Ethnic Fractionalization and Sub-Saharan Violence, 1970-1996

by

Josiah Seale

Submitted to the Department of Political Science
in Partial Fulfillment of the Requirements for the Degrees of

Bachelor of Science of Political Science

and

Master of Science of Political Science

at the
Massachusetts Institute of Technology
June, 2006

Signature of Author .....
Department of Political Science
June 7, 2006

Certified by .....
Roger D. Petersen
Associate Professor of Political Science
Thesis Supervisor

Certified by .....
J. Chappell H. Lawson
Associate Professor of Political Science
Second Reader

Accepted by .....
Roger D. Petersen
Associate Professor of Political Science
Chair, Department Graduate Program Committee

ABSTRACT:

This study examines the statistical correlations between metrics of ethnic fractionalization and categories of violence in Sub-Saharan Africa from 1970 to 1995. By examining these correlations both prior to and after controlling for income, the study is able to determine whether or not various types of conflict are linked to patterns of ethnic grouping. The study uses newer, more refined measures to evaluate the correlations between specific categories of violence and specific measures of ethnic fractionalization. Using simple and multivariate linear regressions, the study examines each of the correlations between a total of twenty-two sub-metrics of four categories of violence, per capita income and metrics of ethnic fractionalization on three tiers. This allows the study to gauge the impacts (both separately and in interaction) of dichotomous top-tier cleavages in deeply divided societies, general ethnic fractionalization and nested ethnic sub-grouping.

The study finds that the majority of the categories of violence used are not correlated with ethnic fractionalization, neither prior to nor after controlling for income. However, successful coups are highly correlated with the presence and distribution of the dichotomous top-tier cleavages in deeply divided societies, with the frequency of these successful coups increasing linearly as the divide approaches a 50/50 split. The frequency of riots is robustly correlated with general ethnic fractionalization, but that ethnic fractionalization has much less predictive power for this correlation.

An implication of these findings for future research is that using these refined definitions allows for a fuller understanding of the behavior being examined: all categories of conflict are not uniformly correlated with all measures of ethnic fractionalization, and vice versa. Hypotheses regarding these correlations must thus enter into detail, both about the claims being made and the groups for which they are made. Additionally, the findings show that income's correlation with conflict and ethnic fractionalization is more complex than has been indicated in much of political science literature, and that further research on the topic is merited.

1. Introduction
Emerging Conventional Wisdom Re-Examined

In *The Oxford Handbook of Political Economy,* Fearon (2006) makes the following statement: "Once one controls for per capita income, neither civil wars nor ethnic civil wars are significantly more frequent in more ethnically diverse countries; nor are they more likely when there is an ethnic majority and a large ethnic minority."

This point of view, which denies any meaningful relationship between civil war and ethnic fractionalization, is having a formative influence on current conventional wisdom in regard to ethnic conflict. It is not certain, however, that this extrapolation from civil war to conflict in general is justified. Does this understanding hold true for other types of violence besides full-blown civil wars? Does it hold true for some ethnic fractionalization patterns but not for others?

Political science literature remains divided about the importance of ethnic fractionalization with regards to its correlation with conflict. Moreover, there have been few studies that quantitatively evaluate the possibility of specific patterns of ethnic fractionalization being correlated with specific types of violence. This study uses newer, more refined measures of violence to evaluate the correlations between specific categories of violence and specific measures of ethnic fractionalization. By doing this, the study can examine the applicability of Fearon's civil war findings in light of a broader understanding of conflict as it correlates to a broader array of fractionalization patterns.

The project addresses the question in a Sub-Saharan context, using data from 1970 to 1995. The study's work is in two parts. Part one examines the correlations between violence and ethnic fractionalization prior to controlling for per capita income. Part two examines the correlations of violence after controlling for per capita income. For each of these parts the project carries out a

multi-step examination of the combinations of possible correlations between twenty-two sub-metrics of four major types of violence at varying levels and 6 ethnic fractionalization patterns. The first part examines the correlation between these metrics of violence and the fractionalization indices and then addresses the problem of outliers in the data set. The second part examines the same correlations while controlling for real GDP per capita.

In the first two sections of this paper I introduce the project and its goals, and then examine literature on ethnic fractionalization and describe its use in the study. In the third section I discuss the types of conflict being evaluated and the metrics used to gauge them in this study. In the fourth portion I discuss the correlations I found in the data, both with and without outliers and with and without controlling for income. In the fifth and final segment I discuss the contributions of the study's findings, along with the implications of the study for further research and its possible limitations.

II. Ethnic Fractionalization

1. Basic Theories

Political science literature remains divided on the topic of ethnic fractionalization and its correlation with conflict. This division stems from a disagreement not only about the importance of ethnic fractionalization in conflict but also about ethnic conflict's true meaning and underlying causes. What is actually happening when an ethnic group arms itself and engages in violent conflict? Is it a group of people who feel a special kinship to one another, banding together to protect deeply felt common values? Is it the result of a disingenuous fabrication that has been knowingly and calculatingly crafted by self-interested elites or even by the combatants themselves? Or is it (perhaps more likely) some combination of these two that leads to a net result falling somewhere on the spectrum between them?

Several basic theories exist as to ethnic fractionalization's implications for violence. As part

of these theories, the bulk of the political science literature on ethnic conflict seems to have taken the terms "ethnic conflict" and "ethnic civil war" to be synonymous. On the one side are those who find that ethnic conflict is generally economically motivated and that ethnic fractionalization has little effect, if any at all. On this side would fall the findings of Fearon & Laitin (2000). In this study, the researchers find that ethnic fractionalization "appear[s] unrelated to either the probability or the level of civil violence in our sample, once we control for income and income growth." (p.26) Fearon (2006) continues along this tack in the quote given at the beginning of the introduction, stating that "[o]nce one controls for per capita income, neither civil wars nor ethnic civil wars are significantly more frequent in more ethnically diverse countries; nor are they more likely when there is an ethnic majority and a large ethnic minority." (p.5 in 2004 Version)

Collier & Hoeffler (2002) go so far as to state that in their analysis, ethnic fractionalization (and especially Africa's ethnic fractionalization) should impede violence. According to this logic, with its high degree of ethnic fractionalization, "the difficulties of coordinating rebellion should be greater in Africa and should make the region safer, ceteris paribus." (p.18) Like Fearon & Laitin, they also examine the specific issue of a fractionalization pattern in which one ethnic group is dominant. They find that "due to its high ethnic fractionalization, only 40% of African countries are characterized by ethnic dominance, whereas in the other developing regions, 57% of the countries are so characterized. This should also lower the risk of civil war in Africa." (p.18) Thus, after completing their evaluation, they find that "[Africa's] distinctive social structure has made it much less prone to conflict[,]" (p.22) and attribute the violence there to Africa's economic characteristics. (p.25)

The relationship described by Bates (2000) has ethnic fractionalization playing a more nuanced role. Bates examines the relationship between ethnicity, violence and protests in Africa

for the period from 1970 to 1995. He maintains that ethnic fractionalization is one of many factors than can contribute to conflict but does not necessarily lead to conflict in and of itself. According to Bates's findings, "while ethnicity may not provide a sufficient condition for political violence, there exist conditions under which it can become dangerous." (p.134) Bates also finds that political violence (revolts and assassinations) and protests (riots and demonstrations) are both related to his measure of ethnicity (the size of the largest ethnic group). His estimates, though, show different sign patterns for the relationships between ethnicity and violence on one hand and ethnicity and protest on the other. In Bates's findings violence is more likely when the population is either highly heterogeneous or highly homogeneous. Protests, however, are least likely in those conditions and most likely when the ethnic fractionalization present in the country is circa 60%-65%. This finding suggests that Bates's technique of analyzing various types of conflict separately may be a profitable area for further exploration, as is undertaken in the present study.

Horowitz (1985) takes a much more psychological perspective on the topics of ethnic fractionalization and ethnic conflict. He finds that "straightforward relationships between economic rivalry and ethnic conflict are difficult to establish.[...] Economic theories cannot explain the extent of the emotion invested in ethnic conflict." (p.134) According to Horowitz, "an adequate theory of ethnic conflict should be able to explain both elite and mass behavior." (p.226) Economic incentives explain elite behaviour but do not explain why the masses also engage in impassioned violence. Even if the masses are being manipulated by the elites, says Horowitz, these elites "could not use antipathy for their own ends unless ethnic feeling were already strong." (p.225) Horowitz invests much more stock in the fabric of ethnic identification itself. According to his formulation, "[e]thnic conflict arises from the common evaluative significance accorded by the groups to acknowledged group differences and then played out in public rituals of affirmation and

contradiction." (p.225) Thus, it is the joint function of comparative worth and legitimacy that more fully accounts for the passionate, symbolic and apprehensive aspects of ethnic conflict.

Mozaffar (2001) points out that ethnic fractionalization is a very broad term, especially in the context of Africa, where many countries are populated by large numbers of small ethnic groups, none of which is large enough to constitute a majority. The ethnic groups in question can be joined, polarized or split into sub-groups, which creates much leeway for different sorts of behaviour. He notes a difference between deeply-divided societies (composed of two major polarized ethnic groups within which internal differences are suppressed) and multiethnic societies, in which many groups may be present. A dominant group may be present in these multiethnic societies, but there is nonetheless more room for variance in ethnic identification than in these deeply-divided societies he identifies. Mozaffar finds that "multiethnic societies are perhaps less prone to (especially large-scale) violence, not only because the very heterogeneity of the group militates against it but also because the formal and informal mechanisms of intra-ethnic 'self-policing' as well as inter-ethnic networks that help to pre-empt inter-communal violence are easier to design and sustain in smaller and locally focused ethic groups." (p.33) In deeply-divided societies, the whole country is dichotomously configured in rigid ethno-political cleavages, which provides less room for leeway. Internal ethnic sub-grouping, he finds, thus vitiates the formation of these volatile large composite ethnic groups and weakens the prospects of sustaining them if and when they are constructed. (p.7)

Although the positions presented are in many ways starkly opposed, each perspective on ethnic conflict is supported with explanatory narratives and empirical data. Having seen that these theories have merit and are supported by empirical evidence of various sorts, the implication for this study is that to continue to champion one side over the others is perhaps less useful than an in-

depth statistical examination of the conflicts and ethnic fractionalization in question. By carrying out a nuanced examination of available data on ethnic fractionalization patterns and on the categories of violence and by examining the details that distinguish the situations, the discipline can begin to better indicate the relationships among these variables.

