values calculated for HEI are then interpolated between survey years and extrapolated to cover the time span of the study, 1986-2005. Simple linear interpolation is used between survey years. Extrapolation is done by adopting the value of HEI from the survey closest in time for previous and subsequent years within the time period. For countries with only one survey, the value of HEI for that survey is used for all years in order to increase the sample size of the study. While this may weaken the validity of my independent variable, it is important to note that inequality often changes slowly over time, and thus the accuracy of the measure still remains reasonably strong (Østby, 2006). Another important limitation of my data is that the sample suffers from potential selection bias. DHS surveys are not conducted in countries that are heavily war-ridden, and for some countries that do host DHS surveys, ethnicity data is censored due to its highly sensitive nature. It should be noted, however, that these selection biases tend to omit countries that potentially would further strengthen my hypotheses, and thus their omission puts a possible downward bias on the results (Oswald, 2010; Østby, 2008).

3.3 Control Variables

I control for other robustly significant variables in the civil conflict literature, focusing on those included in Fearon & Laitin's (2003) model of civil conflict onset. As noted throughout the civil conflict literature, Fearon & Laitin have employed perhaps the most comprehensive and well-supported model of conflict onset to this date, and their approach has been replicated in several other studies of civil conflict initiation (Blattman & Miguel, 2008; Dixon, 2009; Sambanis, 2001; 2002; Thyne, 2006). As noted by Dixon (2009), all seven of the independent variables found by Fearon & Laitin to be significant have been well supported in subsequent literature. Thus, I will include these controls in my analysis. Other authors have shied from including more than three or four control variables in their analyses due to concerns over multicollinearity; however, as can be seen by the covariance table in Appendix C, collinearity does not seem to be a considerable issue for my data (Oswald, 2010; Østby, 2008). Furthermore, rerunning the models with fewer control variables was found to not significantly alter the main results. To deal with possible problems of endogeneity in my models, all time variate control variables listed below are lagged by one year, following the methodology of Fearon & Laitin (2003) and others.

Three of the most commonly controlled for factors in models of civil conflict onset are economic development, population size, and conflict history. The first, economic development, is proxied by *GDP per capita* and is given in constant 2000 US Dollars. The negative effect of GDP per capita on propensity for civil conflict is one of the most widely supported relationships in the literature. The second control variable, log transformed *population size*, is theorized to have a positive relationship with civil conflict onset, as a large population makes it easier for rebels to evade notice and detainment by the central authorities and also provides a larger

recruitment base. Thirdly, I control for the number of years since the previous conflict, *peace years*, at the given violence threshold – so, years since the previous conflict for the conflict variables and years since the previous war for the war variables. This aims to control for any temporal dependence within the data.[8]

Factors concerning the political environment in a country often are associated with the risk of civil conflict breakout. It is theorized that democracy should exhibit a negative relationship with civil conflict, since democracies provide a political environment in which all citizens have some political power, reducing discrimination and repression of certain groups. A dummy variable for *democracy* is thus included, measured by a score of +6 or higher on the Polity IV scale. It is also hypothesized that autocracies may be associated with less civil conflict, since an authoritative state is well equipped to suppress rebellion. To test this line of thinking, a control for political systems that are neither democratic nor autocratic, known as anocracies, is included. *Anocracy* should thus exhibit a positive relationship with civil conflict, and it is coded as a dummy variable for political systems that lie between -6 and +6 on the Polity IV scale. A third control is included for political environments that are unstable, since recent changes in the central government may indicate disorganization and potential to be overthrown by rebels. Thus, a dummy variable for *regime change* is included that indicates a movement along the Polity IV scale of 3 or more points over the prior three years. Finally, it is hypothesized that the government of a new state may be weak within the first few years of formation, as it has not yet settled into a stable structure and may still have new or untested military operations. Since this

[8] I also ran my analyses with cubic splines to further control for temporal dependence, following the methodology of Østby (2008) and others; however, inclusion of the splines did not significantly alter my findings, so I excluded them from the results shown in this paper. For an explanation of the rationale for cubic splines, see Beck, Katz, & Tucker (1998).

may leave the government vulnerable to rebellion, a dummy variable for *new state* is included, coded for countries in the first two years of existence.

Lastly, three controls related to geography and natural resources are included. Countries that depend heavily on oil exports tend to have weaker centralized governments, as rulers can depend on oil rents and do not require elaborate bureaucratic systems to raise tax revenues. Furthermore, capturing oil revenues may serve as an incentive to gain control of the state.[9] Thus, *oil production per capita*, as measured by the annual oil production in metric tons divided by the population, is included in my analysis. Similar to the rationale for a large population, rough terrain should favor rebellion as it makes it easier for rebels to evade notice and capture from the central government. I follow Fearon & Laitin (2003) by controlling for the percent of a country covered in *mountainous terrain*, according to the coding of geographer A. J. Gerard. Finally, it is predicted that a country that includes some territories separated from the state's center by water or distance (for example, East Pakistan (now Bangladesh) from West Pakistan or an island nation such as the Philippines) may be at higher risk for rebellion, as the government may not be able to retain control over distant or separate territories. To control for this, a dummy for *noncontiguous state* is included.

