

# What's Population Got to Do with it?: The Onset of Civil War in Sub-Saharan Africa

Nobuko Mizoguchi  
Department of Demography  
University of California at Berkeley  
nobukom@demog.berkeley.edu

March 7, 2008

## Abstract

Sub-Saharan Africa has witnessed an upsurge of civil wars in the past two decades. While many theories exist about the different causes of civil war, population factors remain under-explored. The purpose of this paper is to examine the factors associated with the onset of civil war in Sub-Saharan Africa. Logistic regression was conducted using both bivariate and multivariate models and using pooled analysis and country-fixed effects. Results of the pooled analysis showed that total population increased the odds of the incidence of civil war while GDP per capita and GDP growth lowered the odds when controlling for other economic, social and political factors. However, results from country fixed effects models showed that only GDP growth had a significant effect. With population size increasing in many Sub-Saharan countries, a closer examination of their demography may provide some clues to the political instability of the region.

## 1 Introduction

Sub-Saharan Africa has witnessed an upsurge of civil wars in the past two decades. While many theories exist about the different causes of civil war, population factors remain under-explored. Most studies have focused on factors such as income, education level, terrain and ethnic fractionalization (Collier and Hoeffler, 2002; Elbadawi and Sambanis, 2002; Fearon and Laitin, 2003). One study that examined the demographic factors focused on the youth bulge (Urdal, 2006). This paper will build upon this body of

work to more closely examine the demographic factors related to the onsets of civil war in Sub-Saharan Africa.

As can be seen in Figure 1, the 1990's experienced one of the highest incidence of civil war in recent history, along with the 1970's. While in the 1970's, conflicts occurred primarily in Asia, the Americas, and the Middle East, Africa had the largest share of civil war incidence in the 1990's. Furthermore, the total incidence of civil war was much greater in the 1990's and 1970's than in other decades.

One demographic hypothesis on the onset of civil war is the youth bulge. This theory suggests that as the proportion of young men rises, the risk of conflict also increases. Some attribute this to biological reasons, while others argue that Easterlin's relative cohort size effect is in play. A high proportion of young men in a population means that they are faced with lack of educational opportunities as children and lack of economic opportunities as they enter the labor force. These conditions lead to wide-spread discontent among the young men, creating conditions for possible insurgencies to occur.

Scarcity of resources also can be traced back to demographic roots. The Malthusian argument is that as population size increases, there is pressure on limited resources which sets into motion a series of checks that pushes the population down to its equilibrium size. Civil war can be regarded as one of the checks. Rapid population growth and increase in population density can also create competition for resources under this theory.

Other potential causes of civil wars include economic factors such as poverty and income inequality, sometimes resulting from rapid economic development. One model suggests that conflicts occur when opportunities are ripe for insurgency groups to flourish, such as access to capital and a rural base (Collier and Hoeffler, 2001).

Some social and cultural factors include ethnic fractionalization and religious fundamentalism. Ethnic fractionalization refers to the level of ethnic heterogeneity in a region. Both ethnic fractionalization and religious fundamentalism may be fueled by economic factors.

Political factors for conflict include ideological extremism, which is sometimes mixed with religious and ethnic factors, and "state failure" or when states become so weak that they are incapable of governing. Democratization is often considered a preventative factor.

Another political factor may be the length of time the state has been in existence. Relatively new states may be naturally more prone to civil war as governments try to establish themselves. Hence it may be part of a natural life cycle of states to have more civil wars in the first decades of their existence. Since most countries in Sub-Saharan Africa became independent