2. Existing Measures of Ethnic Fractionalization

Several measures of ethnic fractionalization have been advanced in political science literature. Perhaps the most common measure is the Ethno-Linguistic Fractionalization (ELF) Index, which was derived from data compiled by Soviet ethnographers in the early 1960s. This set of indices provides one index per country, where the index is calculated by finding the Herfindahl-Hirschmann concentration index for the country's ethnic groups.[1]

Although it has been commonly used, the ELF index has many problems, which Daniel Posner (2004) details. (p.850) Among other critiques, Posner points out that the ELF index is approximately forty years out of date. Additionally, the effect of ethnic sub-grouping was not captured in the way the data were coded. The ELF index provides one index per country, which does not capture the multi-tiered cleavages that can play into ethnic conflict.

Posner proposes an alternate measure of ethnic fractionalization dubbed the "PREG" index for "Politically Relevant Ethic Groups". This index also calculates the Herfindahl-Hirschmann index for each country, but the PREG uses updated data on the ethnic groupings and references secondary sources to provide cleaner and more recent data. As such, the PREG measure provides a large improvement over the ELF index that preceded it. Nonetheless, the PREG index also provides only one measure for each country, and in so doing masks behaviour that is due to ethnic

1 The Herfindahl-Hirschmann index is the sum of the squares of the decimalized percentages of the population of each ethnic group. Thus, a country that was evenly split between two ethnic groups would receive a Herfindahl-Hirschmann index of $(0.5^2 + 0.5^2) = 0.5$ and its Mozaffar-Scarritt index would be $(1/0.5) = 2$ for that tier of aggregation.

sub-grouping or to specific patterns of ethnic dominance.

One of the goals of the project is to examine the effects of using measures of ethnic fractionalization that are more refined than are the ELF or PREG indices, which described each country's fractionalization using a single numeric figure. For this reason, the paper uses a more in-depth approach to ethnic fractionalization developed by Mozaffar & Scarritt (2002). Statistical analyses were conducted using a 2005 update the researchers were kind enough to provide to data from the 2002 Mozaffar-Scarritt set.

The Mozaffar-Scarritt (2002) data set provides three indices per country. Each of the three indices represents the degree of fractionalization on a certain tier.[2] Fragmentation Tier One provides a measure of the presence of a long-standing overarching dichotomous cleavage in the country's population, whereby every member of the population is a member of one of two ethnopolitical groups. This type of split is present in only twelve out of the forty countries used in this study, but is included because they are "uniquely both dichotomous and inclusive of the entire population, thus leaving no neutral ground on which escape or compromise is possible." (.p20) Fractionalization of the kind described in Fragmentation Tier Two is more common: this tier represents those ethnic groups that are neither top-tier groups nor nested sub-groups of other ethnic groups. Largely speaking, this tier represents the amount of general ethnic fractionalization. Ethnic fractionalization at this second, middle tier is present in every country except for four (which only have division at their top tier). In this category are included both those groups that are cohesive and those that represent conglomerations of other groups. These conglomerations are represented by the Fragmentation Tier Three. This term tallies the additional fractionalization that comes with nested ethnic grouping. An example of such groups would be the Marakwet and the Pokot in

2 These Mozaffar-Scarritt indices were obtained by taking the reciprocal of the Herfindahl-Hirschmann index of each country's fractionalization at a given tier.

Kenya; both of these groups are officially a part of the Kalenjin group, but have engaged in violent conflict as the largely agriculturalist Marakwet and the largely pastoralist Pokot have competed for land resources.

The present study does not limit itself to these three indices, though. To provide information on the relationship between fractionalization at the various tiers, three interaction terms were also created.[3] Thus, a country with a large value for the 'tiers 1 & 3' interaction term (i.e. "lvl1n3inter" in the data set) would imply a country with a dichotomous split at the top tier (represented by Tier One fractionalization) , but in which these top two groups were composed of further-fragmented sub-groups (represented by Tier Three). An example Mozaffar gives of a country with such a distribution is Nigeria. He describes the north-south cleavage that has emerged with varying intensity at several points as a Tier One divide. He describes the Hausa/Fulani split in the north and the Yoruba/Ibo split in the south as Tier Two cleavages and the Hausa split into the Sokoto and Kano coalitions as a Tier Three cleavage. (pp. 25, 26) These interaction terms were then included alongside the original fractionalization terms as independent variables. By using these Mozaffar & Scarritt data and by examining the interactions of these terms to gauge the effect of simultaneous fractionalization on the various tiers, the project is able to evaluate the influence of these various patterns in greater detail.

3 These three interaction terms are respectively the products of the indices for the top and middle, the top and bottom and the middle and bottom tiers of aggregation.

III. Violence

General Forms

In the same way that more detailed definitions of ethnic fractionalization may add to our understanding of the relationship between fractionalization and violence, more detailed definitions of violence may deepen our understanding as well. One of the primary observations of this study is that the terms "violence" and "conflict" are extremely broad. As discussed, Bates (2000) shows that some forms of conflict may be differently correlated with ethnic fractionalization patterns than are others. To account for this possibility, the violence data used in the study are derived from the Harvard Africa Research Program's data set on violence in Africa.[4] The present study analyzes a newer version of the data used in Bates's project by examining 22 different categories of violence over a twenty-six year period (1970-1995). In our analyses, these Harvard violence data were collapsed by summing over this twenty-six year period, thus providing an indicator of the prevalence of the particular type of violence in each country during this period. A country that had experienced riots in seventeen of those twenty-five years, for example, would receive a score of seventeen in the "riots" variable, whereas a country that had been completely at peace would have a score of zero.[5]

Although the violence variables analysed in this project may be thought of in many ways, this study suggests an understanding of the variables as being divided into four broad categories that are further divided into more specific forms for the actual analysis. These four categories increase in their level of intensity and organisation, as do the nuanced sub-groups into which they are divided.

The four broad groups into which these sub-groups can be divided are the following: *Low*

4 This updated original Harvard Africa Research Program data set is available at http://africa.gov.harvard.edu/data/violnce2.dta

5 This collapsed data set has been made available at web.mit.edu/jdseale/www/data/

Level, Organized Group, High Level and Coups. Although the lines between these categories are undoubtedly blurred, using this division allows the study to develop a broader understanding of the implications of the types of violence analysed within each of these categories.

The first of these categories, *Low Level*, is divided into six sub-groups. The metrics in this group are of events that can be carried out by unorganized groups of people or by groups of people that do not have a militaristic focus. The division begins with a metric of general unrest and is evaluated as UNREST in the data set. This metric can be seen as the lowest level of violence and simply tallies the number of years in which events such as protests or marches occurred. The second metric tallies the number of events in which shots were fired and can thus be seen as being slightly more intense. This metric is evaluated as SHFI in the data set. The third of these tallies the number of years in which at least one riot occurred and is evaluated as RIOT. The fourth, fifth and sixth of these sub-groups are RURAL2, REGCTY2 and CAPCTY2, which respectively tally the number of years in which fighting in the countryside, fighting in a regional city and fighting in the capital city occurred.

The second of these categories, *Organized Group*, is divided into four sub-groups. The metrics in this category are of events that are carried out by groups that are organised with a military goal in mind. The first of these metrics is GUER, which tallies the number of years in which guerilla or terrorist activities took place. The next metric, DOMGRP, represents the presence and activity of a domestic armed group (note: in this metric the years in which a domestic armed group exists and is fighting some entity receive twice the weight of years in which a domestic armed group exists but is not fighting some entity). The third metric tallies the years in which a foreign entity is stationed within the country and is evaluated as FORGN. The fourth and final metric in this category tallies the years in which some domestic entity (either governmental

or non-governmental) is fighting some foreign force (either governmental or non-governmental). This final metric of the second group is evaluated as FORFGHT.

The third of these four broad categories is *High Level*. The six sub-groups in this category are of those metrics that tally years with various types and levels of large-scale violence. The first of these tallies the years in which the government itself was actively engaged in fighting either one or more domestic armed groups, and/or another country, countries or foreign armed groups. This metric is called GOVCON in the data set. The second of these metrics, REWAR, tallies the years in which a regional war took place. Similarly, MLWAR tallies the years in which multiple regional wars took place. The BIGDED2 metric tallies the years in which mass killings of civilians described as a massacre occurred. The GOVFOR metric tallies the years in which the government itself was fighting any foreign entities, governments or non-governmental groups. The last metric, WAR, tallies the years in which the country took part in at least one war.

The fourth and last broad category contains six metrics that tally events concerning coups d'etat. The first metric is ATCOUP, which tallies years in which attempted coups or plots took place. The second is UNCOUP, which tallies the years in which a thwarted coup took place. The third is STAY, which tallies the years in which either a thwarted coup, and/or an attempted coup or plot took place. The fourth metric is SUCOUP, which measures the years in which a successful coup took place. The fifth metric, GOVINS, tallies the number of years in which any of the above (successful, thwarted, attempted or plotted) took place. The final metric is INSTAB, an experimental metric in which years with successful coups are given twice the weight of thwarted, attempted or plotted coups in the final sum.

Table 1: Violence Data

Group	Code	Concept
Low Level	UNREST	Civil unrest: demonstrations, marches, etc.
	SHFI	Conflict with shots fired
	RIOT	Riots
	RURAL2	Fighting in the countryside
	REGCTY2	Fighting in a regional city
	CAPCTY2	Fighting in the capital city
Organized Group	GUER	Guerrilla or terrorist activity
	DOMGRP	Presence of a domestic armed group; years where fighting occurred weighted twice as much as years where none occurred
	FORGN	Foreign entity stationed within the country
	FORFGHT	Some domestic entity (governmental or non-governmental) fighting some foreign entity (governmental or non-governmental)
High Level	GOVCON	Government fighting domestic or foreign groups
	REWAR	Occurrence of a regional war
	MLWAR	Occurrence of multiple regional wars
	BIGDED2	Large-scale killing of civilians described as massacre
	GOVFOR	Government fighting foreign entity
	WAR	Occurrence of war
Coup	ATCOUP	Occurrence of attempted coups and plots
	UNCOUP	Occurrence of unsuccessful coup
	STAY	Occurrence of unsuccessful coup or attempted coups and plots
	SUCOUP	Occurrence of successful coup
	GOVINS	Occurrence of successful coup, unsuccessful coup, attempted coups and plots
	INSTAB	Identical to GOVINS but successful coups are weighted twice as much as otherwise

IV. Correlations

This section provides a step-by-step description of the factual outcomes of the operations carried out. An explanation of each step and its corresponding reasoning is detailed in Appendix 2. Ultimately the study opts to check each of the sub-metrics of violence against each of the fractionalization patterns. Although this requires a large quantity of regressions, which brings with it the increased probability of Type One false positive answers, the multicollinearity of the data precludes the usage of stepwise regressions and other such techniques.

