Statistical analysis of the causes of civil war:
A plea for the recognition of causal complexity

Sophia du Plessis, Ada Jansen and Krige Siebrits
(Department of Economics, University of Stellenbosch)

INTRODUCTION

Academic research on the causes of civil wars has proliferated during the past two decades. One of the main reasons for this growth in interest is the high incidence of such conflicts. Data on civil conflicts (internal conflicts with 25 or more battle-related deaths in a single year) and civil wars (internal conflicts with one thousand or more battle-related deaths in a single year) reported by Blattman and Miguel (2010: 3-4) starkly portray the scope of the phenomena. More than half of all countries experienced at least one episode of civil conflict in the period from 1960 to 2006, while one third were afflicted by civil wars. Fully 20 percent of all nations experienced ten or more years of civil conflict. There has been a welcome reduction in violent internal conflict in recent years – the number of countries experiencing civil conflict peaked at 51 in 1992 and decreased thereafter to 32 in 2006 – but the phenomenon remains sufficiently widespread and destructive in social, political and economic terms to warrant further study.[1]

The ongoing research programme has produced various theoretical explanations for civil war as well as a large number of attempts to test these theories empirically. It has been a productive endeavour that has yielded many useful insights. Yet the adoption of a wider range of methods may well be a requirement for sustaining the momentum: the multiple regression techniques that have hitherto dominated the empirical part of the research programme are far from ideal for explaining the complexity of the many civil wars with several interlinked causes. The purpose of this paper is to draw attention to the implications for the study of civil war of this feature of regression analysis.

The remainder of the paper consists of two sections, followed by concluding comments. The first part provides a brief review of theoretical and empirical research on the causes of civil wars and explains the difficulty of dealing with the multiple causes of many civil wars in regression analyses. The second reports the findings of a simple set-theoretic analysis of the causes of 23 civil wars in African countries. (The main reasons for focusing on Africa are the relatively high incidence of civil war on the continent[2] and the urgent need for political stability – a necessary condition for accelerated economic growth.) The analysis shows that more than one potential cause were present at the onset of the vast majority of these wars. This suggests the possibility that combinations of factors may have given rise to many of these conflicts.

---

[1] It is obviously very difficult to quantify the costs of civil wars. Skaperdas (2011) provides a useful discussion of the technical issues and review of the available evidence.

[2] Besley and Reynal-Querol (2012: 2) point out that the percentages of all country years since 1950 that constituted conflict years (i.e. years in which battle-related deaths exceeded 1 000) were 8.5 percent for African countries and 5 percent for other countries.

REVIEW OF THEORETICAL AND REGRESSION-BASED EMPIRICAL RESEARCH

In a wide-ranging survey paper, Blattman and Miguel (2010: 9-17) distinguish three sets of explanations for the outbreak of civil wars:

- One strand of the literature models armed conflict as the result of competition for resources such as government revenue, mineral rents and foreign aid. This approach builds on economic models of contests.
- A second approach attempts to explain why rational actors sometimes engage in armed conflict despite its potential costliness and destructiveness. Two sets of reasons have been offered why conflict situations are not necessarily solved by means of bargaining. Some theories ascribe engagement in destructive conflict by rational actors to information asymmetries: parties may have private information regarding their military capabilities and misrepresent such information for strategic purposes. The alternative argument is that constraints on contract enforcement and conflict resolution give rise to commitment problems, which make settlements between warring parties incomplete contracts.
- A third approach to the causes of civil war studies the formation and endurance of groups involved in such conflicts, as well as the mechanisms used to motivate members to take part in dangerous activities. Contributions on participation have identified various mechanisms for solving the collective-action problem such as financial and material incentives, grievances caused by extreme inequality and unfulfilled expectations of economic advancement, and coercion. Models of the formation and coherence of competing coalitions emphasise intra-group conflict over the distribution of the surplus and the importance of property-rights norms and other mechanisms for mitigating such conflict and ensuring group stability. Given the relatively high prevalence of ethnically based rebellions against states, many such models emphasise factors related to ethnic mobilisation and conflict.

Differences in the scope for measuring key concepts and the availability of relevant data have led to marked variance in the extent to which each of these sets of theoretical explanations has been subjected to empirical testing. Nonetheless, such testing has identified a long list of correlates of civil war: negative income shocks, low per capita incomes, weak state institutions, large populations, sparsely populated peripheral areas, strong dependence on natural resources, mountainous terrain, war-prone neighbours, recent political instability, and small government military forces (cf. Blattman and Miguel, 2010: 22-24). By contrast, high levels of income inequality and ethno-linguistic fractionalisation are generally not statistically significant predictors of civil conflict.