3.4 Empirical Model

Due to the binary nature of my dependent variables, I employ a logistic regression model to test my hypotheses. This model predicts the odds of civil conflict onset by fitting data to a logistic curve. Following the methodology of Østby and others, I use country-clustered standard errors, which relaxes the assumption of completely independent observations from the same

[9] An alternative explanation is provided by Wimmer & Min (2006), who hypothesize that states that depend on oil rents instead of a constituent tax base may develop higher degrees of clientelism, especially along ethnic lines. This ethnic exclusion may provide the impetus for mobilization against the state, leading to higher prevalence of civil conflict.

countries over time (Sambanis, 2001; Østby, 2008). It should be noted however, that running the regressions using regular robust standard errors instead of robust clustered standard errors did not significantly alter my results. Analyses were performed using STATA, version 11.1. Descriptive statistics for each variable can be found in Appendix D.

4. Results

As noted previously, I employ logistic regression models to test my hypothesized relationships between HEI and civil conflict. I start by presenting the results of my analysis of female HEI, measured for 44 countries from 1986 to 2005. To test my first hypothesis – that higher levels of HEI are associated with a higher propensity for civil conflict – I regress the onset of civil conflict and civil war on female HEI and my control variables. The results of these regressions can be seen in Model 1 and Model 2 presented in Table 2 below.

In the first model, I find female HEI to be significant and positively correlated with increased propensity for civil conflict onset. This finding supports the analysis done by Østby (2008) on a smaller set of countries and years, in which she found HEI to be positive and significantly related to conflict at the 10% level. The marginal effect of HEI is also strong, as a 0.1 increase in the value of HEI (on a 0 to 1 scale) is associated with a 26% increase in the odds of civil conflict onset. This means that the odds of civil conflict onset for the maximum level of observed HEI in this dataset are 3.28 times higher than the odds of conflict for the average level of HEI. In my model of civil war initiation (Model 2), female HEI is not found to be significant. This may suggest that a high level of HEI for females may mobilize groups to engage in minor conflicts, but it is not enough to justify large-scale conflicts that are much more costly.

TABLE 2: Conflict Onset in 44 Developing Countries, 1986 – 2005

	Model 1	*Model 2*	*Model 3*	*Model 4*
	Civil Conflict	**Civil War**	**Ethnic Civil Conflict**	**Ethnic Civil War**
Female HEI	2.303**	1.990	4.466***	3.848**
	(0.916)	(1.304)	(0.960)	(1.577)
GDP per capita†	-0.241	-0.260	-0.950***	-1.235***
	(0.193)	(0.255)	(0.362)	(0.382)
Population (ln)†	0.006	0.126	-0.263	0.003
	(0.159)	(0.217)	(0.187)	(0.296)
Peace years (Civil conflict)	0.015		0.030*	
	(0.012)		(0.016)	
Peace years (Civil war)		-0.012		0.004
		(0.013)		(0.018)
Democracy†	-1.025	-0.846*	-0.684	-1.166**
	(0.627)	(0.491)	(0.817)	(0.569)
Anocracy†	-0.027	0.027	-0.020	-0.337
	(0.430)	(0.669)	(0.532)	(0.773)
New state	3.509***	3.031***	6.950***	7.867***
	(1.149)	(0.995)	(1.792)	(1.272)
Regime change	-0.421	0.049	-0.268	0.236
	(0.323)	(0.576)	(0.407)	(0.789)
Oil production per capita†	0.165	0.231	0.690***	0.884***
	(0.155)	(0.202)	(0.216)	(0.183)
Mountainous terrain (ln)	-0.008	0.178	-0.069	0.138
	(0.096)	(0.165)	(0.183)	(0.284)
Noncontiguous state	2.118***	1.852**	4.247***	4.213***
	(0.719)	(0.912)	(0.846)	(1.009)
Constant	-3.474	-5.425***	-1.681	-4.585
	(1.258)	(1.901)	(1.548)	(3.059)
N Observations	834	834	834	834
N Conflict onsets	39	20	28	13

Logit estimates with robust country-clustered standard errors in parentheses.
*$p < 0.10$; **$p \leq 0.05$; ***$p \leq 0.01$
† Lagged one year

Turning to ethnic conflict, Model 3 from Table 2 shows that female HEI is strongly associated with the onset of ethnic conflict and is significant at the 1% level. It is also significant in magnitude, as an increase in the level of HEI of 0.1 (on a 0 to 1 scale) is associated with a 56% increase in the odds of ethnic conflict initiation. This means that the odds of ethnic conflict onset at the highest observed level of female HEI are 10 times the odds of conflict onset at the

mean value of HEI. This result lends support to my second hypothesis that high levels of HEI affect propensity for ethnic conflict at a higher magnitude than they affect civil conflict. Furthermore, this relationship holds for higher thresholds of violence. Female HEI significantly predicts the onset of ethnic civil war, whereas it was not found to be significant in the model of overall civil war. The percent change in odds associated with an increase in female HEI is less for ethnic civil war than for ethnic civil conflict but is still greater than for overall civil conflict – an increase of .1 is associated with a 47% increase in the odds of ethnic civil war onset. This translates into the odds of onset at the highest level of observed female HEI being 7.27 times the odds of onset at the average level of HEI.

In the four models presented in Table 2, only *new state* and *noncontiguous state* are consistently significant in predicting the onset of conflict, and they both contribute heavily in magnitude to the odds of conflict onset. For countries in my dataset, being in the first two years of existence increases the odds of civil conflict and civil war onset by 33 and 20 times, respectively, and increases the odds of ethnic civil conflict and ethnic civil war by 1043 and 2611 times, respectively. Being a noncontiguous state increases the odds of civil conflict and civil war onset by 8.3 and 6.4 times, respectively, and increases the odds of ethnic civil conflict and ethnic civil war by 70 and 68 times, respectively. The extreme nature of these results is likely due to the small size of my dataset, which includes a higher percentage of noncontiguous countries and newly formed countries than a more representative sample; however, the results still signal a strong relationship that has been supported both theoretically and empirically in conflict literature.