The first step of empirical analysis identifies correlations that demonstrate statistical significance using Ordinary Least Squares regressions. After this a two-step process is used to address the effect of outliers on these correlations. The study then examines the effect of holding income constant by performing multivariate versions of the same correlations with the added independent variable of real GDP per capita. Finally, due to the quantity of regressions examined, further analysis is carried out on all of the correlations that remain significant after their outliers have been dropped to examine both the robustness and the nature of these correlations.[6]

Step 1 Findings: General Correlations

Prior to Controlling for Income

In the first step, the twenty-two metrics of violence were individually regressed on the six measures of ethnic fractionalization for a total of one hundred thirty-two regressions. Twenty of these correlations demonstrated a P>|t| value of less than 0.05.[7] A table of data from these regressions is included as Table 2.

---

6 This further analysis is carried out by checking for heteroskedasticity and by comparing the the kernel densities, the normal probability plots and the quantile plots of the residuals with those of the expected normal distributions. This is done both before and after the outliers have been dropped from each correlation.

7 See Appendix 1 for explanation.

Six metrics of violence lead to these twenty correlations. The category of violence that stands out most markedly is the Coup category. Four of the six metrics measure aspects of coups d'etat; three of these coup metrics are correlated with four measures of ethnic fractionalization and one is correlated with three. The genocide metric of mass killings of civilians (from the broad category of metrics examining high levels of violence) is also correlated with three measures of ethnic fractionalization. The riot metric (from the broade category of metrics examining low levels of violence) is correlated with two measures of ethnic fractionalization. The one broad category of violence that does not show correlation with any ethnic fractionalization pattern is the Organized Group set of metrics.

Table 2: Univariate Analysis

Dep. Var.	Indep. Var.	P Value	Adj. R Squared
atcoup	fragtot	0.043	0.0824
atcoup	lvl2n3inter	0.027	0.1015
atcoup	lvl1n2inter	0.025	0.1043
atcoup	lvl1n3inter	0.020	0.1150
bigded2	fraglvl2	0.018	0.1187
bigded2	lvl2n3inter	0.011	0.1380
bigded2	lvl1n2inter	0.004	0.1850
govins	lvl2n3inter	0.048	0.0777
govins	lvl1n2inter	0.027	0.1024
govins	fraglvl1	0.026	0.1029
govins	lvl1n3inter	0.026	0.1039
instab	lvl2n3inter	0.042	0.0830
instab	lvl1n2inter	0.019	0.1157
instab	lvl1n3inter	0.015	0.1266
instab	fraglvl1	0.008	0.1508
riot	lvl1n3inter	0.017	0.1210
riot	fragtot	0.016	0.1241
sucoup	lvl1n2inter	0.030	0.0977
sucoup	lvl1n3inter	0.013	0.1314
sucoup	fraglvl1	0.001	0.2305

2. Addressing Outliers

The correlations found must be further examined prior to drawing conclusions from the results of these regressions. The regressions in Table 1 are statistically significant but their robustness has yet to be examined. At this point the regressions must be evaluated with regards to

the possible impact of any major outliers. Regressions that are dependent on the placement of their outliers for their statistical significance are not considered to be robust in this study; a correlation must demonstrate statistical significance both with and without its outliers to be considered robust. In order to identify outliers with excess leverage, I constructed bar graphs of all the independent variables were constructed. I also constructed and examined scatterplots of the all correlations.[8]

The Uganda Problem

As will be seen in this section, Uganda's leverage as a Tier Two general fractionalization outlier is great enough to mask the behaviour of the remainder of the data in regressions using terms that involve Tier Two as an independent variable.

The first set of outliers to be evaluated are those outliers in the independent variable. In this case, Uganda presents a problem as an extremely ethnically diverse country. An analysis of Tier Two, which represents general ethnic fragmentation, reveals Uganda's extreme fractionalization (see Figure 1, Tier Two Ethnic Fractionalization). Uganda's Tier Two fractionalization score is 12.238, which is greater than four times the standard deviation (of 2.034) from the mean of 3.227. The next highest score after Uganda for that tier is 6.0937, for the Democratic Republic of the Congo, which is approximately three standard deviations closer to the mean. This suggests that Uganda's presence as an outlier within level 2 could be skewing the data analysis, as confirmed in subsequent analyses. Further examination also reveals that this skewed level 2 is a cofactor in all of the significant correlations for the metric of mass killings of civilians ("bigded2").

8 Scatterplots and graphs have been made available at web.mit.edu/jdseale/www/graphs/

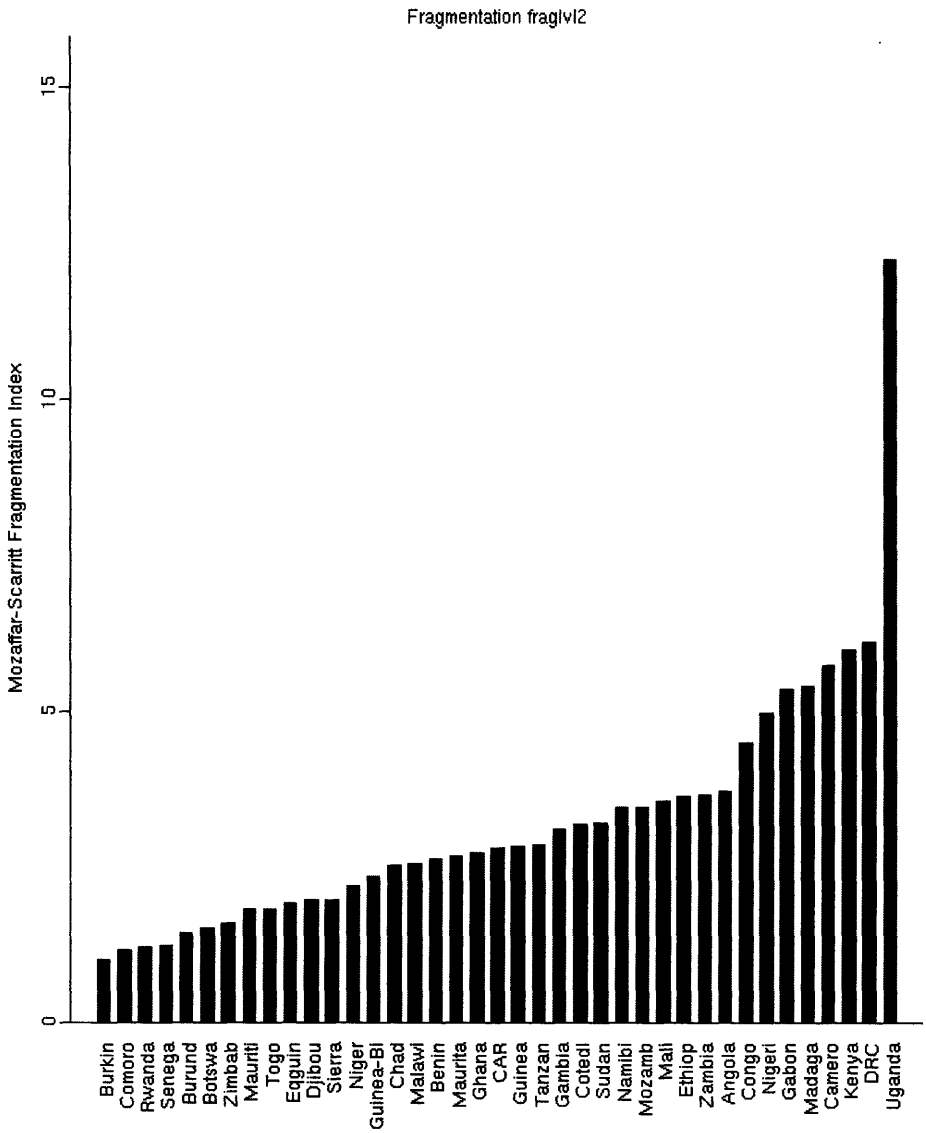


Figure 1: Mozaffar-Scarritt Tier Two Fragmentation

Removal of Independent Variable Outliers

This step examines the correlations without outliers with excess leverage (as opposed to excess influence). To determine whether the extreme fractionalization of Uganda inappropriately skewed the correlations, the study next examines the effects of removing the Uganda observation from the data set. This allows the study to determine the degree to which the statistically significant correlations were dependent on its extreme values. The fear is that Uganda's extremely high degree of second-tier fractionalization and unstable government during the period analysed have confounded the findings by exerting undue leverage on those regressions involving the second tier. For example, it is possible that Uganda's extreme position may have caused otherwise robust correlations to lose statistical significance.

The same set of simple regressions was conducted again, this time without Uganda. The hypothesis in so doing is that dropping Uganda will have the greatest effect on those regressions involving the second tier of fractionalization.

Interestingly, performing these regressions without Uganda reveals it to be a key observation in more than just the correlations involving the second tier of ethnic fractionalization. Upon examining the results once Uganda is removed, shown in Table 2, only seven correlations appear as statistically significant, as opposed to the twenty correlations when Uganda is included.

Thirteen previously-significant correlations have now lost their statistical significance. Correlations with genocide disappear completely for all tiers. With regards to coups d'etat, two of the four metrics (a metric of government instability and the metric tallying attempted coups) also do not appear at all. Each of these had been significantly correlated with four measures of fractionalization. Moreover, the majority of the previously significant correlations disappear from

each of the remaining coup metrics.

Seven relationships retain their statistical significance. For the two remaining coup metrics, only those correlations that had had the highest percentages of explanatory power retained statistical significance. The two remaining correlations, between the general coup metric that gives higher weight to successful coups ("instab") and successful coups ("sucoup") are correlations with Tier One, which represents the degree to which the group demonstrates ethnic fractionalization at a dichotomous top level. The graphs for all of the regressions from this round have been made available online.[9]

Table 3: Correlations Found when Uganda is Excluded

Dep. Var.	Indep. Var	P Value	Adj. R^2
instab	fraglvl1	0.050	0.0777
riot	lvl2n3inter	0.000	0.3100
riot	fraglvl2	0.002	0.2141
riot	lvl1n2inter	0.000	0.3528
riot	lvl1n3inter	0.001	0.2676
riot	fragtot	0.002	0.2241
sucoup	fraglvl1	0.008	0.1566

Examining the quantity of correlations that remain demonstrates to some degree the weakness of the correlations that disappeared. Whereas twelve coup d'etat/fractionalization relationships showed statistical significance when the regressions were conducted with all forty observations, for example, only two relationships retain significance when Uganda is left out and the regressions are conducted with data from the other thirty-eight countries.