Much could be said about the specification and measurement of variables in these studies, as well as the strategies for identifying relationships between the incidence of civil war and the various causal variables (cf. Sambanis, 2004). The issue highlighted in this paper is that the multiple regression techniques used to identify these correlates rest on strong assumptions regarding the nature of independent variables. Ragin (2008: 112) summarises these assumptions as follows:

> Typically, each causal variable is thought to have an autonomous or independent capacity to influence the level, intensity, or probability of the dependent variable. Most applications of conventional quantitative methods assume that the effects of the independent variables are both linear and additive… the impact of a given

independent variable is the same not only across the value of the other independent variables, but also across their different combinations.

These assumptions manifest in the empirical literature on the origins of civil war in a tendency to develop explanations based on dominant causes: regression analysis leads researchers to develop and test competing explanations of civil war, instead of exploring possible interactions.[3] Collier and Hoeffler's (2004) emphasis on motivational factors (divided into "grievance factors" and "greed factors") is a clear example, as is that of Fearon and Laitin (2003) on feasibility factors. Yet it is increasingly acknowledged that researchers should take the possibility that complex combinations of factors cause civil wars more seriously. Sambanis's (2004) critique of aspects of statistical studies of civil war, for example, includes several references to the value of case-study methods for exploring interactions between causal factors. Similarly, Blattman and Miguel (2010: 31) suggest that contextual factors may influence the links between civil war and some causal variables and argue that theoretical modelling and empirical testing should pay close attention to interaction effects between variables.

PRELIMINARY EVIDENCE FROM CIVIL WARS IN AFRICA

This section tentatively reinforces the argument that the causes of civil war should be studied as combinations by presenting the preliminary findings of a simple set-theoretic analysis of African countries. (The main reasons for focusing on Africa are the relatively high incidence of civil war on the continent[4] and the urgent need for political stability – a necessary condition for accelerated economic growth.) For each African country, each five-year period from 1970 to 2014 is defined as a case marked by civil war or the absence thereof as well as the presence or absence of certain causal factors. Scores of zero or one are assigned for each variable or factor for each case: a score of zero indicates that the variable or factor was either absent from the case or too feeble to have been a likely cause of the civil war, while variables that received scores of one are deemed to have been likely causes of wars.[5] The results of the analysis show all the combinations of causal factors present in the various cases and the frequencies of their occurrence; this makes it possible to draw preliminary inferences regarding the importance of studying the causal factors in combination.

The tentative nature of the analysis and findings should be stressed at the outset. The chosen causal variables are the same ones that yield disparate results in regression analysis and some of the data used to quantify these variables may be of dubious quality. Furthermore, the methods used to assign scores to the various causal factors are inevitably subjective (although much effort went into the choosing of criteria and their

[3] It should be noted that multiplicative interaction terms make it possible to capture the effects of combinations of causal variables in regression analyses (cf. Scruggs, 2007: 312-316). Technical factors, however, limit the scope for handling causal complexity by these means. Interaction terms can be difficult to interpret and often give rise to collinearity and degrees-of-freedom problems, especially when several (including higher-order ones) are included in models estimated with modest-sized samples (Brambor, Clark and Golder, 2005; Ragin, 1987: 15, 65-66; 2008: 9, 113).

[4] Besley and Reynal-Querol (2012: 2) point out that the percentages of all country years since 1950 that constituted conflict years (i.e. years in which battle-related deaths exceeded 1 000) were 8.5 percent for African countries and 5 percent for other countries.

[5] The method is based on elements of crisp-set qualitative comparative analysis – a set-theoretic method developed by Ragin (1987) and later augmented by the more versatile fuzzy-set qualitative comparative analysis (cf. Ragin, 2000).

consistent application). In addition, the validity of the findings has not been established by means of detailed study of the individual cases. The purpose of the exercise is not to provide an accurate explanation of civil war in Africa, but to demonstrate that the causal factors used in most regression analyses often appear in combinations and in such cases may exert their influence jointly rather than individually.

Data and variables

We use the *Major Episodes of Political Violence (MEPV) and conflict regions, 1946-2015* database (Marshall, 2016) to identify African countries that experienced civil and ethnic wars in the period from 1970 to 2014. To facilitate comparisons, we also compile a list of African countries not afflicted by internal conflict. Table 1 lists the two groups.