GDP per capita and *oil production per capita* are found to be significant in the models of ethnic conflict and war but not in those for overall conflict and war. A one-dollar increase in per

capita GDP is associated with 61.3% decrease in the odds of ethnic civil conflict onset and a 70.9% decrease in the odds of ethnic civil war onset. This supports much of the literature on civil conflict that predicts a decrease in conflict with higher levels of economic development. High levels of oil production, theorized to increase the risk of conflict due to dependence on oil rents, is found to increase the odds of ethnic civil conflict by 99% per extra metric ton of oil per capita, and increase the odds of ethnic civil war by 142% for the same increment.

Analyzing the same models, but using male HEI and 32 countries from 1986 to 2005, I am able to compare the differences between female and male HEI and their relationship with civil conflict onset. Models 5 and 6 in Table 3 again test my first hypothesis, and similar to the results from the female HEI data, it can be seen that male HEI is marginally significant in predicting the onset of civil conflict and not significant in predicting civil war. This shows some support that male and female measures of HEI seem to exhibit parallel effects on civil conflict. The economic significance of male HEI in Model 5 is also very similar to that for female HEI in Model 1; an increase of 0.1 in the value of HEI (on a 0 to 1 scale) is associated with a 25% increase in the odds of civil conflict onset, whereas it was a 26% increase for female HEI. This begins to draw into question my third hypothesis, which posits that male HEI will be higher in magnitude than female HEI in predicting civil conflict onset.

To further test my second hypothesis, I model the relationship between male HEI and my controls and ethnic civil conflict and ethnic civil war. Once again, ethnic civil conflict is significantly linked with the measure of HEI, this time being for males. However, the magnitude of the significance is smaller in Model 7 than it was for the parallel model of female HEI. A 0.1 increase in the value of male HEI is associated with a 53% increase in the odds of ethnic civil conflict onset (compared to a 56% increase for female HEI). This means that the odds of ethnic

TABLE 3: Conflict Onset in 32 Developing Countries, 1986 – 2005

	Model 5	*Model 6*	*Model 7*	*Model 8*
	Civil Conflict	**Civil War**	**Ethnic Civil Conflict**	**Ethnic Civil War**
Male HEI	2.238*	0.812	4.276***	2.744
	(1.312)	(1.644)	(1.189)	(1.840)
GDP per capita†	-0.929**	-0.687	-1.556***	-1.123***
	(0.463)	(0.430)	(0.463)	(0.297)
Population (ln)†	-0.167	-0.230	-0.340	-0.302
	(0.275)	(0.435)	(0.258)	(0.022)
Peace years (Civil conflict)	0.032**		0.340**	
	(0.475)		(0.014)	
Peace years (Civil war)		-0.004		0.208
		(0.017)		(0.913)
Democracy†	-2.080	---	-1.903	---
	(1.640)		(2.245)	
Anocracy†	0.316	0.597	0.212	0.208
	(0.475)	(0.851)	(0.529)	(0.913)
New state	7.434***	5.053***	10.325***	7.732***
	(1.888)	(1.127)	(2.032)	(1.363)
Regime change	-0.314	0.527	-0.183	0.955
	(0.375)	(0.702)	(0.458)	(0.816)
Oil production per capita†	0.467*	0.544*	0.837***	0.952***
	(0.254)	(0.301)	(0.258)	(0.182)
Mountainous terrain (ln)	0.147	0.588**	0.209	0.746**
	(0.120)	(0.255)	(0.176)	(0.319)
Noncontiguous state	3.182***	1.803***	4.105***	2.032***
	(0.772)	(0.677)	(1.013)	(0.637)
Constant	-1.760	-2.988	-0.603	-3.321
	(2.415)	(3.154)	(2.642)	(3.907)
N Observations	609	609	609	609
N Conflict onsets	28	12	22	8

Logit estimates with robust country-clustered standard errors in parentheses.
*$p < 0.10$; **$p \leq 0.05$; ***$p \leq 0.01$
† Lagged one year
--- Democracy removed from civil war models due to perfect collinearity

civil conflict onset for the maximum level of observed male HEI in this dataset are 8.48 times higher than the odds of conflict for the average level of male HEI. Contrary to my findings for female HEI, however, the male HEI is not significant in the model of ethnic civil war. This fits with the overall trend of a weaker relationship between HEI and high levels of conflict, likely

because rebels may need a considerably larger motive to engage in such costly activity. This result also further disproves my third hypothesis, as male HEI is not significant in predicting ethnic civil war, but female HEI is significant. One important consideration is the smaller size of the data set for males, as I was limited in scope by the availability of surveys. It may also be unsound to compare male and female estimates from different data sets, as they reflect a different cross-section of countries. Further research into the relationship between female and male HEI will be needed to better inform the accuracy of my third hypothesis.

Similar to the models of female HEI and conflict onset, *new state* and *noncontiguous state* are consistently significant in Models 5-8. For countries in the male dataset, being in the first two years of existence increases the odds of civil conflict and civil war onset by 1692 and 156 times, respectively, and increases the odds of ethnic civil conflict and ethnic civil war by 30492 and 2281 times, respectively. Being a noncontiguous state increases the odds of civil conflict and civil war onset by 24 and 6.1 times, respectively, and increases the odds of ethnic civil conflict and ethnic civil war by 61 and 7.6 times, respectively. Again, the extreme results of the magnitude of new state are likely due to the size of the sample, which may not be representative of all developing countries. Other variables reaching significance in the models include *GDP per capita* for ethnic and civil conflict and ethnic civil war, similar to the female models, and *oil producation per capita* for all models. The number of *peace years* is found to be significant for civil conflicts, but not for civil wars, whereas the percentage of *mountainous terrain* is significant in the models of civil war but not civil conflict. The number of peace years may be significant for civil conflict because of the more frequent nature of conflicts in compariso to wars. Additionally, mountainous terrain may be significant for civil war but not for conflict because only in large scale operations would hiding in the mountains be advantageous.