However, these thirteen correlations that disappeared should not be automatically rejected en masse. The purpose of this second step, of examining the effect of removing Uganda from the data set, was not to discount correlations that had been previously discovered. That task is left for the post-regression analysis. Rather, the purpose is to make sure that Uganda's presence had not

9 These graphs have been made available at web.mit.edu/jdseale/www/graphs/

caused otherwise robust correlations to lose significance. Judgement of the robustness of the outcomes of both the first and second steps must be reserved until the post-regression analysis is complete.

Indeed, removing Uganda causes new correlations to appear. The metric for riots not only retains the two correlations that had appeared earlier but also reveals three new correlations with ethnic fractionalization, this time with the second tier (on which Uganda had been the outlier) and its two interaction terms with the first and third tiers (see Table 2).

Some of these new correlations would seem to be quite meaningful. In these regressions, for example, the correlation between riots and the interaction term for the first and second tiers has the highest adjusted R^2 value of 0.35. This noteworthy percentage of variance explained deserves further exploration.

Upon further examination, though, the graphs for this regression reveals that the problem with outliers runs deeper than was previously anticipated. This can be seen, for example, in Figure 2 [Riots Regressed on the Interaction Between Tiers 1&2 (excluding Uganda)]. It appears Uganda was not the sole outlier causing this interaction term not to have statistical significance in the first round of regressions; Madagascar and Nigeria also appear to be fairly removed from the bulk of the observations. All of the outliers must be removed for the evaluation of the correlation's robustness to be meaningful. The same correlation without the Madagascar and Nigeria observations, as seen in Figure 3, reveals a much clearer picture of the correlation between riots and the First and Second Tier interaction.

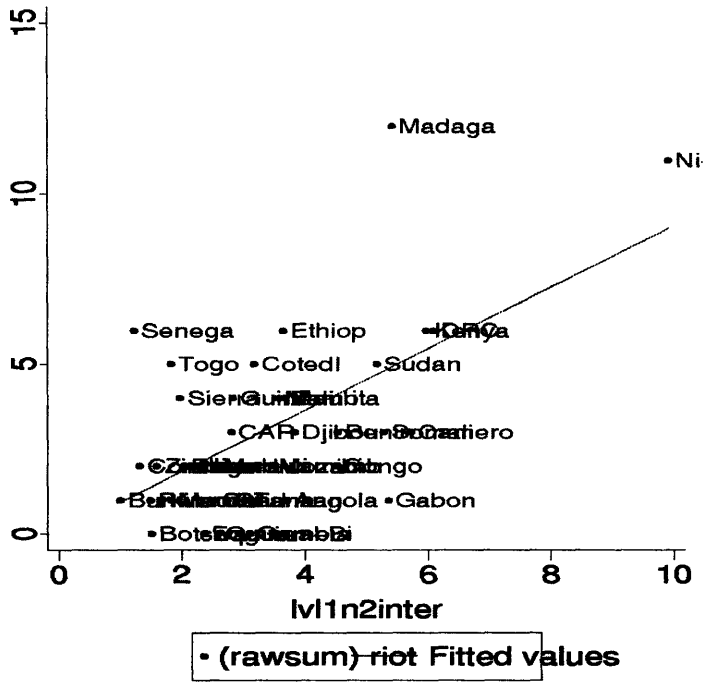


Figure 2: Riots Regressed on the Interaction Between Tiers 1 & 2 (Excluding Uganda)

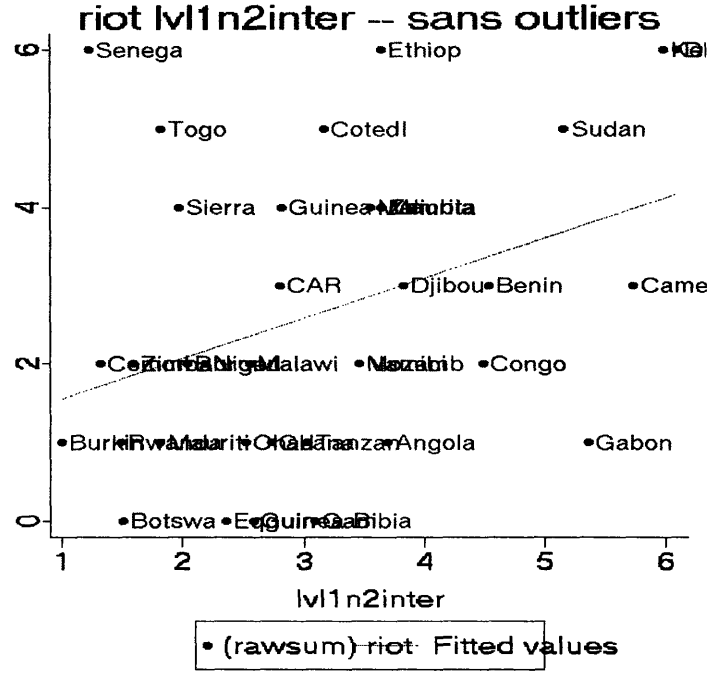


Figure 3: Riots Regressed on the Interaction Between Tiers 1 & 2 (Excluding All Outliers)

Removal of All Outliers

For this reason, this next step examines the correlations without outliers with excess leverage and without outliers with excess influence. To determine which outliers to remove from each correlation, the first twenty statistically significant correlations were once again conducted, this time with the full data set including Uganda. The regressions for the three new riot correlations from the second round were also conducted again, this time on the full data set. This allows the study to examine the original correlations from the full data set and determine which observations have a disproportionate share of influence and leverage in each regression. These observations can be removed and the nature of the underlying correlation examined, if any exists at all.

Table 4: Univariate Analysis Without Outliers

Dep. Var.	Indep. Var.	Adj. R^2	P Value	Observations
govins	fraglvl1	0.2104	0.003	36
instab	lvl1n3inter	0.1031	0.032	36
instab	fraglvl1	0.3147	0.000	36
riot	fraglvl2	0.0870	0.045	36
riot	lvl1n3inter	0.0977	0.036	36
riot	fragtot	0.1046	0.031	36
riot	lvl1n2inter	0.1161	0.024	36
riot	lvl2n3inter	0.1619	0.009	36
sucoup	fraglvl1	0.6926	0.000	33

Determining which observations have a disproportionate share of influence and leverage was carried out by calculating Cook's distance for each of the observations and determining which of these had values greater than $4/n\text{-}k\text{-}1$, in this case 4/(39-1-1). By removing these outliers the study can make an exploratory foray to determine which of these correlations remain significant after all of their outliers have been removed. In this way the study can separate the cases in which the original statistical significance was reflective of the bulk of the data from those in which it was an artefact stemming from the positioning of its outliers.

Each of these twenty-three regressions was performed again, both with and without the outliers in each correlation. The before-and-after scatter plots of these regressions have been made

available online.[10] All of those regressions that demonstrated P>|t| values greater than 0.05 when all their outliers with undue influence or leverage have been dropped are not considered to be robust. This allows the study to consider robust only those correlations that are statistically significant both with and without their outliers. All correlations showing a statistical significance that is dependent on outliers are ultimately considered spurious, although the full post-regression analysis is still conducted on all of them.

Nine of the twenty-three correlations pass this test for robustness. The data from the regressions performed without outliers is summarized in Table 4. Four metrics of violence are responsible for these nine correlations. Of these four metrics, three coup-related metrics —sucoup, instab, and govins— account for a total of four correlations. The remaining five correlations are with the "riot" metric.

The further post-regression analysis conducted on all of the correlations is meant to further reveal the robustness of the correlations, their respective degrees of heteroskedasticity and the nature of the residuals from the regressions. All of the data and all of the graphs from the further post-regression analyses have been made available online.[11]

Although there are fewer correlations for each of the coup variables, the three correlations that explain the highest percentage of the variance are all coup-related. Moreover, all three of these correlations are with the level of top-tier fractionalization. The strongest underlying correlation after all outliers have been dropped is the relationship between the number of successful coups and the degree of top-tier fractionalization. This correlation explains 69% of the variance after the outliers have been dropped. This same correlation of variables explained 23% of the variance prior to dropping the outliers, also the highest percentage from that round.

10 These graphs have been made available at web.mit.edu/jdseale/www/graphs/
11 See web.mit.edu/jdseale/www/post.regression.analyses/

The other two coup metrics with robust correlations are "instab" and "govins". The "instab" variable is correlated with two fractionalization metrics and the "govins" variable with one. Of the two "instab" correlations, the highest percentage of variance explained is found in the correlation with the top tier, which is also the case with the one "govins" correlation. These "instab" and "govins" correlations respectively explain 31% and 21% of the variance, the second and third highest percentages observed in this examination without outliers. They respectively explained 15.4% and 10.5% of the variance in their correlations prior to dropping the outliers.

The "riot" metric was correlated with the highest number of ethnic fractionalization measures, showing correlations with the bottom two tiers of fractionalization and with all three interaction terms. The percentage of variance explained was lower than in the coup correlations, though. The highest percentage of variance explained, even after dropping the outliers, was 16.2%.

Summary of Part I Findings

This first portion of the study finds that the robust simple correlations can be grouped into two general categories, one demonstrating stronger correlations than the other. The first and stronger of the relationships that emerges is between a top-tier of dichotomous ethnic fractionalization and coups d'etat. The degree of top-tier fractionalization has significant explanatory power in these correlations. The second and weaker of these relationships is between general ethnic fractionalization and a low level of violence, namely riots. These metrics are correlated with several metrics of lower-level ethnic fractionalization. In fact, the riot metric is correlated with all of the fractionalization metrics except for top-tier fractionalization. Nonetheless, all of these riot correlations have less explanatory power than those found for the coup metrics.

Step 2 Findings: Controlling for GDP Per Capita

A similar sequence of operations was carried out for the second part of this study, which examines the effect of controlling for income. The same 132 regressions were once again conducted, this time controlling for income. The study uses multivariate linear regressions to hold each country's income per capita constant while examining these same combinations of possible relationships between the twenty-two sub-types of violence and patterns of ethnic fractionalization.

The income data for the period were taken from Heston, Summers & Aten (2002)'s Penn World Table. The value used is the constant real GDP[12] per capita figure and it is taken from the period from 1970 to 1995 for each country. In a manner analogous to that used with the violence data, these figures were also collapsed by finding the mean value for the constant real GDP per capita for the period. Although they will not be addressed in this study, select variables beyond GDP per capita were created and/or included as a portion of a separate forthcoming study on the interplay of various other economic variables and ethnic fractionalization as predictors of violence. These data have also been included in the data sets made available with this study.[13]

Three countries (Djibouti, Somalia and the Sudan) were excluded from the data analysed because of a lack of constant real GDP data. Djibouti was not an autonomous state until its 1977 independence from France and Somalia at the date of this writing has not yet resolved the question of its northern province that has declared its independence as Somaliland. The data necessary to calculate the constant real GDP per capita are not available for the period in the Sudan.