### Number of Civil War Onsets by Decade and Region, 1820–1999

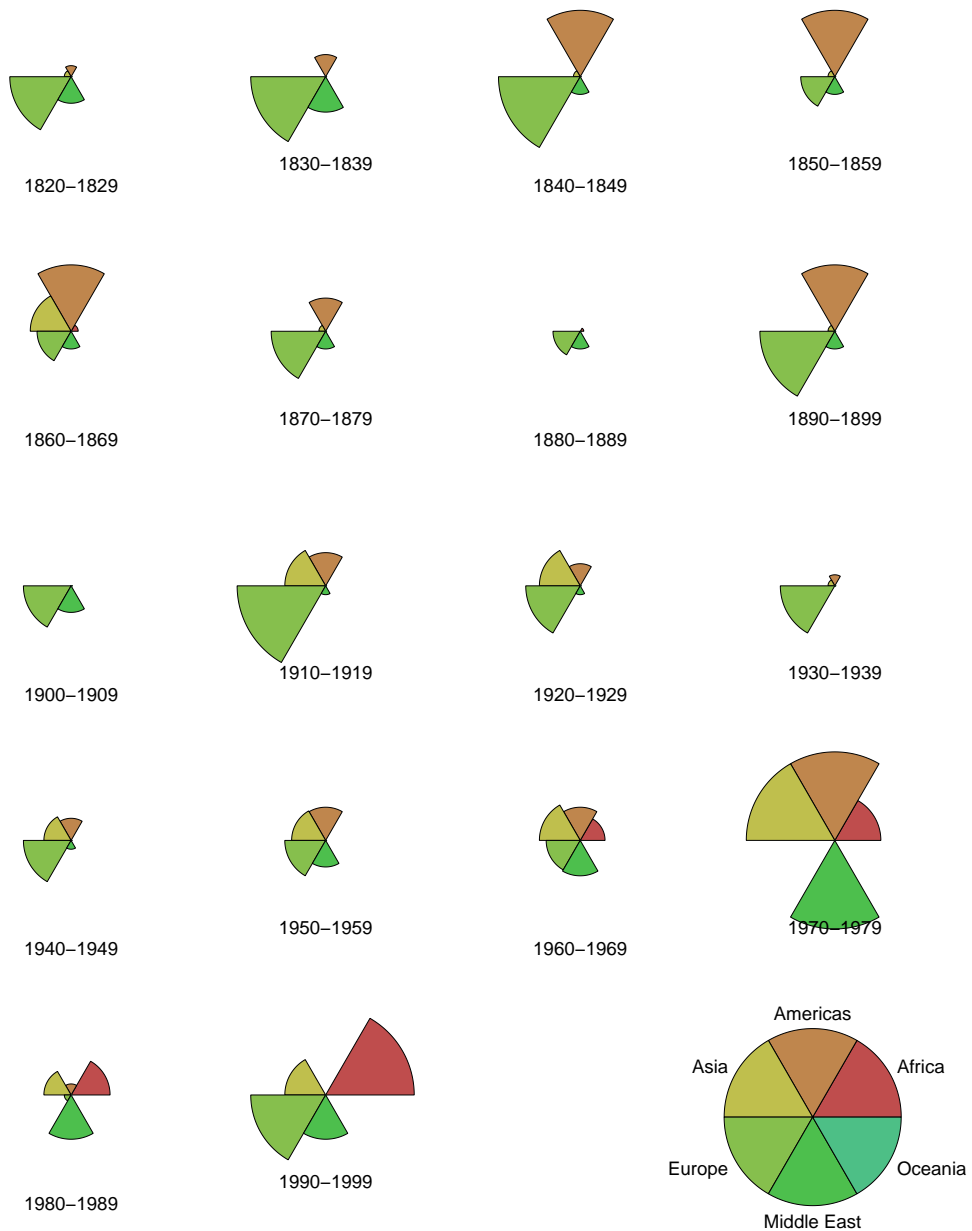


Figure 1: Number of Civil War Onsets by Decade and Region, 1820-1999. The radius of each segment represents the number of civil war onsets for the region and the decade. A total of 419 civil wars were included in the dataset, 152 in Europe, 91 in the Americas, 60 in Africa, 64 in Asia, and 52 in the Middle East. These represent civil wars only and do not include international or colonial conflicts. These charts only show the number of conflicts that started. Therefore, do not capture the magnitude or the duration of the conflicts.

in the latter half of the 20th century with the exceptions of Ethiopia and Libya, they fall into this category of young states.

This paper seeks to examine the association between demographic factors and the incidence of civil war.

## 2 Methods and Data

Descriptive analysis and logistic regression were used to examine the factors associated with the onsets of civil war. Logistic regression was conducted using both bivariate and multivariate models and using pooled analysis and country-fixed effects. Robust estimation of standard errors was used in the pooled analysis in order to account for different variance-covariance matrices for each country.