6. Conclusion

Building primarily on the work of Østby (2008), this study has demonstrated that examining horizontal inequalities, especially in relation to social factors such as education, is an important extension of the classic civil conflict literature. In contrast to prominent studies that have found only insignificant relationships between inequality, ethnic heterogeneity, and civil conflict, horizontal educational inequality was found to be marginally significant in predicting civil conflict and strongly significant in predicting ethnic civil conflict in this study. Furthermore, while Østby was able to establish a link between female HEI and conflict, I was able to extend that relationship to male HEI and higher levels of conflict, increasing the significance of the results.

There are several limitations to my study that I have discussed throughout my paper. One of the biggest restrictions is the small sample size, as 44 countries may not be representative of the population of developing countries for female estimates, and 32 countries for male estimates are likely even less representative. This could lead to skewed estimates in some of my models. The calculation of HEI is also somewhat limiting, as it can only be constructed for country-years in which a DHS survey has taken place, and values must be interpolated and extrapolated from those years. A more accurate measure would be calculated from data on each country-year, but that type of data is not currently available. Promisingly, DHS surveys continue to expand their coverage, so future studies of a similar nature could employ more years and more countries in their analyses.

Future extensions of this study could include other measures of HI concerning economic, political, or other social factors. Establishing the links between these measures and ethnic conflict would be worthwhile, especially given some of the significant results found in this

paper. Furthermore, exploring the relationship between female and male indicators of HI for different socio-economic factors would add to our understanding of the causes of civil conflict. An analysis of the policy implications of these results, while outside of the scope of this paper, would aid greatly in understanding the importance of HI, and would be an interesting complement to the empirical results of my study. Furthermore, it would allow for these results to be applied towards actual policy-making that could aid in reducing the prevalence of civil conflict in developing countries.

6. References

Annan, Kofi. 1999. "Peace and Development: One Struggle, Two Fronts." Available at: http://www.un.org/News/Press/docs/1999/19991019.sgsm7187.doc.html.

Beck, Nathaniel, Jonathan 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. Available at: http://www.jstor.org/stable/2991857.

Blattman, Christopher, and Edward Miguel. 2008. "Civil War" ed. N Shanmugaratnam. *Journal of Economic Literature* 48(1): 3-57. Available at: http://pubs.aeaweb.org/doi/abs/10.1257/jel.48.1.3.

Brockerhoff, M., and P. Hewett. 2000. "Inequality of child mortality among ethnic groups in sub-Saharan Africa." Bulletin of the World Health Organization 78(1): 30-41. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2560588&tool=pmcentrez&rendertype=abstract.

Caselli, Francesco, and Wilbur John Coleman II. 2006. "On the theory of ethnic conflict." *NBER working paper* (732). Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=893780.

Cederman, L-E, and Luc Girardin. 2007. "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies." *American Political Science Review* 101(01): 173. Available at: http://www.journals.cambridge.org/abstract_S0003055407070086.

Collier, Paul. 2006. "Economic causes of civil conflict and their implications for policy." *World Bank*. Available at: http://users.ox.ac.uk/~econpco/research/pdfs/EconomicCausesofCivil Conflict-ImplicationsforPolicy.pdf.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and grievance in civil war." *Oxford Economic Papers* 56(4): 563-595. Available at: http://oep.oupjournals.org/cgi/doi/10.1093/oep/gpf064.

Dixon, Jeffrey. 2009. "What Causes Civil Wars? Integrating Quantitative Research Findings." *International Studies Review* 11(4): 707-735. Available at: http://doi.wiley.com/10.1111/j.1468-2486.2009.00892.

Esteban, Joan, and Debraj Ray. 1999. "Conflict and Distribution." *Journal of Economic Theory* 87(2): 379–415. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0022053199925498.

Esteban, Joan, and Debraj Ray. 2006. "A model of ethnic conflict." *Framework* (0241070). Available at: http://pareto.uab.es/wp/2007/70107.pdf.

Esteban, Joan, and Debraj Ray. 2008. "Polarization, Fractionalization and Conflict." *Journal of Peace Research* 45(2): 163-182. Available at: http://jpr.sagepub.com/cgi/doi/10.1177/0022343307087175.

Fearon, James D. 2004. "Ethnic Mobilization and Ethnic Violence." *Political Science*. Available at: http://www.seminario2005.unal.edu.co/Trabajos/Fearon/Ethnic mobilization and ethnic violence.pdf.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(01): 75. Available at: http://www.journals.cambridge.org/abstract_S0003055403000534.

Gurr, Ted Robert. 2000. *Peoples Versus States: Minorities at Risk in the New Century*. US Institute of Peace Press. Available at: http://books.google.com/books?hl=en&lr=&id=6-hEA_wWpyMC&pgis=1.

Hegre, H., and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution* 50(4): 508-535. Available at: http://www.jstor.org/stable/27638504.

Montalvo, José G, and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review* 95(3): 796-816. Available at: http://www.atypon-link.com/AEAP/doi/abs/10.1257/0002828054201468.

Murshed, S. Mansoob. 2002. "Conflict, Civil War and Underdevelopment: An Introduction." *Journal of Peace Research* 39(4): 387-393. Available at: http://jpr.sagepub.com/cgi/doi/10.1177/0022343302039004001.

