

CODEBOOK

*VIOLENCE AND RESTRAINT IN CIVIL WAR:  
CIVILIAN TARGETING IN THE SHADOW OF INTERNATIONAL LAW*

Jessica A. Stanton

(Cambridge University Press, 2016)

## Table of Contents

<i>I. Overview of Data Set</i> .....	1
<i>II. Comparisons to Other Data Sets on Violence against Civilians</i> .....	1
<i>III. Identifying Variables</i> .....	4
<i>IV. Variables Measuring Civilian Targeting and Restraint toward Civilians</i> .....	6
<i>V. Variables Measuring Conflict Characteristics</i> .....	12
<i>VI. Variables Measuring Rebel Group Characteristics and Behavior</i> .....	15
<i>VII. Variables Measuring Government Characteristics and Behavior</i> .....	18
<i>VIII. Additional Coding Notes</i> .....	23
<i>References</i> .....	25

## *I. Overview of Data Set*

The book presents an original data set on government and rebel group violence against civilians in all civil wars from 1989 to 2010. The list of cases is drawn from the Uppsala Conflict Data Program (UCDP)/Peace Research Institute, Oslo (PRIO) Armed Conflict Dataset, which defines internal armed conflict as “a contested incompatibility that concerns government or territory or both where the use of armed force between two parties results in at least 25 battle-related deaths” and the conflict “occurs between the government of a state and internal opposition groups” (Gleditsch et al. 2002, 618–19). In accordance with standard definitions of civil war, I include all cases that accumulated at least 1,000 battle-related deaths (Fearon and Laitin 2003; Doyle and Sambanis 2006). In countries where multiple rebel groups challenged the government simultaneously, I consider each conflict a separate civil war, as long as each rebel group maintained separate leadership and participated in fighting that produced 1,000 battle-related deaths.

To identify conflicts that caused at least 1,000 battle-related deaths, I use data on the best estimate of battle-related deaths (variable name: *bdeadbest*) from the PRIO Battle Deaths Dataset, 1946-2008 Version 3.0 (Lacina and Gleditsch 2005); this version of the PRIO Battle Deaths Dataset provides information on battle-related deaths that occurred prior to 1989, whereas later versions of this data set do not. For some cases, the PRIO Battle Deaths Dataset, Version 3.0 groups together multiple rebel groups and estimates battle-related deaths for the conflict as a whole. For these cases, I use Version 5.0-2012 to calculate battle-related deaths for separate rebel groups.

Following Sambanis (2004) and Doyle and Sambanis (2006), I relax the 1,000 battle-related deaths threshold to include the following conflicts that fall just below the threshold. Gathering data on conflict-related deaths is challenging; precise estimates are difficult to obtain, making it likely that conflicts falling just below the 1,000-death threshold did, in fact, cause 1,000 deaths. These conflicts are likely to be similar in character to conflicts above the 1,000-death threshold. Thus, the data set includes the following four conflicts with battle-related death estimates between 700 and 1,000:

- Djibouti/FRUD (1991-1994) – 740 battle-related deaths (Version 3.0)
- Egypt/Gama’a al-Islamiyya (1993-1998) – 918 battle-related deaths (Version 3.0)
- Georgia/Ossetia (1992-1994) – 750 battle-related deaths (Version 3.0)
- Moldova/Dneister Republic (1992) – 700 battle-related deaths (Version 3.0)

Variables measuring government and rebel group strategies of violence and restraint are coded as missing in cases for which I was unable to obtain detailed information on government or rebel group military actions or was unable to differentiate between the actions of multiple warring parties, even after consulting multiple secondary sources as well as conducting exhaustive searches of English-language news media for reporting on these conflicts. The following 12 cases are coded as missing for these variables:

- Burma/Kachins (1961-1992)
- Chad/MPS (1982-1990)
- Chad/MDJT (1999-2005)
- Colombia/EPL (1965-1989)
- Congo-Brazzaville/Cocoyes (1997-1999)
- Georgia/Ossetia (1992-1994)
- Guinea/RFDG (2000-2001)
- India/Manipur, PLA (1982-2007)
- India/Manipur, UNLF (1993-2009)
- India/Tripura, ATTF (1992-1999)
- Iran/KDPI (1979-1996)
- Lebanon/multiple factions (1975-1990)

If I were to keep these cases in the data set, they might skew the results; the lack of information regarding government and rebel group behavior would make it appear as if there was no civilian targeting to report.