The database created for this project consists of a separate observation recorded for each year of each country. The dependent variable in the logistic regression was a dummy variable indicating whether there was an onset of civil war during the year. The data on the conflicts are from the expanded version of the Correlates of War datasets created by Kristian Skrede Gleditsch and covers conflicts between 1816 and 2002. The Correlates of War dataset covers only up to 1992. Conflicts are entered in the database if the annual average battle deaths by all participants exceed 1,000. States are defined as those having at least 500,000 total population and have diplomatic recognition by at least two major powers or membership in the United Nations or the League of Nations.

Demographic variables included were total population (in millions), population growth (annual percent), population density (people per square km), and percent youth in the adult population. Percent youth in adult population was defined as the number of young people ages 15-24 years old as a proportion of the total adult population ages 15 and up. This was done to be consistent with Urdal's (2006) measure of youth bulge. Total population, number of youth aged 15-24, and number of adults aged 15 and up were obtained from the U.S. Census, International Data Base. Population growth and population density measures were taken from the World Bank World Development Indicators dataset 1960-2002.

Also included in the model were economic factors. To measure income, gross domestic product (GDP) at purchasing power parity (PPP) per capita in constant 2000 international dollars was included in the model. Measures of poverty and income inequality were unavailable for most countries, so were excluded. Also included were measures of GDP growth as an indicator

of how fast the economy is developing, which may trigger resource scarcity issues. Foreign aid was measured by foreign aid per capita in current US dollars. Foreign debt was measured by total external debt outstanding in current US dollars. Data for the economic indicators are from the World Bank World Development Indicators dataset 1960-2002.

To measure ethnic fractionalization, an index developed by Krain (1997) was included in the model. The index is based on the proportions of ethnic groups in the population and is in a range from 0 to 1. A low score means homogeneous population with high proportion of one ethnic group and a very small portion of a minority group. A high score means that there are many groups with relatively equal percentages of the population. This index does not capture how power and resources may be distributed along these groups. The data was only available to 1990. However, since the index seemed to be relatively stable over time, the last recorded entry for each country was assumed to hold for the following decade. In addition, to make the results of the logistic regression more easily interpretable, the index was multiplied by 100 so that the odds ratios reflect one percent change in the index.

For the political factors, level of democratization and age of the state were included in the model. The polity score in the Polity IV dataset was used as an indicator of democratization. The polity score is an index based on levels of institutionalized autocracy or authoritarianism and institutionalized democracy. Mostly, this score looks at political participation and competition, such as free elections. The hypothesis is that conflicts occur in states with medium levels of polity. The polity scores were split into three categories of low, medium, and high. The high polity score was used as a baseline for the analysis. By categorizing into larger groups, I also hope to reduce the level of arbitrariness in the assignment of the scores. The Correlates of War state system membership dataset was used to calculate the number of years states have been in existence.

### **3 Results of Descriptive Analysis**

Results of the descriptive analysis showed that civil war and non-civil war countries displayed no major differences in the observed variables. In addition, much variation exists among the Sub-Saharan African countries for the indicators of interest, making it difficult to draw definitive conclusions.

### 3.1 Comparison of civil war and non-civil war countries

First, Sub-Saharan African countries were examined by those that experienced civil war and those that did not, to see whether there are any apparent differences in the social, economic, and demographic characteristics. The series of box and whisker plots in Figures 2 through 4 show the results, covering the period from 1960 to 2002. Forty-three countries are included in the analysis. Twenty were in the “civil war” category (807 country years) and 23 were in the “peaceful” category (892 country years). The countries were considered to be civil war countries if they experienced an onset of civil war during 1960-2002. Countries were considered in the peaceful category if they did not experience any conflict during the same time period. Only one country, Eritrea, did not experience civil war but experienced conflict during that time period. It was excluded in the analysis.

These figures show no strong trends out of any of the indicators. The median total population, age dependency ratio, GDP per capita, foreign aid per capita, and ethnic fractionalization levels are slightly higher in civil war countries. The external debt and polity score are both slightly lower in the civil war countries, but not by very much. Predictably, life expectancy is lower in civil war countries. Median population density, population growth, GDP growth are about the same. Contrary to what was expected, the median “age” of the state, or the number of years that states have been existence are slightly higher in civil war countries.

### 3.2 Country-specific comparisons

Because these broad categorizations of civil war and non-civil war countries may mask the trends in individual countries, country-specific trends for these indicators were examined. This analysis showed wide variations within civil war and non-civil war countries. No conclusive trend appeared from the country-specific comparisons of demographic and economic factors. Even the polity score, representing the level of democratization showed no clear differences among civil war and non-civil war countries.