Østby, Gudrun. 2005. "Inequality, Institutions, and Instability: Horizontal Inequalities, Political Institutions, and Civil Conflict in Developing Countries 1986–2003." *Political Institutions, Development and a Domestic Civil Peace Workshop in Oxford*. Available at: http://www.prio.no/files/file47438_ostby_paper_piddcp_oxford_2005.pdf.

Østby, Gudrun. 2006. "Horizontal Inequalities, Political Environment and Civil Conflict: Evidence from 55 Developing Countries." *Centre for Research on Inequality, Human Security and Ethnicity (CRISE). Working Paper* 28(28). Available at: http://www.crise.ox.ac.uk/pubs/workingpaper28.pdf.

Østby, Gudrun. 2008. "Polarization, Horizontal Inequalities and Violent Civil Conflict." *Journal of Peace Research* 45(2): 143-162. Available at: http://jpr.sagepub.com/cgi/doi/10.1177/0022343307087169.

Østby, Gudrun, Ragnhild Nordås, and Jan Ketil Rød. 2006. "Regional Inequalities and Civil Conflict in 21 Sub-Saharan African Countries, 1986–2004." *47th Annual Convention of the International Studies Association, San Diego, CA* (April): 22–23. Available at: http://www.prio.no/files/file47720_ostby_nordas_rod_nicosia2006.pdf.

Oswald, Casey. 2010. "Horizontal Inequality, Economic Shock, and Civil Conflict." *Carroll Round Proceedings* 5: 198-210. Available at: http://www7.georgetown.edu/sfs/carrollround/CRP-V.pdf#page=213.

Sambanis, Nicholas. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil War." *Defence and Peace Economics* 13(3): 215-243. Available at: http://www.informaworld.com/openurl?genre=article&doi=10.1080/10242690210976&magic=crossref||D404A21C5BB053405B1A640AFFD44AE3.

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(3): 259-282. Available at: http://jcr.sagepub.com/cgi/doi/10.1177/0022002701045003001.

Stewart, Frances. 2000. "Crisis Prevention: Tackling Horizontal Inequalities." *Oxford Development Studies* 28(3): 245-262. Available at: http://www.informaworld.com/openurl?genre=article&doi=10.1080/713688319&magic=crossref||D404A21C5BB053405B1A640AFFD44AE3.

Stewart, Frances. 2001. *Horizontal Inequalities: A Neglected Dimension of Development*. United Nations University, World Institute for Development Economics Research. Available at: http://www3.qeh.ox.ac.uk/pdf/qehwp/qehwps81.pdf.

Stewart, Frances, Graham Brown, and Luca Mancini. 2005. "Why Horizontal Inequalities Matter: Some Implications for Measurement." *International Meeting on Gini and Lorenz in Commemoration of their Centenary Scientific Research* (19). Available at: http://www.cpc.unc.edu/measure/training/mentor/materials/basic-me-concepts-portuguese/workingpaper19.pdf.

Thyne, Clayton L. 2006. "ABC's, 123's, and the Golden Rule: The Pacifying Effect of Education on Civil War, 1980?1999." *International Studies Quarterly* 50(4): 733-754. Available at: http://doi.wiley.com/10.1111/j.1468-2478.2006.00423.

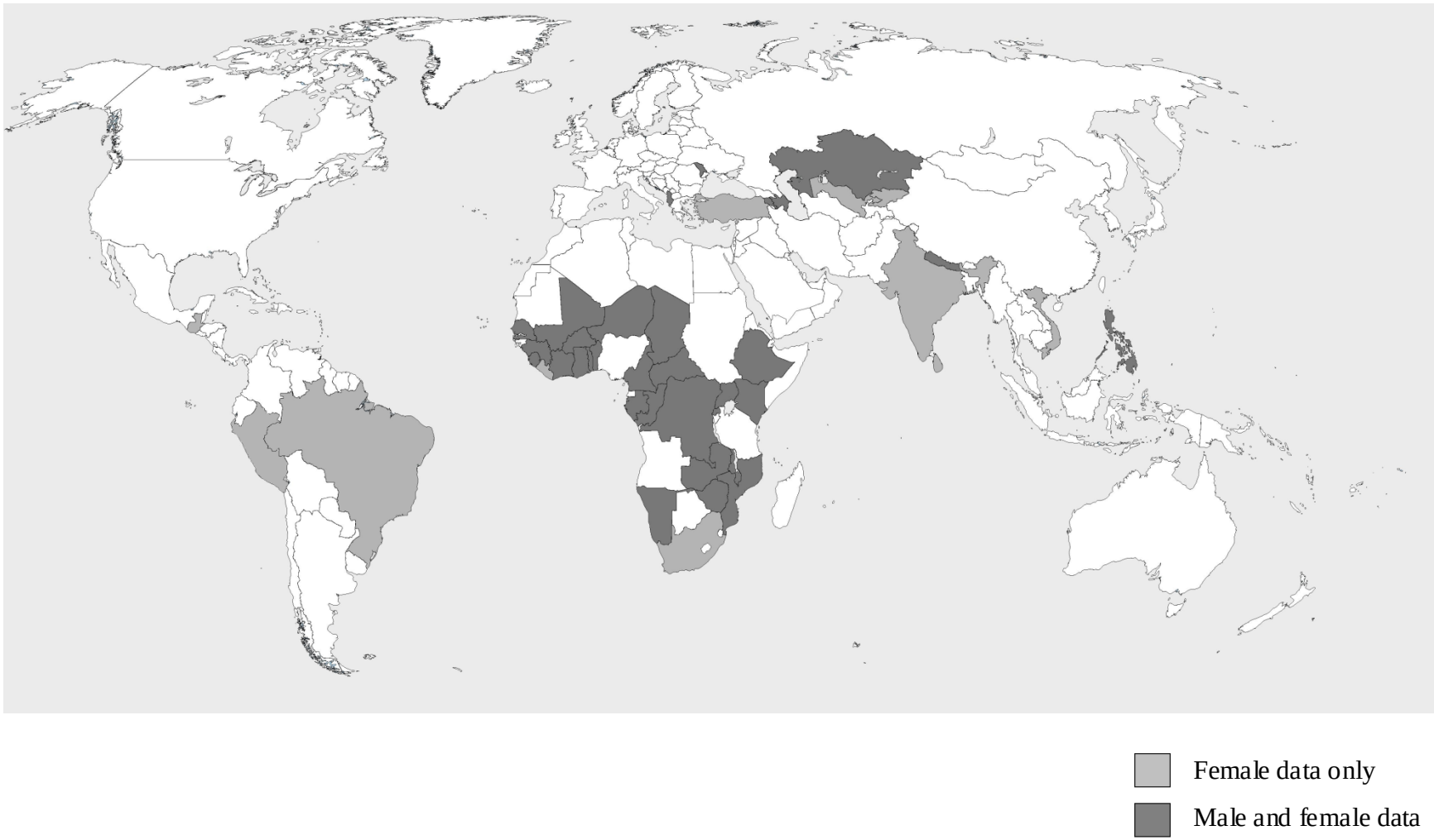
Wimmer, Andreas, L-E Cederman, and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." *American Sociological Review* 74(2): 316-337. Available at: http://asr.sagepub.com/cgi/doi/10.1177/000312240907400208.