The data set codes the main variables of interest for the remaining 103 cases of civil war. The full list of cases, as well as the coding for the variables measuring government and rebel group violence and restraint is provided in the book, *Table 2.3 on pp.84-88*.

## ***II. Comparisons to Other Data Sets on Violence against Civilians***

Several other scholars have developed data sets on civil war violence against civilians; in the discussion that follows, I compare my data set to each of these existing data sets.

### ***Valentino, Huth, and Balch-Lindsay (2004)***

Valentino et al. examine cases of intentional mass killing of civilians during interstate and civil wars. The Valentino et al. data are similar to the data introduced in this book in that we both examine intentional violence against civilians, excluding accidental violence against civilians or collateral damage. The Valentino et al. data, however, differ in several key respects. First, Valentino et al. focus on a much higher threshold of violence, coding only cases of mass killing, defined as 50,000 intentional civilian deaths over a five year period or less. Second, Valentino et al. code cases of government mass killing, but do not code rebel group violence against civilians. Third, the Valentino et al. data set includes both interstate and civil wars. And finally, the Valentino et al. data ends in the year 2000, whereas the data presented in this article continue through 2010.

For the period of time that the Valentino et al. data overlap with my data (1989 to 2000), Valentino et al. identify 11 cases of government mass killing in civil wars. Of these 11 cases, I code government violence as occurring in nine cases. The two cases for which our coding differs (Afghanistan and El Salvador) are conflicts that began in the 1970s and in which much of the violence against civilians took place in the early years of the conflict. Because my data set focuses on violence occurring from 1989 on, it does not capture the violence that occurred in these two cases. For this same time period of overlap between the two data sets, my data set codes government violence for a number of additional cases beyond the 11 cases of mass killing identified by Valentino et al. This is because, in many cases, governments used significant violence against civilians, but this violence did not meet the Valentino et al. threshold of 50,000 civilian deaths.

### ***Eck and Hultman (2007) – UCDP One-Sided Violence Data Set***

The existing data set most similar to the one presented in this book is Eck and Hultman's (2007) data set on one-sided violence in civil wars. Like the data set presented in this article, the Eck and Hultman data set focuses on civil wars. The time periods covered in the two data sets are nearly the same; Eck and Hultman's original data set covered the period from 1989 to 2004, but has since been extended. In addition, the two data sets share a similar definition of violence against civilians. Eck and Hultman define one-sided violence as "the intentional and direct use of violence" against civilians by a government or other organized group (p.235). The focus on intentional, direct violence excludes violence that is accidental or the result of collateral damage, much like the data set presented in this book.

Eck and Hultman's one-sided violence data differ in two key respects from the data presented in this book. First, one-sided violence must reach a threshold of 25 civilian deaths per year to be included in the data set. The Eck and Hultman data thus excludes violence that leads to significant destruction of civilian targets, but does not kill large numbers of civilians. My data set, in contrast, includes forms of violence that intend to harm civilians through the destruction of civilian targets, but may not always lead to immediate civilian deaths – for example, scorched earth campaigns or the deliberate bombing or shelling of civilian areas. Rebel groups may use small-scale bombs placed in public areas that intend to kill large numbers of civilians, but for a variety of reasons (timing, poor-functioning bombs, etc.) may not always succeed in doing so.