## 4 Results of pooled regression analysis

### 4.1 Bivariate analysis

Bivariate analyses were conducted before proceeding to multivariate analysis. Results of the bivariate analysis are shown in Table 1. The following variables showed statistically significant ( $p$ -value  $< 0.05$ ) relationship to

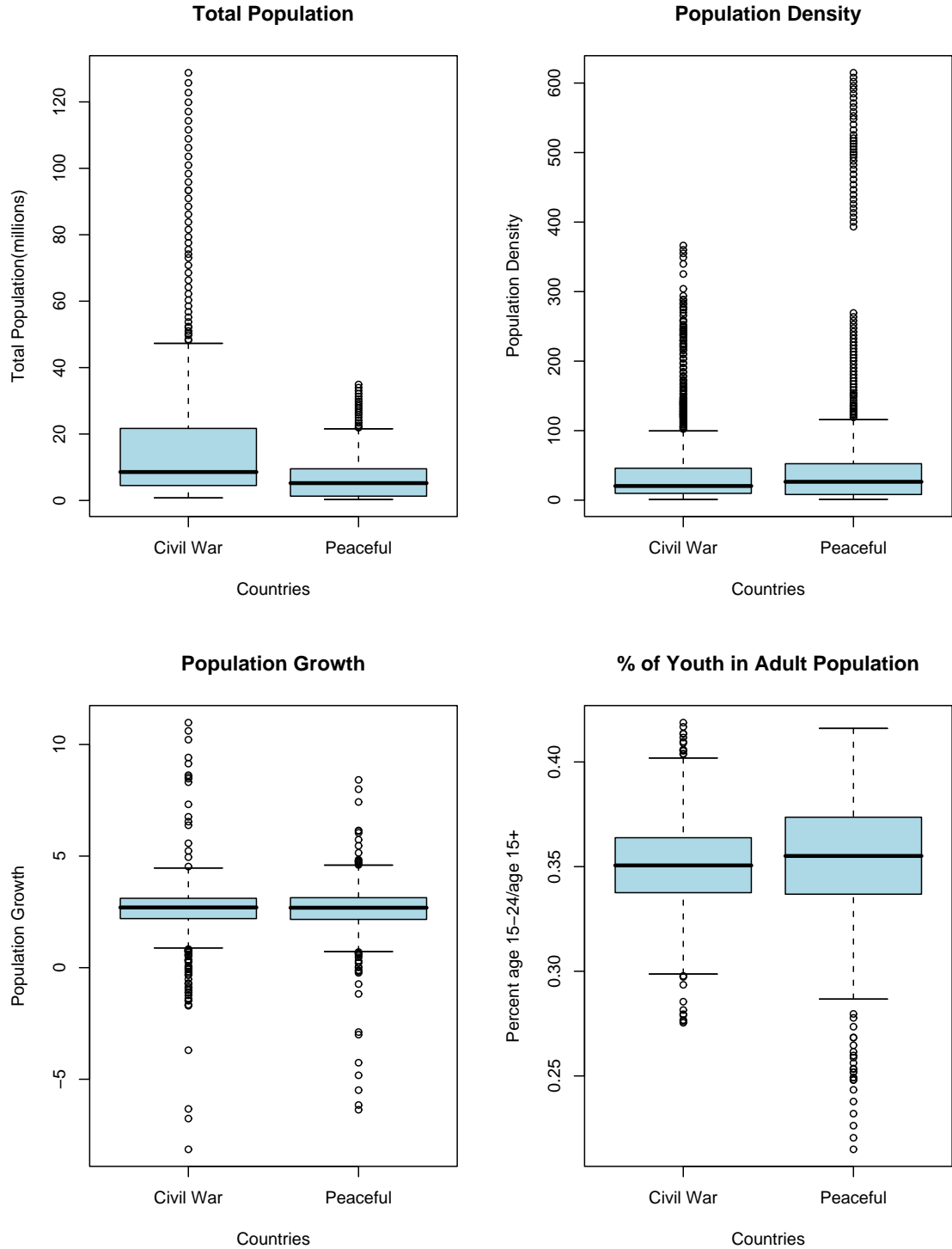


Figure 2: Comparison of possible explanatory factors for onset of civil war by civil war and non-civil war countries in Sub-Saharan Africa, 1960-2002. The thick bar represents the median. The top and bottom of the box represent the first and third quartiles.