Table 1

Civil wars in Africa (1970-2014) and the control-group countries

Civil and ethnic wars (years in which wars broke out in brackets):
Algeria (1991); Angola (1975); Burundi (1972; 1993); Central African Republic (2006); Chad (2006); Congo (1997); Cote d'Ivoire (2000); Democratic Republic of Congo (1977; 1996); Ethiopia (1974); Guinea-Bissau (1998); Liberia (1985; 1990; 2001); Mauritania (1975); Morocco (1975); Mozambique (1981); Nigeria (2009); Rwanda (1990; 2001); Sierra Leone (2001); Somalia (1988); Sudan (1983); Uganda (1971; 1981).
Countries that did not experience civil or ethnic wars:
Benin; Botswana; Burkina Faso; Cameroon; Cape Verde; Comoros; Djibouti; Egypt; Equatorial Guinea; Eritrea; Gabon; Gambia; Ghana; Guinea; Kenya; Lesotho; Libyan Arab Jamahiriya; Madagascar; Malawi; Mali; Mauritius; Namibia; Niger; Sao Tome and Principe; Senegal; Seychelles; South Africa, Swaziland, Tanzania, Togo, Tunisia, Zambia, Zimbabwe.

Source: Marshall (2016)

The analysis revolves around four sets of factors that could possibly explain the onset of armed conflicts: economic grievances, political grievances, greed and the feasibility of initiating a civil war. Each factor consists of one or more variables often used in studies of the causes of civil war (cf. Blattman and Miguel, 2010: 22-24). The remainder of this section discusses each set of factors separately, focusing on their constituent variables, the methods used to assign scores to each variable, and the aggregation of these scores. The data sources are listed in Appendix Table 1.

Economic grievances

The possibility that poor economic performance could cause unhappiness that manifests in armed conflict is captured by two variables: the growth rate of per capita household consumption expenditure at constant 2005 prices (a measure of income growth), and life expectancy at birth of males in years (an indicator of living standards).

Political grievances

Suppression of the political rights of a population (or parts thereof) can also give rise to armed conflict. In our dataset, the likelihood of serious political grievances is proxied by

the dichotomous coding of democracy in the Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2010 Dataverse (cf. Boix et al, 2014). The data provides a dichotomous coding of democracy (i.e. 1 if the country is democratic, 0 if otherwise).

Greed

We identified two factors to capture the presence of greed-related causes of civil war: access to government resources (proxied by general government final consumption expenditure as a percentage of GDP), and the availability of lucrative natural-resource rents (proxied by the value of mining and utilities expressed as a percentage of GDP).

Feasibility

This factor combines certain aspects that are often regarded as determinants of the feasibility of armed rebellion against governments.

Demographic elements

- Relatively high number of potential combatants: we assume that the likelihood of conflict is increased if the male population in the age group 15-49 is a high proportion of the total male population.

Topography

- Mountainous terrain: we assume that rebel activity is facilitated the land area of a country is mountainous.

We code each of the possible contributing factors by considering whether or not there was a change in the level of the variable, or whether or not a negative shock occurred. The former is defined as a change in the level of the variable a year before the beginning of the five-year period, or a change in the level of the variable during the previous five-year period. A shock is defined as a negative change in the variable a year before the beginning of the five-year period, or a negative change in the variable during the previous five-year period.

RESULTS

This section commences with a brief comparative analysis of how the explanatory factors differ between countries that experienced a civil war and those who did not during the period 1974 to 2010 (note we use this time period as we are restricted by data availability). Table 2 provides a description of how the average values for the explanatory factors were determined, whilst Table 3 provides reports a crude score of these values between the two groups of countries.

Table 2

Calculation of average scores of explanatory factors

Variable	Description
Low levels of consumption	Average ranking of civil war countries before civil war vs average ranking of non-civil war countries
Negative consumption shock	Average growth rate of civil war countries before the first civil war vs average growth rate of non-civil war countries
Low life expectancy	Average life expectancy in civil war countries before the first civil war vs average life expectancy in non-civil war countries
Males aged 15-49	Average percentage males in age group 15-49 in civil war countries before the first civil war vs average percentage males in age group 15-49 in non-civil war countries
Mountainous terrain	Average for civil war vs non-civil war countries
General government expenditure	Average general government expenditure as percentage of GDP in civil war countries before the first civil war vs the same in non-civil war countries
Mining and utilities	Mining and utilities as a percentage of GDP in civil war countries before the first civil war vs the same in non-civil war countries
Political grievance (democracy)	We use the binary score from Boix, et al (2014). "1" indicates democracy and "0" autocracy. Our score here indicates the average number of years for the civil war countries with democracy before a civil war vs the average number of years for the non-civil war countries with democracy