Appendix A: Countries and DHS Surveys included in the analysis

Country	*Year*	*Country*	*Year*
Albania	2008*	Mali	1987
Armenia	2000*		1995*
Azerbaijan	2006*		2001*
Benin	1996*		2006*
	2001*	Moldova	2005*
	2006*	Mozambique	1997*
Brazil	1991	Namibia	1992
	1996		2000*
Burkina Faso	1993*	Nepal	1987
	1998*		1996
	2003*		2001*
Cameroon	1991*		2006*
	1998*	Niger	1992*
	2004*		1998*
Central African Republic	1994*		2006*
Chad	1996*	Peru	1991
	2004*		2000
Cote d'Ivorie	1994*	Philippines	1993
	1998		1998
Democratic Republic of the Congo	2007*		2003*
Ethiopia	2000*		2008
	2005*	Republic of the Congo	2005*
Gabon	2000*	Rwanda	1992*
Ghana	1988	Senegal	1986
	1993*		1992*
	1998*		1997*
	2003*		2005*
	2008*	Sierra Leone	2008*
Guatemala	1987	South Africa	1998
	1995	Sri Lanka	1987
	1998	Togo	1988
Guinea	1999*		1998*
	2005*	Trinidad and Tobago	1987
India	1992	Turkey	1993
	1998	Uganda	1988
Kazakhstan	1995		1995*
	1999*		2000
Kenya	1989	Uzbekistan	1996
	1993*	Vietnam	1997
	1998*		2002
	2003*	Zambia	1992
	2008*		1996*
Kyrgyz Republic	1997		2001*
Liberia	1986		2007*
Malawi	2000*	Zimbabwe	1988
	2004*		1994*
			1999

* Survey years including both male and female data (otherwise female only)

Appendix B: Geographic Distribution of Survey Countries



Appendix C: Covariance Tables (female and male data)

	FHEI	GDPPCL	LPOPL	CPYRS	WPYRS	DEML	ANOCL	NWST	RGCHG	OILPCL	LMTN	NCON
Female HEI (FHEI)	1.000											
GDP per capita (GPDPCL)	-0.213	1.000										
Population size (LPOPL)	0.215	-0.131	1.000									
Conflict Peace Years (CPYRS)	-0.142	0.126	-0.177	1.000								
War Peace Years (WPYRS)	-0.203	0.131	-0.175	0.690	1.000							
Democracy (DEML)	0.003	0.282	0.238	0.088	0.065	1.000						
Anocracy (ANOCL)	0.001	-0.117	-0.115	-0.104	0.016	-0.562	1.000					
New state (NWST)	-0.075	0.101	-0.047	-0.072	-0.104	-0.005	0.029	1.000				
Regime change (RGCHG)	0.072	-0.144	-0.048	0.018	0.043	-0.078	0.195	-0.045	1.000			
Oil production (OILPCL)	-0.201	0.698	-0.396	0.148	0.116	-0.030	0.037	-0.006	-0.026	1.000		
Mountainous terrain (LMTN)	0.239	0.008	0.391	-0.213	-0.394	0.057	0.012	0.056	-0.068	-0.254	1.000	
Noncontiguous state (NCON)	-0.094	0.017	0.452	-0.209	-0.295	0.225	-0.172	0.029	-0.003	-0.025	0.189	1.000