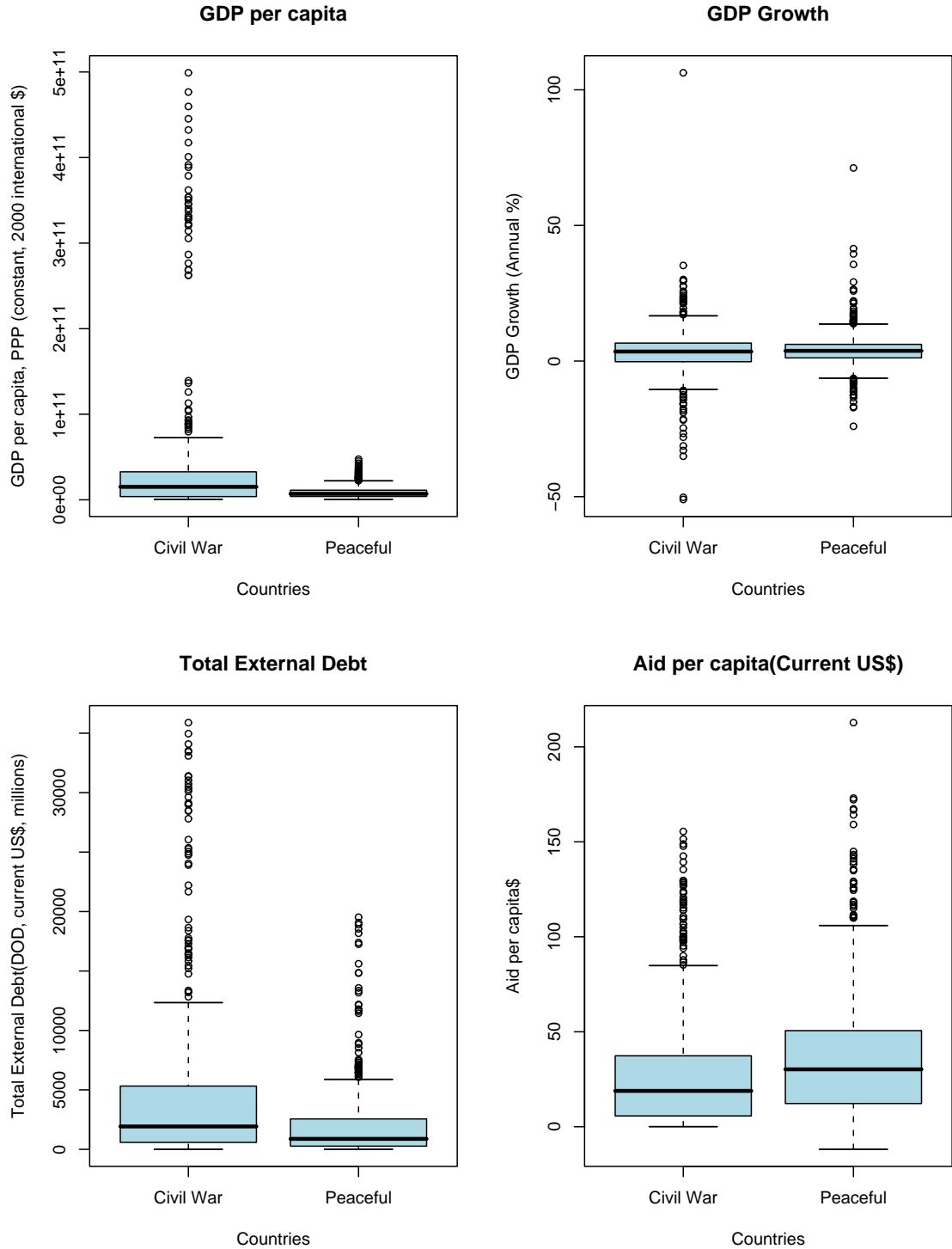


Figure 3: Comparison of possible explanatory factors for onset of civil war by civil war and non-civil war countries in Sub-Saharan Africa, 1960-2002. The thick bar represents the median. The top and bottom of the box represent the first and third quartiles.