Examining the correlations using the full data set reveals that thirteen statistically significant correlations remained so after controlling for income. This figure represents seven fewer correlations than the twenty that were found prior to controlling for income. Additionally,

12 These were calculated with a Laspeyres fixed base index that uses 1996 as its reference year. Using this figure allows the study to evaluate income per capita's effect while controlling for market fluctuations.

13 This collapsed data set has been made available at web.mit.edu/jdseale/www/data

controlling for income does not cause any ethnic fractionalization pattern to take on significance as a predictor for any new types of violence. Data on the correlations found are presented in Table 4.

Although seven correlations lose significance, the first and most relevant observation is that controlling for income does not eliminate any of the six violence metrics that were significant when the correlations were examined using the whole data set. Four coup metrics, riots and genocide demonstrated statistical significance prior to controlling for income; these same metrics demonstrate significance after controlling for income.

Table 5: Correlations Found When Controlling for Income

Dep. Var.	Beta 1	Beta1 P Val.	Beta 2	Beta2 P Val.	Adj. R Squared
atcoup	rgdpl	0.053	lvl2n3inter	0.0470	0.1692
atcoup	rgdpl	0.052	lvl1n2inter	0.0390	0.1771
bigded2	rgdpl	0.273	lvl2n3inter	0.0200	0.1412
bigded2	rgdpl	0.204	fraglvl2	0.0170	0.1486
bigded2	rgdpl	0.272	lvl1n2inter	0.0060	0.1941
govins	rgdpl	0.028	lvl1n2inter	0.0420	0.1985
govins	rgdpl	0.049	fraglvl1	0.0200	0.2292
instab	rgdpl	0.025	lvl1n2inter	0.0300	0.2165
instab	rgdpl	0.045	fraglvl1	0.0030	0.3012
riot	rgdpl	0.231	lvl1n3inter	0.0400	0.1197
riot	rgdpl	0.189	fragtot	0.0290	0.1335
sucoup	rgdpl	0.073	lvl1n2inter	0.0490	0.1539
sucoup	rgdpl	0.135	fraglvl1	0.0000	0.3829

The effect of controlling for GDP per capita is minimized even further when Uganda's effect is considered. In this case, Uganda and its extreme ethnic diversity present the same statistical problem as they did prior to controlling for GDP per capita. Examining the multivariate correlations again without Uganda reveals that only seven correlations retain their statistical significance. (See Table 6) These seven statistically significant correlations are between the same metrics of fractionalization and of violence as those that were seen before controlling for GDP per capita but did similarly exclude Uganda in the evaluation. (See Table 3)

Just as in that case, three of these seven correlations did not appear in the previous round

that was conducted with the full data set. Just as in the simple linear correlations, all three are correlations between the riot metric and fractionalization metrics involving the second tier.

As was the case with the corresponding simple linear regressions, this step is merely meant to reveal correlations that might have been obfuscated by the presence of confounding outliers with undue leverage. Although the large number of correlations that lost significance does demonstrate their weakness to some degree, the vanished correlations cannot yet be discounted. Determining which of the correlations are robust must once again be left until the post-regression analysis has been completed.

Table 6: Correlations Found Controlling for Income Without Uganda

Dep. Var.	Beta 1	Beta1 P Val	Beta 2	Beta 2 P Val	Adj. R Squared
instab	rgdpl	0.047	fraglvl1	0.0330	0.2051
riot	rgdpl	0.106	fragtot	0.0020	0.2577
riot	rgdpl	0.053	fraglvl2	0.0010	0.2808
riot	rgdpl	0.143	lvl1n3inter	0.0010	0.2922
riot	rgdpl	0.088	lvl2n3inter	0.0000	0.3452
riot	rgdpl	0.069	lvl1n2inter	0.0000	0.3897
sucoup	rgdpl	0.143	fraglvl1	0.0020	0.2900

Final Outcomes: Correlations Adjusted for Both Outliers and Income

The most significant correlations between violence and fractionalization in this study are those that continue to demonstrate significance when both outliers and GDP are taken into consideration. In a procedure analogous to that conducted with the simple linear regressions, all fifteen multivariate regressions that have demonstrated statistical significance thus far are once again conducted on the full data set. For rigour's sake, all other correlations that have demonstrated significance at any point are also fully analyzed while controlling for GDP per capita. This allows the study to examine correlations that had previously been rejected because they relied on their outliers for significance. Controlling for income may cause these correlations to take on actual

significance; the control's effect on these correlations must thus also be evaluated.

As with the correlations that did not control for income, in this step the correlations are examined separately while using the full data set. All outliers[14] are then identified and dropped for each correlation. As with the simple linear correlations, in each case the study can now make an exploratory foray to determine which of the correlations remain significant after all of their outliers have been removed. The same criterion for robustness is used here; only those correlations demonstrating statistical significance both with and without their outliers are considered to be robust. The full post-regression analysis is nonetheless carried out on all the correlations that have demonstrated statistical significance at any point.

The effects of performing these operations can be seen in Table 7. Ethnic fractionalization remains as a significant predictor in five of the nine correlations that were found significant prior to controlling for income. Per capita income takes on significance along with fractionalization in one out of these five. Income also takes on significance as a predictor for two of the other four correlation. Neither ethnic fractionalization nor income are significant for the remaining two.

For further analysis, the correlations can be divided into four groups, depending on the independent variables that show significance when the correlations are examined without their outliers. In each correlation ethnic fractionalization is either significant as a predictor or not significant; likewise with the income variable. A 2x2 matrix of these groups is shown as Table 8.

The first group is of those correlations for which both income and ethnic fractionalization are statistically significant predictors. Only one correlation meets this criterion. This correlation is between real GDP per capita, top-tier fractionalization and the "instab" metric of governmental instability. Interestingly, the "instab" metric was included as an experimental coup d'etat metric

14 Observations with a Cook's distance greater than 4/(n-k-1), in this case 4/(37-2-1), are considered to be outliers.

Table 7: Results of Dropping All Outliers from Each Correlations While Controlling for Income

Dep. Var.	Beta 1	Beta1 P Val	Beta 2	Beta 2 P Val	Observations	Adj. R Squared	Correlations
instab	rgdpl	0.031	fraglvl1	0.004	34	0.3209	BOTH
riot	rgdpl	0.089	lvl1n2inter	0.033	34	0.1422	FRAG
riot	rgdpl	0.073	fraglvl2	0.026	34	0.1545	FRAG
riot	rgdpl	0.096	lvl2n3inter	0.011	34	0.1951	FRAG
sucoup	rgdpl	0.115	fraglvl1	0.000	34	0.7815	FRAG
atcoup	rgdpl	0.037	fragtot	0.824	35	0.0806	RGDPL
atcoup	rgdpl	0.020	lvl2n3inter	0.960	34	0.1147	RGDPL
atcoup	rgdpl	0.021	lvl1n2inter	0.915	34	0.1149	RGDPL
atcoup	rgdpl	0.024	lvl1n3inter	0.537	33	0.1239	RGDPL
govins	rgdpl	0.037	lvl1n3inter	0.742	35	0.0851	RGDPL
govins	rgdpl	0.029	lvl2n3inter	0.701	35	0.0863	RGDPL
govins	rgdpl	0.032	lvl1n2inter	0.774	36	0.0821	RGDPL
govins	rgdpl	0.040	fraglvl1	0.069	33	0.1912	RGDPL
instab	rgdpl	0.027	lvl2n3inter	0.766	36	0.0884	RGDPL
instab	rgdpl	0.036	lvl1n3inter	0.626	35	0.0912	RGDPL
instab	rgdpl	0.019	lvl1n2inter	0.588	35	0.1155	RGDPL
bigded2	rgdpl	0.294	fraglvl2	0.942	36	-0.0253	NOT
bigded2	rgdpl	0.294	lvl1n2inter	0.893	36	-0.0249	NOT
bigded2	rgdpl	0.285	lvl2n3inter	0.546	36	-0.0140	NOT
riot	rgdpl	0.143	lvl1n3inter	0.101	34	0.0893	NOT
riot	rgdpl	0.117	fragtot	0.054	34	0.1188	NOT
sucoup	rgdpl	0.115	lvl1n2inter	0.174	34	0.0744	NOT
sucoup	rgdpl	0.086	lvl1n3inter	0.095	34	0.1026	NOT

that includes both successful coups and unsuccessful or attempted coups and plots in its tally, but that gives successful coups twice the weight of the others.

When income is held constant the correlation accounts for 30.12% of the variance before outliers are dropped. Dropping the outliers in the correlation does not significantly increase its predictive power by bringing it to a value of 32.09%. Nonetheless, the fact that its predictive power increases at all is noteworthy, given that it implies that the outliers were masking the correlation's strength as opposed to producing it.

The second group is of those correlations for which income is not a significant predictor but ethnic fractionalization is significant. Four correlations fall into this category. Riots account for three of these four correlations; the riot metric is correlated with Tier Two fractionalization and

Table 8

Summary of Study Findings:
Impact of Ethnic Fractionalization and GDP per Capita on
Metrics of Violence in Sub-Saharan Africa
(1970-1995)

Correlation of Income using GNP per capita	Correlation of Fractionalization with Specific Metrics of Violence: Correlated	Correlation of Fractionalization with Specific Metrics of Violence: Not Correlated
Correlated	instab on fraglvl1	atcoup on fragtot atcoup on lvl1n2inter atcoup on lvl1n3inter atcoup on lvl2n3inter govins on fraglvl1# govins on lvl1n2inter govins on lvl1n3inter govins on lvl2n3inter instab on lvl1n2inter instab on lvl1n3inter# instab on lvl2n3inter
Not Correlated	riot on lvl1n2inter riot on lvl2n3inter riot on fraglvl2 sucoup on fraglvl1	bigded2 on fraglvl2 bigded2 on lvl1n2inter bigded2 on lvl2n3inter riot on fragtot# riot on lvl1n3inter# sucoup on lvl1n2inter sucoup on lvl1n3inter

Legend
correlation with fractionalization disappeared after controlling for income

with Tier Two's interactions with the first (top) and third (bottom) tiers. The successful coup metric is correlated with Tier One fractionalization.

In terms of predictive power for the correlations in this group, the ethnic fractionalization pattern that best explains the variance in riots is the interaction between general ethnic fractionalization and ethnic fractionalization when ethnic sub-grouping is taken into account. This correlation explains 34.52% of the variance when only Uganda is dropped but only 19.51% of the variance in the quantity of riots when all outliers are dropped.