	MHEI	GDPPCL	LPOPL	CPYRS	WPYRS	DEML	ANOCL	NWST	RGCHG	OILPCL	LMTN	NCON
Male HEI (MHEI)	1.000											
GDP per capita (GPDPCL)	-0.210	1.000										
Population size (LPOPL)	0.222	-0.401	1.000									
Conflict Peace Years (CPYRS)	0.019	0.160	-0.341	1.000								
War Peace Years (WPYRS)	0.064	-0.002	-0.205	0.644	1.000							
Democracy (DEML)	-0.189	0.012	0.053	-0.012	-0.041	1.000						
Anocracy (ANOCL)	0.031	0.029	0.048	-0.015	0.174	-0.454	1.000					
New state (NWST)	-0.083	0.167	-0.061	-0.081	-0.119	0.045	0.027	1.000				
Regime change (RGCHG)	0.085	-0.084	0.025	0.023	0.078	0.019	0.179	-0.050	1.000			
Oil production (OILPCL)	-0.188	0.806	-0.443	0.183	0.079	-0.134	0.108	0.001	-0.015	1.000		
Mountainous terrain (LMTN)	-0.075	0.007	0.410	-0.221	-0.393	-0.077	0.052	0.054	-0.042	-0.186	1.000	
Noncontiguous state (NCON)	-0.181	0.139	0.293	-0.209	-0.309	0.200	-0.149	0.057	0.037	0.002	0.255	1.000

Appendix D: Descriptive Statistics (female and male data)

Variable (Female)	*N*	*Mean*	*Std. dev.*	*Min.*	*Max.*
Civil Conflict Onset	846	0.046	0.210	0	1
Civil War Onset	846	0.024	0.152	0	1
Ethnic Civil Conflict Onset	846	0.033	0.179	0	1
Ethnic Civil War Onset	846	0.015	0.123	0	1
Female HEI	846	0.285	0.224	0.000	0.800
GDP per capita	839	2.832	2.794	0.171	19.561
Population (ln)	839	9.338	1.277	6.781	13.886
Peace Years (Civil Conflict)	846	12.911	15.372	0	59
Peace Years (Civil War)	846	21.267	18.336	0	59
Democracy	834	0.308	0.462	0	1
Anocracy	834	0.415	0.493	0	1
New State	846	0.017	0.128	0	1
Regime Change	846	0.187	0.390	0	1
Oil Production per capita	839	0.634	2.193	0	16.407
Mountainous Terrain (ln)	846	1.915	1.484	0	4.313
Noncontiguous State	846	0.065	0.247	0	1

Variable (Male)	*N*	*Mean*	*Std. dev.*	*Min.*	*Max.*
Civil Conflict Onset	616	0.045	0.208	0	1
Civil War Onset	616	0.019	0.138	0	1
Ethnic Civil Conflict Onset	616	0.036	0.186	0	1
Ethnic Civil War Onset	616	0.013	0.113	0	1
Male HEI	616	0.249	0.182	0.144	0.749
GDP per capita	611	2.158	2.365	0.338	14.007
Population (ln)	611	9.117	0.941	6.781	11.337
Peace Years (Civil Conflict)	616	13.731	14.560	0	59
Peace Years (Civil War)	616	22.604	16.750	0	59
Democracy	609	0.200	0.401	0	1
Anocracy	609	0.452	0.498	0	1
New State	616	0.016	0.126	0	1
Regime Change	616	0.227	0.419	0	1
Oil Production per capita	611	0.631	2.349	0	16.407
Mountainous Terrain (ln)	616	1.667	1.446	0	4.313
Noncontiguous State	616	0.057	0.232	0	1

Appendix E – Data Appendix

Variable Code Descriptions (Final Data Sets):

year: Year
cowcode: Country code
country: Country name

Dependent Variables:
newonset: Coded 1 for years in which a new war starts
newethonset: Coded 1 for years in which a new ethnic war starts
onsetintensity: Coded 1 for years in which a new low intensity war starts (below 1000 battle-deaths per year); 2 for new high intensity war (above 1000 deaths)
newethintensity: Coded 1 for years in which a new low intensity ethnic war starts (below 1000 battle-deaths per year); 2 for new high intensity ethnic war (above 1000 deaths)
onset1: Coded 1 for years in which a new low intensity war starts (below 1000 battle-deaths per year)
onset3: Coded 1 for years in which a new high intensity war starts (above 1000 battle-deaths per year)
newethonset1: Coded 1 for years in which a new low intensity ethnic war starts (below 1000 battle-deaths per year)
newethonset3: Coded 1 for years in which a new high intensity ethnic war starts (above 1000 battle-deaths per year)

Independent Variable:
HEIf: measure of female horizontal educational inequality (0 to 1 scale, 0 being exactly equal)
HEIm: measure of male horizontal educational inequality (0 to 1 scale, 0 being exactly equal)
HEI2: squared measure of HEI

Control Variables:
gpdcap: GDP per capita (constant 2000 US dollars)
gdpcapl: GDP per capita (constant 2000 US dollars), lagged 1 year
popavg: Population
lpop: ln(Population)
lpopl: ln(Population), lagged 1 year
oilpc: Oil Production per capita
oilpcl: Oil Production per capita, lagged 1 year
anoc: Coded 1 for anocracy, indicated by -6 to +6 on Polity IV scale
anocl: Coded 1 for anocracy, lagged 1 year
democl: Coded 1 for democracy, indicated by +7 to +10 on Polity IV scale, lagged 1 year

regchg3: Measure of instability (regime change), coded 1 if a change in Polity IV score of 3 points or more over prior 3 years
mtnest: Percent of country in mountainous terrain
lmtnest: ln(% mountainous terrain)
ongoingwar: Coded 1 if already an ongoing war in a year
ncontig: Coded 1 if country is non-contiguous
nwstate: Coded 1 if country is in the first 2 years of existence

Time Controls:
npeaceyears: Number of years since last war
nspline*: Splines for npeaceyears
hpeaceyears: Number of years since last high intensity war (above 1000 battle-deaths per year)
hspline*: Splines for hpeaceyears