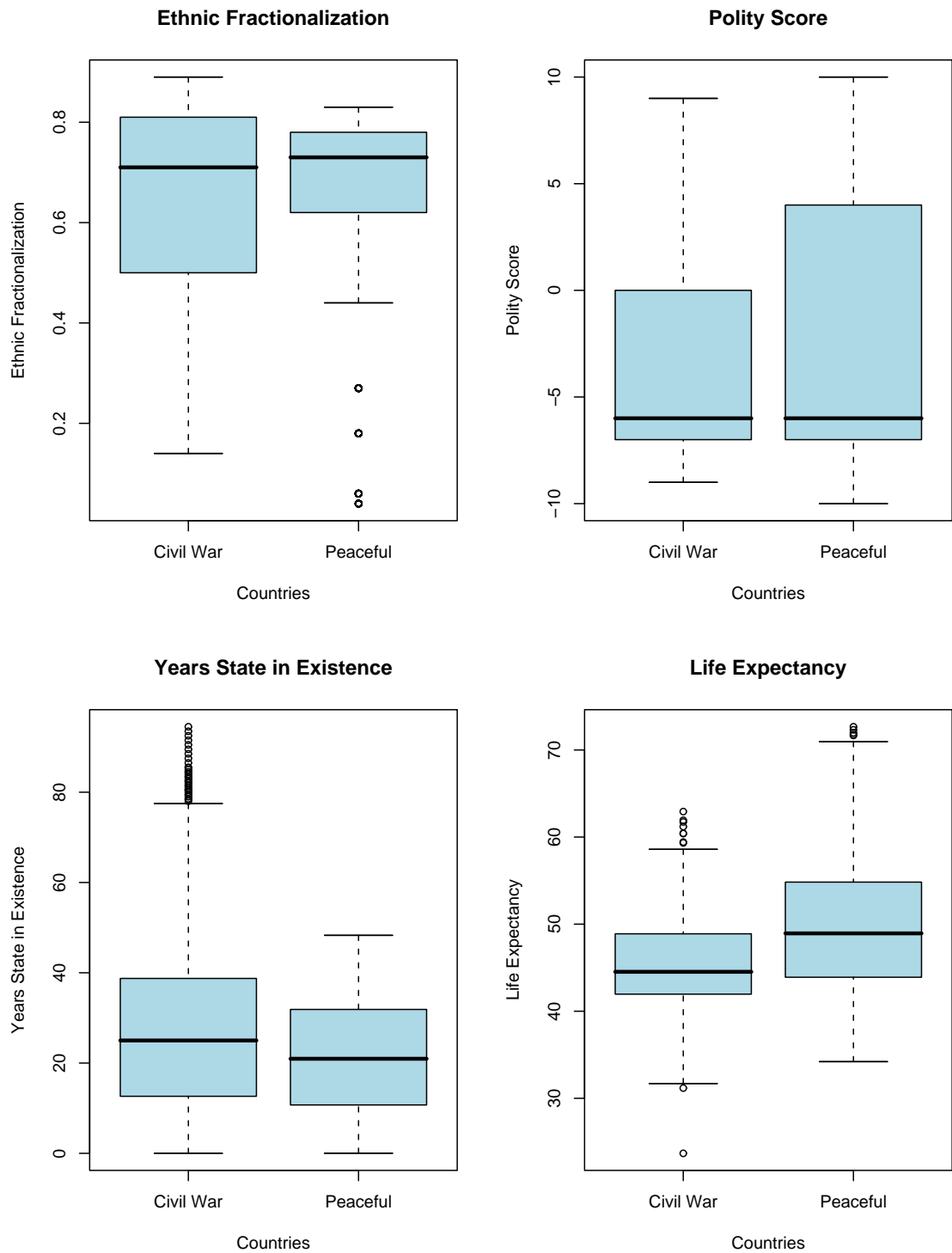


Figure 4: Comparison of possible explanatory factors for onset of civil war by civil war and non-civil war countries in Sub-Saharan Africa, 1960-2002. The thick bar represents the median. The top and bottom of the box represent the first and third quartiles.

the onset of civil war: total population size, GDP growth, external debt, medium polity score, state age, and the decade 1990's (with 1960's as the baseline). For total population size, the odds increased by 1.02 times with every million increase in population size. Population growth was significant at the .10 level. The percent of youth in adult population did not show any significance. For GDP growth, the odds ratio was less than one, indicating a negative relationship. A million dollar increase in external debt slightly increased the odds of civil war incidence. Foreign aid per capita was significant at the .10 level and showed a negative association. State age was positively associated, as was medium polity and the decade 1990's.