Second, the way in which violence is coded differs between the Eck and Hultman data set and the data set presented in this book. The Eck and Hultman data set provides count data on the number of civilians killed by each actor in each year, during incidents of one-sided violence. Count data, however, present a number of challenges. Civilian deaths are often not well documented; sources may note, for example, that some number of civilians were killed, but may not provide specific estimates. The availability of information on civilian deaths also varies significantly across countries and across time, making it difficult to use count data as a basis for cross-country comparison. In addition, data on the number of civilians killed exclude forms of violence that impose significant destruction, but do not kill large numbers of civilians.

The data set presented in this book addresses many of these concerns by focusing on forms of violence against civilians. For most civil wars, sources report on general patterns of violence, even if they do not provide estimates of the numbers of civilians killed. For this reason, the data set presented in this book is less likely to omit inadvertently cases of violence against civilians and thus, offers a more reliable means of comparing violence against civilians across cases of civil war. In addition, the data set presented in this book includes a wider range of cases of violence against civilians, including cases that involve forms of violence, such as scorched earth campaigns or small-scale bombings, which do not always produce high civilian casualties in the short term.

A closer comparison of the two data sets illustrates some of these differences. In 26 of the 52 cases that I code as involving rebel group violence against civilians, Eck and Hultman record fewer than 500 one-sided violence fatalities for the conflict. Among rebel groups, much of the mismatch between my data set and the Eck and Hultman data set is the result of cases in which rebel groups used small-scale bombs to attack populated civilian targets or used scorched earth campaigns. As noted above, these forms of violence may intend to harm or kill large numbers of civilians, but may not always lead to high numbers of civilian deaths in the short term.

In 22 of the 50 cases that I code as involving government violence against civilians, Eck and Hultman record fewer than 500 one-sided violence fatalities for the conflict. For example, in the Eck and Hultman data, the Turkish government is not listed as responsible for any one-sided violence fatalities, but in my data set Turkey is coded as having engaged in violence against civilians. The Turkish government deliberately burned thousands of Kurdish villages during its fight against the PKK; but the Turkish government evacuated villages before destroying them, thus limiting the number of civilian casualties.

Nearly all of the cases coded as restraint in my data set are also coded as having low numbers of one-sided civilian fatalities in the Eck and Hultman data set. In four cases (three government cases and one rebel group case), however, I code a government or rebel group as having used restraint,

but Eck and Hultman attribute more than 500 one-sided violence fatalities to the government or rebel group. The three government cases are: Nepal, Guatemala, and the Democratic Republic of the Congo (for the years 1998 to 2000). The one rebel group case is LURD in Liberia.

In these cases, most of the violence perpetrated against civilians involved the selective targeting of individuals believed to be collaborating with the opponent. In most cases involving high levels of selective violence, the belligerent also used other forms of violence against civilians and, thus, is not coded as exercising restraint in my data set. But these four cases are somewhat unusual, in that these four armed groups did not use any other forms of violence against civilians; thus, these four cases are not coded as cases of violence in my data set. In robustness checks, I use an alternate measure of restraint that code these as cases of violence rather than restraint. The results do not change substantially.

### ***Raleigh (2012)***

Raleigh uses data on violence against civilians from the Armed Conflict Location and Event Dataset (ACLED). The ACLED data differ in several key respects from the data introduced in this book. First, the ACLED data set is still in progress, so data were not available as of 2016 for all countries in the world. Second, the ACLED data set's coverage begins in 1997 for African countries, later for countries in the Asia and the Middle East. Third, the ACLED data set provides event data; for each country, it codes each incident of violence as a separate event. While at first glance, the ACLED event data appear to allow for a more fine-grained analysis of the patterns of violence against civilians, the way in which these event data are categorized in the ACLED data set makes such an analysis difficult, as I discuss in the next two paragraphs.

The ACLED data set includes all incidents of violence occurring in a country, including incidents of violence that did not involve civilians; although ACLED does provide a variable that identifies incidents of violence against civilians, ACLED's categorization of violence against civilians is much broader than the categorization I use. ACLED categorizes as violence against civilians all incidents in which civilians were the victims of violence, including incidents in which violence against civilians was unintentional or was the result of collateral damage.