The remaining correlation is between top-tier fractionalization and the tally of successful coups that occurred during the period. This correlation, which examines the effect of top-tier fractionalization approaching a 50/50 split, has the highest predictive power found in the study. Prior to dropping any outliers, this correlation predicts 38.29% of the variance in the occurrence of successful coups. When the three observations that score as outliers are dropped, though, the percentage jumps to 78.15%. Although this percentage is comparatively very significant, questions remain regarding a potential endogeneity problem when identifying ethnic groups for purposes of the fractionalization indices. (see LIMITATIONS for further discussion)

The third category is of the converse correlations, in which income is statistically significant as a predictor but ethnic fractionalization is not. Of special interest in this category are those correlations in which ethnic fractionalization was a significant predictor prior to controlling for income, but in which fractionalization's impact disappears when the GDP variable is introduced.

Two of the eleven correlations in this category are of this nature. The first of these correlations is between the "govins" coup variable and income when top tier fractionalization is held constant. This correlation is especially noteworthy because of its relationship with the

successful coup metric. The "govins" metric is the coup variable that includes all types of coup activity regardless of its success or failure. This correlation is the one that best explains the "govins" variance both before and after addressing the outliers. It predicts 22.92% of the variance prior to dropping the outliers and 19.12% after doing so.

This relationship between income, coups d'etat and top-tier ethnic fractionalization bears more research. This becomes especially apparent upon examining the second of the two correlations in which fractionalization ceased to be a predictor and income appeared as being significant. This correlation is between the experimental governmental instability ("instab") coup metric and income while holding constant the interaction term for top-tier fractionalization with ethnic sub-grouping ("lvl1n3inter"). Top-tier fractionalization was significant for "instab" when sub-grouping's effect is not considered but ceases to be so for the interaction term between the two.

The success or failure of coup attempts appears to be correlated with top-tier ethnic fractionalization, although the directionality here is unclear. Like the government instability metric, the experimental "instab" metric also gauges all types of coup activity but gives successful coups twice the weight in the tally. The correlations become even more interesting when we recall that the sole correlation for which both income and fractionalization are predictors is the correlation between "instab" and top-tier fractionalization. Thus, the best predictor for overall coup activity is income when top-level cleavage is held constant, but the best predictor for successful coups is top-level cleavage when income is held constant. The only correlation in which both of these predictors are significant is with a combination of these two coup metrics: overall coup activity when successful coups are given greater weight.

Thus, two mechanisms appear to be at work with coups d'etat. This hypothesis is borne out by an examination of the error component of the correlation between successful coups and top-tier fractionalization when income is held constant. Shown in Figure 4 is a plot of the kernel density of the regression's residuals. This figure provides a smooth approximation of the errors in the predicted values. The expected plot is a normal curve centred at 0; this expected curve is overlaid on the actual values found. The expected curve is what would be obtained when most of the predictions for the amount of coups are correct. What was actually found was that although the correlation is linear, the bulk of the predictions are too high by ~.5 coups on average and a significant minority of the predictions are too low, also by ~.5 coups on average. This is the pattern that would be obtained if two separate processes were at work, each one responsible for a certain number of coups and one responsible for more coups than the other. Further examination of these complex interrelationships is needed to determine the mechanism at work here.

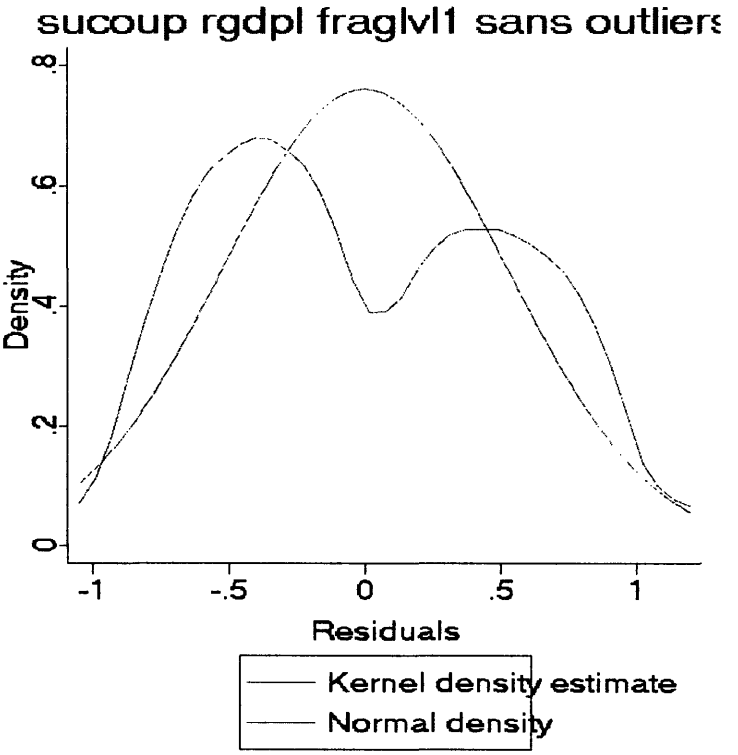


Figure 4: Kernel Density of sucoup--income--tier one Correlations Without Outliers

Although income is statistically significant as a predictor in the other nine correlations, ethnic fractionalization was never robustly significant, neither in the simple linear nor the multivariate regressions. The analysis of these correlations falls beyond the scope of this paper,

which is focused on correlations with ethnic fractionalization. Further work in this area is nonetheless merited. All eleven correlations in this category are between coup metrics and income when various fractionalization patterns are held constant. Income is a statistically significant predictor in twelve of the thirteen correlations found with coup metrics. Ethnic fractionalization is a predictor in only two of these correlations, but these two correlations are the ones with the strongest predictive power.

The fourth and final category is comprised of those correlations in which neither income nor any of the metrics of ethnic fractionalization are robust predictors. Seven correlations fall into this category. Two of the correlations were robustly significant prior to controlling for income. The effect of holding income constant, though, reveals that neither income nor ethnic fractionalization is a significant predictor. These correlations are between riots and Tier Three fractionalization and between riots and the interaction between Tiers One and Three. The remaining five correlations never showed robust significance at any point; they appear in the study only because the positioning of outliers at some point marked them for further analysis that later found them to be spurious.

V. Discussion

By using finer-grained tests to examine the correlations between categories of violence and various ethnic fractionalization patterns, this study is able to demonstrate that controlling for income does not cause ethnic fractionalization's effect to disappear in all cases. With the study's results, it is possible to show that violence is not statistically correlated with ethnic fractionalization in the majority of the cases examined. A minority of categories of violence, however, did emerge as being robustly correlated with measures of ethnic fractionalization.

Holding income constant causes ethnic fractionalization's impact to disappear in a large portion of these correlations but the majority of the correlations nonetheless retain their significance.

Returning to the applicability of Fearon's findings on civil wars and ethnic civil wars to ethnic conflict in general, this study shows that the relationship between ethnic fractionalization and ethnic conflict is too complex for such broad generalizations about the various types of conflict and the various ethnic fractionalization patterns to be made. This study confirms the Bates (2000) finding that different sorts of ethnic conflict are differently correlated with ethnic fractionalization patterns. It also confirms the Mozaffar (2001) finding that deeply divided societies (the top-tier fractionalization index used) are differently correlated with violence than are multi-ethnic societies.

In terms of specifics, this study finds that relatively few types of violence are found to be correlated with ethnic fractionalization and still fewer remain significant after evaluating the correlations without the presence of their outliers. Two patterns of ethnic fractionalization are found to be robustly correlated with two general types of violence, both with and without controls for per capita income. Four of nine individual correlations disappear upon controlling for income. Nonetheless, the five correlations that remain are for the same two kinds of violence as before controlling for income.

The first of these violence metrics is the tally of riots that occurred during the period. These riots are correlated with general ethnic fractionalization. Three of the five significant correlations are between riots and the fractionalization metrics that include the second tier, which is representative of general ethnic fractionalization. The strongest of these riot correlations is with general ethnic fractionalization when its interaction with sub-divisions in these ethnic groups is also considered.

The second type of violence that stands out is with respect to coups d'etat, specifically successful coups d'etat. These two remaining significant correlations are between coup metrics and the degree of top-tier fractionalization that is present. The first of these coup metrics represents the number of successful coups that occurred, the second is an experimental metric that represents the overall tally of coups when successful coups are given twice the weight that is given to the others. This second metric is also significantly correlated with per capita income.[15]

Although the study finds that controlling for income does not cause these metrics to disappear completely, it does find that doing so reveals a complex interrelationship among income, top-tier ethnic fractionalization and the success and failure of coups d'etat. Three coup metrics are found to be correlated with income in a total of twelve correlations. The coup metric is also correlated with top-tier fractionalization in only one of these twelve. In only one case is a coup metric correlated with fractionalization but not income. Thus, coups d'etat are found to be correlated with income in twelve correlations and top-tier ethnic fractionalization in two correlations. Nonetheless, these two correlations with ethnic fractionalization have higher predictive power and lower P values than any of the twelve correlations with income.

In all of the correlations found, ethnic fractionalization is positively correlated with the violence metric being examined. When significant, income is always negatively correlated with violence. However, the directionality of income's correlation is not examined in the study. Statistically speaking, it is possible that income is negatively affected by violence, that tensions ease as financial security increases or that measures of both of these occur in unknown proportions.

The ethnic fractionalization pattern that best predicts the prevalence of any sort of violence is with the presence of a dichotomous top tier of ethnic fractionalization. Violence observed

15 See Appendix 2 for further examination of the correlations found.

increases as the split approaches a 50/50 divide. When controlling for income but not addressing any outliers, this dichotomous pattern explains 38.29% of the variance in the quantity of successful coups that occur. When three outliers are dropped from the correlation, this pattern explains 78.15% of the variance. A robust correlation was found for an additional experimental coup metric, in which a successful coup is given twice the value of an unsuccessful coup, such that the impact of successful coups is greater but the impact of unsuccessful coups is still examined. Although these correlations with coups bear further investigation, it is necessary to take into account the possibility of endogeneity, which is discussed in the Limitations section of this study.

The second and broadest tier of general ethnic fractionalization, in which groups are neither top-tier nor sub-groups of other groups, and its interaction terms representing its presence along with a top-level split and its presence along with the presence of ethnic-subgroups, are all correlated with riots. The percentage of violence explained by the interaction between general fragmentation and ethnic sub-grouping is higher than that explained by general ethnic fragmentation alone. The percent of variance explained by these correlations is comparatively low, though, never reaching above 20% in the top underlying correlation and 14%-15% in the other two. All of these percentages of variance explained represent a drop by 40% to 50% when the outliers are dropped from the regressions.

Some of the results from this study were unexpected. Specifically, examining the errors in the predicted values that emerge from the successful coup metric's underlying correlation indicates the presence of two general processes or variables at work in these successful coups d'etat. Speculation as to the nature of the variables reaches beyond the scope of this study, which only claims to have found both top-tier ethnic fractionalization and income to be intimately related with coups d'etat.