Table 1. Bivariate Logistic Regression Results on Civil War Onset, 1960-2002, Pooled Analysis

Variable	Odds Ratio	Std. Error
GDP per capita	1.0000	0.0001
GDP growth(%)*	0.9200	0.0159
External debt(mil)*	1.0001	0.0000
Foreign aid per capita $\diamond$	0.9909	0.0055
Total population(mil)*	1.0187	0.0059
Population growth(%) $\diamond$	1.1206	0.1115
Population density	1.0021	0.0012
% Youth in adult pop.	3.5572	6.0995
State age*	1.0001	0.0000
Medium polity*	2.1635	0.2756
High polity	0.1374	1.0207
Ethnic fractionalization	0.9991	0.0062
Decade 1970	1.2680	0.4819
Decade 1980	1.1027	0.4896
Decade 1990*	2.5184	0.4345
Decade 2000	0.3203	1.0743

\*p-value <0.05

$\diamond$ p-value <.10

## 4.2 Multivariate analysis

Four different models were examined for the multivariate analysis. The results are shown in Table 2. The first model included all twelve independent

variables that were explored in the bivariate analysis. Under this model, the only factors that remained significant at  $p\text{-value} < 0.05$  were GDP per capita and GDP growth. The odds of starting a civil war decreases slightly with a dollar increase in GDP per capita and a percent increase in GDP growth.

One problem with using these economic factors to model the onset of civil war is that the direction of causality is unclear. An increase in GDP can reduce the onset of civil war but civil war can also increase or decrease the GDP. It can be increased through increased military spending or decreased if the economy becomes crippled as a result. Therefore, in model 1L, I lagged the economic indicators in model 1 to three years back. When lagged variables were used, only the GDP per capita remained significant. The odds ratio was less than one, indicating that a dollar increase in GDP slightly decreased the odds of starting a civil war.

In the second model, external debt, foreign aid per capita, and state age were removed from the model. These three variables were highly correlated with many of the other variables. After these variables were removed, GDP growth and total population showed significant results. With every one million increases in population, the odds of a start of a civil war are 1.02 times greater. GDP per capita was no longer significant. GDP growth showed a significant negative association.

In model 2L, lagged economic variables were used for GDP and GDP growth, plus all of the other variables in model 2 were included. Using the lagged variables, total population was the only variable that remained significant. The odds of starting a civil war was 1.017 times greater with a million increase in population.

## 5 Results of fixed effects regression analysis

The pooled results may be biased since repeated observations of a country are not independent of each other. Therefore, further analysis was conducted using country-fixed effects. The results are shown in Table 3.

When country-fixed effects were used on Model 1, only GDP growth was statistically significant. The odds ratio went up slightly to .9393. Nothing was significant in the lagged version of Model 1 (Model 1L). With Model 2, GDP growth is still significant. However, when lagged variables were used in Model 2L, nothing was significant.

## 6 Limitations

There are many limitations to this analysis. Omitted variable bias is one of them. Many factors were not included in the models because of the data were not easily available, such as poverty and income inequality. In addition, there are other variables, such as prevalence of HIV/AIDS that may be correlated to many of the variables in the model, such as foreign aid per capita. I also did not control for history of recent conflicts.

Another issue is that of reverse causality. I attempted to address this through the use of lagged variables in the pooled analysis. Using the longitudinal design of fixed effects may eliminate some of the problems with unobserved heterogeneity.

There may also be interaction occurring in the variables which was not explored at this time. It may also be more informative to look at prevalence of conflicts in addition to incidence of violence.

## 7 Conclusion

What role does demography play in the onset of civil war? Two of the three models in the pooled analysis showed that population size was an important factor. However, because of country specific variability as was seen in the descriptive analysis, the country fixed effects models may be more relevant. Of these models, model 2L, with lagged GDP may be the most appropriate, since it addresses the issue of reverse causality and excludes correlated variables. However, this model did not show any significant results. Surprisingly, the percent youth among the adult population, which was an indicator for a youth bulge, did not show significance in any of the models. With population size increasing in many Sub-Saharan countries, a closer examination of their demography may provide some clues to the political instability of the region. Further research is needed to better understand the roles demography plays in the onset of civil wars.