Finally, ACLED does not differentiate between different types of violence against civilians. Incidents in which civilians were tortured or beaten enter into the ACLED data set equally with incidents in which civilians were killed. And in some cases, a set of incidents is grouped together and entered into the ACLED data set as a single incident of violence against civilians. For example, in the Democratic Republic of the Congo conflict with the AFDL, the AFDL's killing of 25,000 Hutu refugees, which took place in a series of incidents, is coded in the ACLED data set as one incident occurring during March 1997. The problem with grouping these incidents together and coding them as a single incident of mass violence against civilians is that this "incident" then enters into the data set equally with less severe incidents. Data analysis using the ACLED data set would treat these two "incidents" the same.

### *III. Identifying Variables*

**countryid**                      country ID

**Source:** Correlates of War (COW) Country Codes; available at <http://www.correlatesofwar.org>

**government**                      name of government actor

**rebels**                              name of rebel group actor

**startmo**                              month in which the conflict began

**startyr**                              year in which the conflict began

**endmo**                                month in which the conflict ended

**endyr**                                year in which the conflict ended

**post89**                                conflict began after 1989

0 = conflict began on or before December 31, 1989

1 = conflict began on or after January 1, 1990

**Source:** start and end dates rely on data from the UCDP/PRIO Armed Conflict Dataset and the Duration Data v1-2006b (Gleditsch et al. 2002; Gates and Strand 2006). These data are supplemented with information from secondary sources where necessary.

Start date indicates the date on which the conflict first accumulated 25 total battle-related deaths. If the *first* battle-related deaths occurred within the 12 months immediately prior to the date on which the conflict reached 25 total battle-related deaths, then the start date indicates the date on which the *first* battle-related deaths occurred.

End date indicates the last date of fighting prior to a period of inactivity. If the last battle-related deaths occurred within the 12 months immediately following, then the end date indicates the date on which the last battle-related deaths occurred.

In some cases, UCDP/PRIO codes conflicts as having several different phases, each with distinct start and end dates. If the rebel group and the government do not change across these phases of conflict, the phases are coded as a single conflict, using the start date from the first phase of conflict and the end date from the last phase of conflict.

**asia** conflict located in Asia

0 = conflict not located in Asia

1 = conflict located in Asia

**eeurope** conflict located in Eastern Europe

0 = conflict not located in Eastern Europe

1 = conflict located in Eastern Europe

**lamerica** conflict located in Latin America

0 = conflict not located in Latin America

1 = conflict located in Latin America

**nafrme** conflict located in North Africa or the Middle East

0 = conflict not located in North Africa or the Middle East

1 = conflict located in North Africa or the Middle East

**ssafrica** conflict located in Sub-Saharan Africa

0 = conflict not located in Sub-Saharan Africa

1 = conflict located in Sub-Saharan Africa

**distanceUSA** distance in kilometers between Washington, D.C. and the capital city in the rebel group's country of origin

**logdistanceUSA** distance in kilometers between Washington, D.C. and the capital city in the rebel group's country of origin, logged

**Source:** Minimum Distance Data (Gleditsch and Ward 2001).

#### ***IV. Variables Measuring Civilian Targeting and Restraint toward Civilians***

In measuring civilian targeting and restraint toward civilians, I include as civilian targeting attacks on civilian populations and civilian targets – for example, civilian residential areas, schools, hospitals, camps for internally displaced persons or refugees, as well as civilian infrastructure sites (for more on infrastructure, see the discussion of terrorism below). My definition of civilian populations and civilian targets is similar to the definition used in international humanitarian legal treaties. The war crimes provision of the Rome Statute of the International Criminal Court outlaws attacks “against the civilian population as such or against individual civilians not taking direct part in hostilities” (Article 8.2.b.i) or “against civilian objects, that is, objects which are not military objectives” (Article 8.2.b.ii).