Additional Controls:
ethfrac: Measure of ethnic fractionalization (probability that two individuals selected at random will be from different ethnic groups)
western: Western country
eeurop: Eastern Europe
lamerica: Latin America
ssafrica: Sub-Saharan Africa
asia: Asia
nafrme: North Africa and Middle East

Inter- and Extrapolation Process

Appendix A lists the DHS surveys used in the analysis. In order to calculate values across the entire period 1986 – 2005 for all included countries, interpolation and extrapolation were used. The interpolation is a simple linear interpolation, in which all observations between two data points are assigned a value according to its position on a straight line between the two data points. Extrapolation is done by holding the value observed constant for observations prior to the first, or after the final observation. Figure 1 below illustrates how this is done in an example with two data points. For countries with only one data point, all other relevant observations are extrapolated from this observed data point.

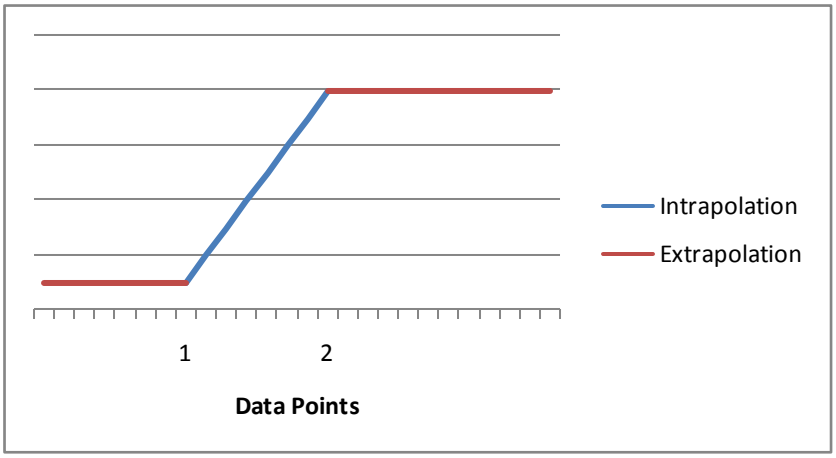


Figure 1: Inter- and Extrapolation

Ethnic war dataset

from Lars-Erik Cederman, Brian Min, and Andreas Wimmer, 2010

"The coding of wars is based on the UCDP/PRIO Armed Conflicts Data Set (ACD) (Gleditsch et al. 2002). ACD defines armed conflict as any armed and organized confrontation between government troops and rebel organizations, or between army factions, that reaches an annual battle-death threshold of 25 people. Massacres and genocides are not included because the victims are neither organized nor armed; communal riots and pogroms are excluded because the government is not directly involved.

"We drew primarily on version 3-2005b of the ACD data set, which provides two levels of conflict identification, a more general war ID number and a disaggregated sub-ID that identifies whenever the constellation of rebel organizations changes completely or when more than 10 years elapse between episodes of violence. We relied on these sub-IDs to construct our own conflict list because we are interested in a disaggregated dependent variable that would allow us to differentiate between conflicts fought by actors claiming to represent different ethnic communities.[10] As a result, we code a larger number of armed conflict onsets than does the original ACD. To preserve comparability with other studies of civil war, we identify high-intensity conflicts as those that reach the standard threshold of 1,000 battle deaths in at least one year. For each conflict, we coded whether actors pursued ethnonationalist aims and if they pursued secessionist objectives.

[10] The ACD data set appears to be more consistent with regard to coding sub-IDs from 1989 onward. We fused sub-IDs that were based on a change in the type of civil war (e.g. internationalized versus non-internationalized conflicts). We also split or merged some wars to be consistent with the rules governing sub-ID coding. A list of our conflicts and how they relate to the ACD war IDs is available upon request.

"Ethnic/nonethnic conflicts are distinguished by the aims of the armed organizations and their recruitment and alliance structures, in line with other ongoing coding projects (Sambanis 2009). Ethnic wars typically involve conflicts over ethnonational selfdetermination, the ethnic balance of power in government, ethnoregional autonomy, ethnic and racial discrimination (whether alleged or real), and language and other cultural rights. We define all other war aims as nonethnic. Examples of nonethnic conflicts include the various military coups staged in Argentina and the civil wars in China, Greece, and Algeria. Regarding recruitment and alliance structures, we define ethnic wars as those fought by armed organizations that recruit fighters predominantly among their own ethnic group and who forge alliances on the basis of ethnic affiliation. For a conflict to be classified as ethnic, armed organizations have to both explicitly pursue ethnonationalist aims, motivations, and interests *and* recruit fighters and forge alliances on the basis of ethnic affiliations.

"We linked all ethnic conflicts to the politically relevant ethnic category in the name of which an armed organization instigated the conflict. We looked at the aims and recruitment patterns of each armed organization separately. In some complex cases (e.g., Afghanistan, Burma, Chad, Uganda, Angola, and Zaire), we disaggregated a conflict into several war fronts with different ethnic claims made on the nongovernmental side. This was necessary when the constellation of rebel organizations changed dramatically over time."

Data set access:

Lars-Erik Cederman; Brian Min; Andreas Wimmer, "Ethnic Armed Conflict dataset", http://hdl.handle.net/1902.1/11797 V1

Østby, Gudrun (2008) 'Polarization, Horizontal Inequalities and Violent Civil Conflict', *Journal of Peace Research* 45(2): 143–16. Available at: http://folk.uio.no/gudruno/.

DHS Survey Data: http://www.measuredhs.com/accesssurveys/start.cfm