Table 2. Multivariate Logistic Regression Results on Civil War Onset, 1960-2002, Pooled Analysis

Variable	Model 1		Model 1L		Model 2		Model 2L	
	Odds Ratio	Robust SE	Odds Ratio	Robust SE	Odds Ratio	Robust SE	Odds Ratio	Robust SE
GDP per capita	0.9991*	0.0003			0.9998	0.0003		
GDP per capita(lag 3 yrs)			0.9991*	0.0003			0.9998	0.0003
GDP growth(%)	0.9208*	0.0370			0.9167*	0.0293		
GDP growth(% , lag 3 yrs)			0.9821	0.0288			0.9740	0.0233
External debt(mil)	1.0002	0.0001	1.0001	0.0001				
Foreign aid per capita	0.9852	0.0094	0.9843	0.0152				
Total population(mil)	0.9701	0.0250	0.9724	0.0257	1.0198*	0.0063	1.0174*	0.0051
Population growth(%)	1.1954	0.1860	1.0062	0.0856	1.2541	0.1671	1.0334	0.0920
Population density	1.0066	0.0040	1.0062	0.0037	1.0053	0.0034	1.0047	0.0037
Youth/adult ratio	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
State age	0.9999	0.0001	0.9999	0.0001				
Low polity	1.3009	1.3404	1.6255	2.0248	5.4148	10.0126	9.3163	23.5147
Medium polity	2.9827	2.4735	3.9127	4.0499	11.1822	19.9138	15.9738	38.9772
Ethnic fractionalization	0.9864	0.0210	0.9898	0.0172	0.9933	0.0139	0.9938	0.0132

\*p-value <0.05

Table 3. Multivariate Logistic Regression Results on Civil War Onset, 1960-2002, Country Fixed Effects

Variable	Model 1		Model 1L		Model 2		Model Model 2L	
	Odds Ratio	SE	Odds Ratio	SE	Odds Ratio	SE	Odds Ratio	SE
GDP per capita	0.9993	0.0014			0.9996	0.0012		
GDP per capita(lag 3 yrs)			1.0005	0.0014			1.0000	0.0011
GDP growth(%)	0.9393*	0.0264			0.9420*	0.0262		
GDP growth(%, lag 3 yrs)			0.9865	0.0251			0.9878	0.0248
External debt(mil)	1.0001	0.0001	1.0001	0.0001				
Foreign aid per capita	1.0100	0.0143	1.0109	0.0132				
Total population(mil)	0.9023	0.0674	0.8927	0.0705	0.9543	0.0371	0.9391	0.0380
Population growth(%)	1.2651	0.1910	1.0787	0.1384	1.1969	0.1754	1.0387	0.1315
Population density	0.9988	0.0145	1.0005	0.0136	1.0026	0.0118	1.0004	0.0113
Youth/adult ratio	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
State age	1.0001	0.0002	1.0000	0.0002				
Low polity	1.7525	2.5714	1.7525	2.5229	3.7722	4.8767	3.3550	4.2957
Medium polity	3.5423	5.2230	4.0979	5.9817	9.0775	11.6705	8.6903	11.2785
Ethnic fractionalization	822037.4	1260000000			1096876	1500000000		

\*p-value <0.05

Note: ethnic fractionalization was omitted in model 3 due to lack of within-group variance.

## 8 References

Collier, P., and A. Hoeffler (2002) On the incidence of civil war in Africa. *The Journal of Conflict Resolution*46(1): 13-28.

Collier, P., and A. Hoeffler (2001) *Greed and grievance in civil war* World Bank Working Paper. <http://go.worldbank.org/GKOR671H90>

Elbadawi, I., and N. Sambanis (2002) How much war will we see? Explaining the prevalence of civil war. *The Journal of Conflict Resolution*46(3): 307-334.

Fearon, J.D. and D.D. Laitlin (2003) Ethnicity, insurgency, and civil war. *The American Political Science Review* 97(1): 75-90.

Krain, M. (1997) State-sponsored mass murder: the onset and severity of genocides and politicides. *The Journal of Conflict Resolution*41(3): 331-360.

Urdal, H (2006) A Clash of Generations? Youth Bulges and Political Violence. *International Studies Quarterly* 50: 607-629.