For the purposes of coding, I consider an individual to be a civilian if s/he is not taking direct part in hostilities and is not a member of an armed opposition group, a member of the government’s military or police forces, or a government official. In addition, I do not code attacks on police or government vehicles or buildings as attacks on civilian targets. Although international legal definitions would consider police officers and government officials not actively taking part in hostilities to be civilians, in the context of a civil war, police and government actors often work closely with military forces in carrying out counterinsurgency operations. This can make it difficult to differentiate between the activities of military forces, police forces, and government officials. Using a narrower definition of civilian thus reduces ambiguity in the coding, allowing for a focus on deliberate attacks on civilian targets.

I identify civilian targets based on the primary purpose or use for that target. For example, any vehicle, building, or place that is not affiliated with the government, police, or military is considered a civilian target. In addition, I consider infrastructure targets, such as oil pipelines, sanitation systems, bridges, railway lines, power lines, and electric stations, to be civilian targets. Although one might argue that the civilian infrastructure is crucial not only to the life of the civilian population and the functioning of the economy, but also to the activities of the military forces, the primary purpose of these infrastructure targets is to provide for the day-to-day functioning of the civilian population and the civilian economy. This is also broadly consistent with the provisions of international humanitarian law, which prohibit attacks on objects that would endanger the civilian population. The Second Additional Protocol to the Geneva Conventions, for example, prohibits attacks on infrastructure targets such as dams or power stations that would endanger the civilian population (Article 15) and prohibits attacks on objects essential to civilian survival, such as crops, drinking water, and irrigation systems (Article 14).

In coding violence against civilians, “I focus on violence that is intentional, in which sources indicate that government or rebel group attacks on civilian populations or civilian targets were deliberate. I do not include incidents of collateral damage, in which civilians are harmed or civilian targets are destroyed during a military engagement with the opponent. To do so would require making judgments about the necessity and proportionality of a military action, as well as whether the actor responsible made adequate attempts to avoid civilian casualties. I am interested in capturing patterns of behavior – forms of violence that are repeated throughout the conflict...Focusing on patterns of conduct limits the need to make judgements about the legality of particular incidents of violence, and reduces the likelihood of attributing individual criminal acts to the armed group as a whole.” (Stanton 2016, 65).

**grestraint**                    government use of restraint

0 = government did not exercise restraint toward civilians

1 = government exercised restraint toward civilians

**rrestraint**                    rebel group use of restraint

0 = rebel group did not exercise restraint toward civilians

1 = rebel group exercised restraint toward civilians

Restraint measures “whether each government and each rebel group refrained from the most severe forms of civilian abuse. A government or rebel group is coded as having exercised restraint if it did *not* use any of the following four forms of violence against civilians: massacres, scorched earth campaigns, forced expulsion of a particular ethnic or religious group from territory, or deliberate bombing and shelling of populated civilian targets.” (Stanton 2016, 68) See pp. 67-70 in the book for additional information on the concept and measurement of restraint.

It is important to note that restraint is not the same as compliance with international humanitarian law. Although groups coded as exercising restraint toward civilians have complied with some provisions of international humanitarian law – namely, provisions prohibiting large-scale attacks on civilian populations and civilian targets – these same groups may not have complied with other provisions of international humanitarian law. For example, groups that exercised restraint by avoiding massacres or bombings of civilians may have engaged in smaller scale attacks against civilians, such as the killing of individual civilians suspected of aiding the opponent.