In sum, ethnic fractionalization can be seen to be generally independent from most types of violence in Sub-Saharan Africa for this period from 1970 to 1995. Nonetheless, there do exist some types that are correlated with violence in a statistically significant way. Coups d'etat are correlated with a dichotomous top-tier of ethnic fractionalization, and coups d'etat that have been successful are especially so correlated. This finding must be interpreted in light of its potential problems with endogeneity, as will be discussed in the Limitations portion of the study. Weaker but still robust are the correlations found between riots and Mozaffar's second tier of fractionalization.

Limitations of This Study

As with any study, the data presented from this study must be interpreted in light of the study's limitations. This portion discusses the limitations that must be borne in mind when considering the correlations described in this study. These limitations that must be addressed are of two types. First, the study must examine potential problems with the meaning of the data. What do the data truly represent? Secondly, the study must address potential problems with the methodology used. What does this methodology imply about the information it provides?

In terms of data, the first possible limitation regards the data on violence. Put simply, the question is this: does a sum of dummy variables over twenty-six years provide a reliable metric for gauging the prevalence of violence in a country? It is quite possible that using this metric might mask the intensity of violence that occurs. For example, a country might hypothetically have a small number of years in which intensely violent riots occurred and wherein many lives were lost, but still receive a score lower than a country with a constant degree of less-violent riots.

Although this complaint is certainly valid, the variables used to measure conflict include violence at several levels of intensity, reaching from general unrest at one end of the spectrum to

full-out genocide on the other. This is far from being a perfect solution, but using other metrics for the intensity of violence stretches beyond the scope of this study.

The second possible data limitation regards the malleable nature of ethnic fractionalization. If ethnic identity truly is as malleable as it is perceived to be in much constructivist literature, the accuracy of the Mozaffar & Scarritt (2002) indices used might vary over the period during which the prevalence of violence was analysed. It must be recalled, however, that the period over which the groups were analyzed is only of twenty-six years, from 1970 to 1995. Mozaffar & Scarritt advance the notion of constrained constructivism, in which would-be cultural entrepreneurs are hampered by already-extant ethnic divisions. It is for this reason that Mozaffar and Scarritt (2002) provide indices at the three different tiers of aggregation used in this study, hoping to capture any changes that might occur as a result of elite manipulation of pre-existing divisions. Additionally, Van Evera (2001) contends that there tends to be considerable short-run stability in ethnic identities once they are constructed. He additionally contends that ethnic divisions are solidified by violent conflict, such that its occurrence tends to cause ethnic divisions to persist. With this in mind, by limiting the scope of this study to a timespan of approximately or just under one generation it is assumed that these ethnic divisions remained approximately constant during the period under analysis and that the Mozaffar & Scarritt (2002) indices accurately reflect the degrees of ethnic fractionalization at their tiers of aggregation.

A further potential problem with the ethnic fractionalization data concerns a possible question of endogeneity. The ethnic fractionalization indices used are based on a definition of 'ethnic group' that looks for those groups that have been meaningful on some political basis, with "participation in collective action or conflict (violent or non-violent)" (p.6) acting as one of the indicators that allows a group to be considered relevant. Thus, it might hypothetically be the case

that to some degree the study is examining the correlation between violence and groups defined by taking part in that same violence. This topic was addressed in an April 12, 2006, telephone conversation with Prof. Mozaffar. In this conversation, Prof. Mozaffar agreed that this could be a potential problem with the data, given that this was one of the indicators used when identifying ethnic groups. Further examination of this endogeneity question is needed in future work.

A third potential limitation with the data concerns the information the interaction terms represent. Although they are meant to capture the effect of simultaneous fragmentation on multiple tiers, it is possible that they do not fully address this situation. It is possible, for example, that the top tier's score would need to be weighted to capture the desired behaviour. As an example, having a country with no top-tier fragmentation (i.e. a score of 1) and a bottom-tier fragmentation score of 3 might have a different outcome than a country with a top-tier score of 1.5 and a bottom tier score of 2. Both of these observations would nonetheless receive an interaction term score of 3. Although the interaction term is able to indicate cases in which both are present or both are absent, the figure does not reflect the proportions of the two variables. This limitation does not invalidate any of the correlations found, but it does imply that the study may very well have missed fractionalization patterns that are negatively correlated with violence. This limitation can be dealt with in further work by experimenting with weighting the fractionalization scores.

Finally, in terms of the methodology's limitations, the largest difficulty for the study lies in gauging the validity of findings that emerge from such a large batch of regressions. Much statistical analysis was carried out to allow the study to rule out seeming correlations that are actually the result of outliers and their placement. Even so, eliminating these false correlations does not guarantee that the findings described will continue in the future. Eliminating the correlations that are dependent on their outliers for statistical significance removes a portion of the

possible spurious correlations, but nothing precludes the random appearance of a seeming correlation that would pass this test. It is quite possible that some of the correlations found are statistically significant in both these cases but will not continue to be so in future years. Further research is needed to confirm these results in other settings or to re-assess the same Sub-Saharan setting from other data.

Appendix 1: Statistical Significance Criteria

The procedure used in this step is perhaps the most important portion of the study, which is to determine which of these correlations demonstrates statistical significance. Here, a decision must be made. Fisher (1925) explains that a P>|t| value of less than 0.05 is the statistic generally used for this purpose.(p.41) Using this 0.05 cutoff score indicates that there is a 1/20 chance that the correlation shown arose randomly. In this case, though, with multiple comparisons being made, the probability of a spurious correlation is actually much higher, given that each individual regression would have a 1/20 chance of having arisen randomly. Thus, two options present themselves: the P>|t| value used as a cutoff score can be lowered or it can be left at 0.05 and additional methods used to gauge its statistical significance.

To avoid incorrectly discarding actual significant correlations, the study opts to use a low initial standard of significance by leaving the P>|t| value at 0.05 and then performs additional evaluations to gauge the significance of those results obtained with this liberal cutoff score. Upon further analysis in this case, the appropriate weight for the Bonferroni correction to the P>|t| value would be very difficult to obtain precisely. The difficulty in calculating the appropriate Bonferroni correction stems from the collinear violence data, the collinear fractionalization data and the unknown degree to which the 132 multivariate regressions will be correlated with their simple linear counterparts. Using a non-weighted Bonferroni correction assumes that the comparisons being made are statistically independent. The appropriate correction coefficient will vary as the statistical independence does. In this case, the appropriate cutoff score could hypothetically vary from 1.8 x 10e-4, which is obtained by dividing 0.05 by 264 (for the Bonferroni correction in which all of the simple linear and multivariate comparisons made are entirely statistically independent from each other) to 0.05, which is obtained dividing it by 1 (for the correction in

which all of the comparisons made are exactly collinear).

A P>|t| value set too low could lead to discarding significant regressions, especially if there are other unknown independent variables at work, which is undoubtedly the case. The correlations that passed the test would undoubtedly be significant, but significant correlations could also be unduly discarded. For these reasons the study uses a value of 0.05, not to indicate significance but rather to eliminate those correlations that were obviously not statistically significant. In this case, to sort out the results that were obviously not statistically significant once these regressions had been performed, a Perl programming language script was used to extract the results of the regressions that had P>|t| values of less than 0.05 for the measures of ethnic fractionalization.[16] The regressions were then further analyzed in order to gauge the robustness of these correlations.

16 This Perl script and the resulting data files of statistically significant regressions have been made available at web.mit.edu/jdseale/www/regressions/

Appendix 2: Regression and Post-Regression Analyses Carried Out

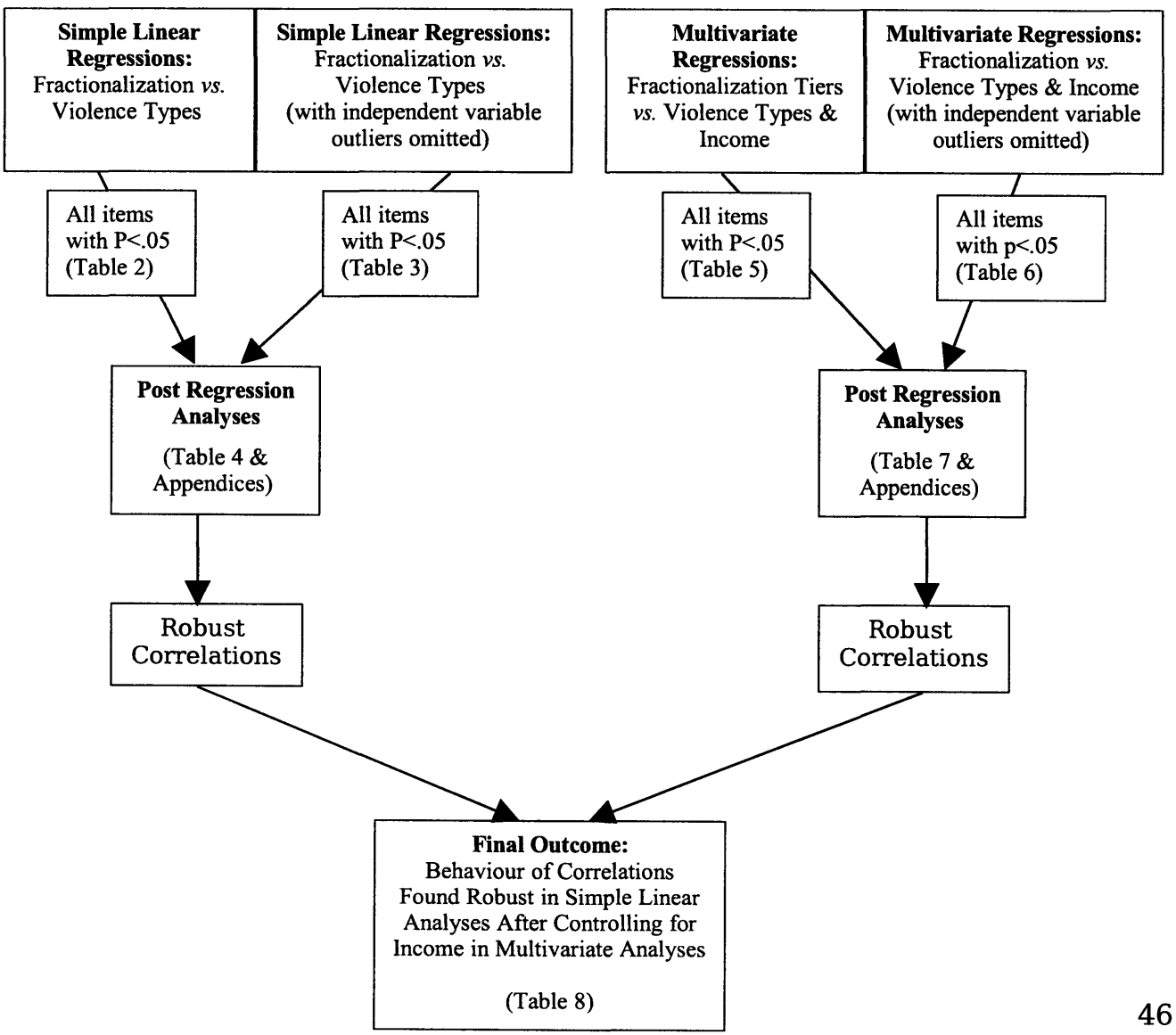
Several options are available to carry out the analysis of these correlations. One option would be to perform a step-wise regression of each of the violence types on the ethnic grouping patterns and the GDP. Unfortunately, this will not provide the desired information. To rigorously examine the data in the manner suggested, each ethnic fragmentation pattern would need to be separately tested against each type of violence; step-wise regressions cannot be used because of the multi-collinearity existing between the tiers of ethnic fractionalization and between the types of violence.