Table 3
Average scores of explanatory factors

Variable	Score	
	Civil-war countries	**Non-civil war countries**
Low levels of consumption	27.92	24.38
Negative consumption shock	0.6%	1.8%
Low life expectancy	45.2	50
Males aged 15-49	44.3%	44.1%
Mountainous terrain	14	19

General government expenditure	14.5%	14.8%
Mining and utilities	13.7	11.3
Political grievance (democracy)	5.5	19.7

The results in Table 3 shows how the variables differ between civil war and non-civil war countries. These crude scores provide us with some guidance on which of these variables should be considered in the subsequent analysis where we apply the crisp fuzzy set technique. If the variables show no statistical significant difference between civil war and non-civil war countries, we will not use the variable in our fuzzy set.

To facilitate the crisp fuzzy set analysis, Table 4 shows the coding we have used to assign binary scores:

Table 4

Variable	Score	
	Year before the five-year period	**During the previous five-year period**
Low levels of consumption	"1' indicates that the country's level of real household consumption per capita was among the lowest third of the 51 countries in the year immediately before the beginning of that five-year period. "0' indicates that the country's level of real household consumption per capita was among the highest two-thirds of the 51 countries in the year immediately before the beginning of that five-year period.	"1' indicates that the country's level of real household consumption per capita was among the lowest third of the 51 countries in each of the five years of the previous five-year period. "0' indicates that the country's level of real household consumption per capita was among the highest two-thirds of the 51 countries in each of the five years of the previous five-year period.
Negative consumption shock	"1' indicates that the country's level of real household consumption per capita decreased in the year immediately before the beginning of that five-year period.	"1' indicates that the country's level of real household consumption per capita decreased during the previous five-year period. "0' indicates that the country's level of real

	"0' indicates that the country's level of real household consumption per capita did not decrease in the year immediately before the beginning of that five-year period.	household consumption per capita did not decrease during the previous five-year period.
Low life expectancy	"1' indicates that the country's average life expectancy for males and females was among the lowest third of the 51 countries in the year immediately before the beginning of that five-year period. "0' indicates that the country's average life expectancy for males and females was among the highest two-thirds of the 51 countries in the year immediately before the beginning of that five-year period.	"1' indicates that the country's average life expectancy for males and females was among the lowest third of the 51 countries in each of the five years of the previous five-year period. "0' indicates that the country's average life expectancy for males and females was among the highest two-thirds of the 51 countries in each of the five years of the previous five-year period.
Males aged 15-49	"1' indicates that the country's percentage of males in the age group 15-49 increased in the year immediately before the beginning of that five-year period. "0' indicates that the country's percentage of males in the age group 15-49 increased in the year immediately before the beginning of that five-year period.	"1' indicates that the country's average life expectancy for males and females decreased during the previous five-year period. "0' indicates that the country's average life expectancy for males and females did not decrease during the previous five-year period.
Mountainous terrain	Average for civil war vs non-civil war countries	
General government expenditure	"1' indicates that the general government expenditure as percentage of GDP decreased in the year immediately before the beginning of that five-year period.	"1' indicates that the general government expenditure as percentage of GDP decreased during the previous five-year period.

	"0' indicates that it did not decrease in the year immediately before the beginning of that five-year period.	"0' indicates that it did not decrease during the previous five-year period.
Mining and utilities	"1' indicates that mining and utilities increased in the year immediately before the beginning of that five-year period. "0' indicates that mining and utilities did not increase in the year immediately before the beginning of that five-year period.	"1' indicates that Mining and Utilities increased during the previous five-year period. "0' indicates that Mining and Utilities did not increase during the previous five-year period.
Political grievance (democracy)	We use the binary score from Boix, et al (2014). "1" indicates democracy and "0" autocracy. "1' indicates that the country was autocratic in the year immediately before the beginning of that five-year period. "0' indicates that the country was not autocratic in the year immediately before the beginning of that five-year period.	

CONCLUSION

The generation of parsimonious, generalizable explanations is a worthy pursuit in scientific research. Such explanations, however, become problematic when they derive from methods that do not fully capture the complexity of real-world phenomena. This paper argues that exclusive reliance on multiple regression analysis could hamper social scientists in their quest to unravel the complex causes of civil wars, because such methods are not designed to study the effects of combinations of interlinked causes. Accordingly, it suggests that deeper understanding of civil wars may require the development or application of alternative techniques better suited to the analysis of such complex causal relationships.