**grestraint\_3cat**                government use of restraint (three categories)

0 = government used higher-casualty forms of violence against civilians

1 = government used lower-casualty forms of violence against civilians

2 = government exercised restraint toward civilians

**rrestraint\_3cat**                rebel group use of restraint (three categories)

0 = rebel group used higher-casualty forms of violence against civilians

1 = rebel group used lower-casualty forms of violence against civilians

2 = rebel group exercised restraint toward civilians

“For this measure, the highest category of restraint is the same as in the dichotomous measure – indicating a government or rebel group that did not use any of the following four forms of violence against civilians: massacres; scorched earth campaigns; forced expulsion of a particular ethnic or religious group from territory; or deliberate bombing and shelling of populated civilian targets. At the other end of the spectrum, the category of highest violence includes governments or rebel groups that used either of the highest casualty forms of violence against civilians: massacres or indiscriminate bombing or shelling of populated civilian targets. The intermediate category captures governments and rebel groups that used lower-casualty forms of violence against civilians – scorched earth campaigns, bombing of economic or infrastructure targets, or forced expulsion of civilians from territory – but did not use either of the highest casualty forms of violence (massacres or bombing of populated civilian targets).” (Stanton 2016, 69–70)

**ghicontrol** government use of high-casualty control

0 = government did not use high-casualty control

1 = government used high-casualty control

**rhicontrol** rebel group use of high-casualty control

0 = rebel group did not use high-casualty control

1 = rebel group used high-casualty control

High-casualty control captures “whether a government or rebel group carried out more than five massacres of civilians during the civil war.” (Stanton 2016, 71) See pp. 71-72 in the book for additional discussion of the measurement of high-casualty control. If a government or rebel group used massacres in combination with forced expulsion of a particular ethnic or religious group, this is coded as a case of high-casualty cleansing, rather than a case of high-casualty control.

**ghicleansing** government use of high-casualty cleansing

0 = government did not use high-casualty cleansing

1 = government used high-casualty cleansing

**rhicleansing** rebel group use of high-casualty cleansing

0 = rebel group did not use high-casualty cleansing

1 = rebel group used high-casualty cleansing

High-casualty cleansing captures “cases in which a government or rebel group forcibly expelled civilians from a particular ethnic or religious group from contested territory *and* also used scorched earth tactics and/or massacres.” (Stanton 2016, 72) See p. 72 in the book for additional discussion of the measurement of high-casualty cleansing.

**ghiterrorism** government use of high-casualty terrorism

0 = government did not use high-casualty terrorism

1 = government used high-casualty terrorism

Government high-casualty terrorism “captures cases in which the government engaged in intentional bombing and shelling of populated civilian targets – for example, the bombing of a village at a time when no rebel military forces or military bases were present and when no military engagement with the rebel group was ongoing. Bombing or shelling directed at the rebel group’s troops, weapons, or military bases is excluded, even if civilians were killed in the crossfire.” (Stanton 2016, 74)

Intentional bombing and shelling includes aerial and naval bombardment as well as artillery shelling. In many cases, governments claim that bombing or shelling of civilian targets is unintentional, insisting that the target of attack was military. For example, governments might assert that rebel forces were stationed near to civilian areas, making it difficult to prevent collateral damage to civilian areas during attacks on rebel group units or bases. To address this issue, I only code a government as having used high-casualty terrorism if sources unaffiliated with either the government or the rebel group (media outlets, NGOs, human rights groups, the United Nations) indicate that the

government was responsible for attacking civilian targets, and if such behavior on the part of the government was a repeated tactic. See pp. 72-74 in the book for additional discussion of the measurement of terrorism.

**rloterrorism** rebel group use of low- or high-casualty terrorism

0 = rebel group did not use terrorism

1 = rebel group used low- or high-casualty terrorism

Rebel group terrorism measures “whether or not a rebel group used small-scale bombs, such as car bombs, suicide bombs, or improvised explosive devices (IEDs), to attack civilian targets; the measure does not include artillery shelling or bombing of towns or cities.” (Stanton 2016, 73) See pp. 72-74 in the book for additional discussion of the measurement of terrorism.

Key to coding this variable is distinguishing between attacks on civilian targets and attacks on government or military targets. Rebel groups often use similar types of small-scale bombs to attack police and military vehicles, police stations, military posts, government offices, and other government or military targets. I code as terrorism only those cases in which the target of attack was an unambiguously civilian target; I identify civilian targets based on the primary purpose or use for that target, as described above.