The solution that eventually emerged was in two parts, both with and without a measure of income to examine the applicability of Fearon (2006)'s civil war observation to other types of ethnic conflict. As such, the first part is composed of simple linear correlations and the second part composed of multivariate correlations with GDP per capita as the additional independent variable. For each of these parts, all of the regressions were individually performed for the combinations of these twenty-two measures of violence on the six patterns of ethnic fractionalization – one hundred thirty two correlations evaluated for each part.

The analyses of both the simple linear and multivariate regressions followed a common thread, which can be seen in Figure 5. First, the study identifies the correlations that merit further examination by demonstrating statistical significance on the full data set. Second, the study identifies additional correlations that merit examination, this time those that demonstrate statistical significance once the independent variable outliers with undue leverage have been dropped. Third, the study drops the outliers in each of these correlations in order to examine the underlying correlations and determine which of these have registered as being statistically significant due to he placement of outliers (as opposed to reflecting a trend the bulk of the observations follow).

Finally, the study examines the nature of the residuals from these regressions, in order to identify those regressions in which the residuals do not follow the expected normal distribution. In this step the study tests the regressions to determine whether or not they present statistically significant degrees of heteroskedasticity, skew and kurtosis. The nature of the residuals and the degree of heteroskedasticity are evaluated both before and after dropping the outliers from the regressions.

Figure 5: Process Utilized for Statistical Analyses



The presence of an outlier in the independent variables, namely Uganda, is responsible for another step in the process of identifying those correlations that merit further analysis. Because Uganda's presence masks the behaviour of the remainder of the observations, the fear is that its positioning has caused the study to miss correlations in which the bulk of the data has followed a significant trend. Because it is a lone outlier, though, the effect of outliers in the independent variables can be assessed by excluding Uganda from the data used and performing the same regressions again.

For this reason, both the preliminary simple linear and the preliminary multivariate regressions were performed twice, once with and once without Uganda. In both the simple linear and multivariate regressions any correlations with ethnic fractionalization P>|t| values less than 0.05 from (either the first or second set of regressions) were deemed worthy of continued examination.

Once these correlations meriting further analysis are identified in these two rounds of preliminary regressions, the evaluation of the correlations themselves can begin. The first task is to examine the effect of the outliers in the regressions themselves, as opposed to the outliers in the independent variable. This analysis is done to determine the direction of the type of error these outliers tend to produce. Do they tend to falsely augment the explanatory power of the regressions, leaning toward Type One "false positive" errors? Or do they rather detract from the accuracy an otherwise powerful explanatory variable, leaning toward Type Two "false negative" errors?

To gauge the impact of outliers in the correlations themselves, a more elaborate process was required. First, following Fox (1997)'s criterion, those observations with a Cook's distance (or "Cook's d") value greater than $4/n\text{-}k\text{-}1$ were dropped. (p.281) The Cook's distance represents a combination of the observation's leverage and influence on the predicted variables. The same

regression was then conducted, this time without the outliers. Having done this, the type of error the outliers tend to cause can be determined by comparing the P>*t* values and the R^2 values from before and after the outliers are dropped. If a regression demonstrates statistical significance before these outliers (in the regression) are dropped and ceases do so once they are removed, the correlation is rejected as not being robust. As to the R^2 values, no claims are made about the validity of these R^2 values in terms of their explanatory power; rather, the information being derived is from their proportions, from before and after the outliers are dropped. If the figure goes down, it is because the outliers have skewed the results of the regression toward producing a Type One "false positive" error. If the R^2 value goes up, though, it is possibly because the outliers represent the influence of other unknown explanatory variables. The importances of these values will be revisited in the findings section.

The full analysis described in this section was performed on all regressions that showed statistical significance at any point in the study. With this said, the study only considers as being robustly significant those correlations that received P>|t| values less than 0.05 on at least two occasions, both in at least one of the preliminary rounds of regressions and in the same regression after any outliers with undue influence and leverage were dropped.

An evaluation of the residuals from these regressions was conducted in parallel to this analysis of the outliers. Both the Breusch-Pagan and White's test were conducted for each of these regressions in order to test for heteroskedasticity. Three additional analyses of the residuals and of the studentized residuals were conducted both before and after the outliers were dropped. These analyses are of kernel density, the standardized normal probability plot (pnorm) and normal quantile probability plot (qnorm). These graphs have been made available online.[17] In spite of the

17 See web.mit.edu/jdseale/www/post.regression.analyses.

relatively low number of observations, the kernel density plot provides a smooth approximation of the residuals overlaid with a normal plot for comparison. The other two graph types are the standarized normal probability plot and the quantile plot. The first of these helps to gauge the robustness of the regressions by allowing comparison of the degree to which the cumulative density of residuals follows that of a normal plot with identical mean and standard deviation. The second, quantile plot, allows for a comparison of the degree to which the quantiles (area under the plot) of the residuals from the regressions follow the quantiles of a normal plot with identical mean and standard deviation as the inverse normal of the quantiles grows. These standardized normal probability plots and normal quantile-quantile plots also provide graphic indications of skew, kurtosis and the appropriateness of OLS linear regressions for the data.

The post-regression analysis performed also examines the heteroskedasticity of these correlations with fractionalization to be examined, as well as the normality of the error component of these regressions. The study considers heteroskedasticity tests with P-values less than 0.05 to be significant. Although the normality of the error component is difficult to gauge with so small a sample size, the kernel density plots of the residuals are provided as a smooth approximation drawn from the available data.

Examining the residuals and studentized residuals for normality reveals several things about the five correlations in question. Kernel density plots of the studentized residuals show that the skew in the plots is always to the left, implying that the correlations found tend to overshoot due to a minority of observations with especially high degrees of violence.

None of the correlations ultimately demonstrated meaningful heteroskedasticity. The correlations were evaluated using both the Breusch-Pagan/Cook-Weisberg test and White's general test for heteroskedasticity, with a P-value of 0.05 being used as the cutoff score.

Table 9: Tests for Heteroskedasticity

Dep. Var.	Indep. Var. 1	Indep. Var. 2	Breusch-Pagan	White	Breusch-Pagan *(no outliers)*	White *(no outliers)*
instab	rgdpl	fraglvl1	0.8712	0.9935	0.9601	0.4323
riot	rgdpl	fraglvl2	0.0002	0.1347	0.8929	0.8442
riot	rgdpl	lvl1n2inter	0.0005	0.0574	0.9788	0.7463
riot	rgdpl	lvl2n3inter	0.0005	0.1492	0.5615	0.6656
sucoup	rgdpl	fraglvl1	0.4583	0.8946	0.2158	0.0066

The Breusch-Pagan test found the three riot correlations to demonstrate significant heteroskedasticity prior to dropping their outliers. Because this heteroskedasticity disappeared when outliers were dropped, though, it does not appear to indicate a general pattern in the data. Rather, it is likely due to the influence of such extreme observations as Uganda, which has already been discussed.

Conversely, White's test found heteroskedasticity in the correlation between successful coups and top-tier fractionalization after outliers were dropped. Two observations can be made on this point. First, the two-peaked kernel density plot of the residuals has already been discussed and may be responsible for this heteroskedasticity. Thus, caution is already strongly advised when attempting to use the standard error from this correlation for predictive purposes. This is confirmed by the second observation, which concerns the relative accuracy of the two tests used to evaluate heteroskedasticity. The Lyon and Tsai (1996) study finds that the Breusch-Pagan test is more reliable than White's general test. In this case, the Breusch-Pagan test does not find heteroskedasticity in the correlation. As such, it would appear that the correlation demonstrates some characteristic that White's test considers to be indicative of heteroskedasticity but that Breusch-Pagan is able to discount. Having already made the observation of the two-peaked nature of the kernel density plot, it is very possible that it is this characteristic that is causing this correlation to score positive for heteroskedasticity in White's test but not in Breusch-Pagan.

Bibliography

Bates, Robert. 2000. Ethnicity and Development in Africa: A Reappraisal. *The American Economic Review*. Vol. 90 No. 2. Papers and Proceedings of the One Hundred Twelfth Annual Meeting of The American Economic Association. (May): 131-34.

Collier, Paul and Anke Hoeffler. 2002. On the Incidence of Civil War in Africa. *The Journal of Conflict Resolution*, Vol. 46. No.1 Understanding Civil War. (Feb): 13-28.

Fearon, James. 2006. Ethnic Mobilization and Ethnic Violence. In The Oxford Handbook of Political Economy, ed. Barry Weingast and Donald Wittman. No. 46. Oxford: Oxford University Press

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

Fox, John. 1997. Applied Regression Analysis. Thousand Oaks: *SAGE Publications.*

Heston, Alan, Robert Summers and Bettina Aten. 2002. Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP). (October)

Horowitz, Donald. 1985. *Ethnic Groups in Conflict. Berkeley: University of California Press*

Lyon, John D. and Chih-Ling Tsai. 1996. A Comparison of Tests for Heteroscedasticity. *The Statistician*. Vol.45 No. 3. 337-349

Mozaffar, Shaheen. 2001. Are Multi-Ethnic Societies and Deeply-Divided Societies Different? Prepared for presentation at the 97th Annual Meeting of the American Political Science Association.

Mozaffar, Shaheen and James Scarritt. 2002. Constructivism, Rationalism and The Construction of A Data Set on Ethnopolitical Groups and Cleavage Patterns in Africa. Prepared for presentation at the 97th Annual Meeting of the American Political Science Association.

Van Evera, Steve. 2001. Primordialism Lives! *Comparative Politics Section Newsletter* 12, no. 1 (Winter): 20-22