REFERENCES

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003): Fractionalization. *Journal of Economic Growth*, 8: 155-194.

Besley, T. and M. Reynol-Querol (2012): The legacy of historical conflict: evidence from Africa. *STICERD - Economic Organisation and Public Policy Discussion Papers Series 036*. London: London School of Economics (Suntory and Toyota International Centres for Economics and Related Disciplines).

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

Boix, C., M. Miller, K. Michael and S. Rosato (2014): Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2010, doi:10.7910/DVN/28468, Harvard Dataverse, V1, UNF:5:x3P4QDm6R349bpyygm/g1Q==.

Brambor, T., W.R. Clark and M. Golder (2005): Understanding interaction models: improving empirical analyses. *Political Analysis*, 13: 1–20.

Collier, P. and A. Hoeffler (2004): Greed and grievance in civil war. *Oxford Economic Papers*, 56: 563-595.

Collier, P. and A. Hoeffler (2009): Beyond greed and grievance: feasibility and civil war. *Oxford Economic Papers*, 61: 1-27.

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

Greig, M. and A. Enterline (2012): *National material capabilities data*. Correlates of War Project. Denton, TX: University of North Texas. Available at: http://dvn.iq.harvard.edu/dvn/dv/cow.

Marshall. M.G., K. Jaggers and T.R Gurr (2011): Polity IV: Political regime characteristics and transitions, 1800-2010. Online database. Center for Systemic Peace. Available at: www.systemicpeace.org/polity/polity4.htm.

Marshall, M.G. (2016): Major episodes of political violence, 1946-2015. Online database. Vienna, VA: Centre for Systemic Peace. Available at: http://www.systemicpeace.org/warlist.htm.

Organisation for Economic Co-operation and Development (2013): *Total official flows*. Paris: Organisation for Economic Co-operation and Development. Available at: http://stats.oecd.org/Index.aspx?DataSetCode=REF_TOTALOFFICIAL

Ragin, C.C. (1987): *The comparative method: moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.

Ragin, C.C. (200o): *Fuzzy-set social science.* Chicago: University of Chicago Press

Ragin, C.C. (2008): *Redesigning social inquiry: fuzzy sets and beyond*. Chicago: University of Chicago Press.

Sambanis, N. (2004): Using case studies to expand economic models of civil war. *Perspectives on Politics*, 2: 259-279.

Scruggs, L. (2007): What's multiple regression got to do with it? *Comparative Social Research*, 24: 309-323.

Skaperdas, S. (2011): The costs of organized violence: a review of the evidence. *Economics of Governance*, 12: 1–23.

United Nations Organisation (2011): *World population prospects: the 2010 revision*. New York: United Nations Organisation (Department of Economic and Social Affairs, Population Division). Available at: http://esa.un.org/wpp/.

United Nations Organisation (2013): *National accounts: estimates of main aggregates*. New York: United Nations Organisation (Statistics Division). Available at: http://unstats.un.org/unsd/snaama/dnlList.asp.

World Bank (2013): *World development indicators 2012*. Washington, D.C.: The World Bank. Available at: http://data.worldbank.org/data-catalog/world-development-indicators

Appendix Table 1

Data sources

Indicator	Source
Economic grievance factors:	
Per capita GNI at current prices (US dollars)	United Nations Organisation (2013)
Growth rate of per capita household consumption expenditure at constant 2005 prices (percent)	United Nations Organisation (2013)
Life expectancy at birth of males (years)	World Bank (2013)
Political grievance factors:	
Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2010 Dataverse	Boix, et al (2014)
Greed factors:	
Total natural resources rents (percent of GDP)	World Bank (2013)
General government final consumption expenditure (percent of GDP)	United Nations Organisation (2013)
Foreign aid (percent of GDP)	Organisation for Economic Co-operation and Development (2013)
Feasibility factors: Demographic and social	
Population density (persons per square kilometre)	World Bank (2013)
Population share of males aged 15-49 (percent of total population)	United Nations Organisation (2011)
Ethnic, linguistic and religious fractionalisation indices	Alesina, Devleeschauwer, Easterly, Kurlat and Wacziarg (2003)
Feasibility factors: Military power	
Military personnel (percent of males aged 15-49)	Greig and Enterline (2012)
Government military spending (percent of GDP)	Greig and Enterline (2012)
Feasibility factors: Topography	
Mountainous terrain (percent of land area)	Fearon and Laitin (2003)