The following two measures differentiate between rebel group low- and high-casualty terrorism by focusing on the primary target of attack. Although some rebel groups using terrorism attack both low- and high-casualty targets, the following measures categorize groups according to their most common type of bombing target.

**rloterrorism** rebel group use of low-casualty terrorism

0 = rebel group did not primarily use low-casualty terrorism

1 = rebel group primarily used low-casualty terrorism

This measure reflects cases in which the rebel group primarily used low-casualty terrorism, rather than high-casualty terrorism. Rebel group low-casualty terrorism “includes cases in which the rebel group bombed infrastructure targets, such as power stations, oil pipelines, or bridges; although attacking these types of targets imposes high costs on civilians by disrupting electric and sanitation services or impeding transportation, casualties are rare.” (Stanton 2016, 73) See pp. 72-74 in the book for additional discussion of the measurement of terrorism.

**rhiterrorism** rebel group use of high-casualty terrorism

0 = rebel group did not primarily use high-casualty terrorism

1 = rebel group primarily used high-casualty terrorism

This measure reflects cases in which the rebel group primarily used high-casualty terrorism, rather than low-casualty terrorism. Rebel group high-casualty terrorism “includes cases in which the rebel group bombed populated civilian targets, such as buses, restaurants, and markets.” (Stanton 2016, 73) See pp. 72-74 in the book for additional discussion of the measurement of terrorism.

To distinguish between *Rebel Group High-Casualty Terrorism* and *Rebel Group Low-Casualty Terrorism*, I focus on the target of attack. Although these variable names refer to low and high casualties, I do

not use a particular threshold number of casualties to differentiate cases of low- or high-casualty terrorism; rather, I differentiate these cases based on whether the target was a target likely to produce low or high casualties. High-casualty civilian targets are populated public civilian targets such as markets, shopping centers, restaurants, hotels, buses, trains, bus or train stations, or any other public place likely to be occupied by significant numbers of civilians. Public places used for government or military purposes, such as police stations or government office buildings, are not included as high-casualty civilian targets. Low-casualty civilian targets are civilian targets that tend not to be occupied by significant numbers of civilians. The majority of these targets are infrastructure targets, such as those described above. Although an attack on a power station or an oil pipeline may kill a small number of civilians – for example, employees working at the power station or on the pipeline – the intended impact of such attacks is to interrupt the provision of civilian services or the functioning of the economy through the destruction of the physical infrastructure itself, rather than to kill large numbers of civilians.

All of these terrorist attacks could be considered part of a strategy of guerrilla warfare. Guerrilla warfare involves small-scale, hit-and-run attacks on the opposing side, with the aim of gradually wearing down the opponent over time through mounting costs. It is possible to use a strategy of guerrilla warfare by attacking only government and military targets and without attacking any civilian targets; such a strategy would involve small-scale, hit-and-run attacks on targets such as police or military vehicles, police or military forces out on patrol, and police and military posts or supplies. It is also possible to use a strategy of guerrilla warfare that combines small-scale, hit-and-run attacks on government and military targets with similar types of attacks on civilian targets. Terrorist attacks, thus, can be considered a form of guerrilla warfare distinguished by the type of target.<sup>1</sup>

**g5massacre**                      government massacres of civilians

0 = government did not massacre civilians

1 = government massacred civilians

**r5massacre**                      rebel group massacres of civilians

0 = rebel group did not massacre civilians

1 = rebel group massacred civilians

“Massacres are defined as the deliberate killing of a group of five or more civilians at close range in a single incident. Civilians killed by bombs or killed in the crossfire during a battle are not included. Only those groups that were responsible for more than five massacres during the course of the conflict are coded as having used this form of violence.” (Stanton 2016, 68)

---

<sup>1</sup> For a detailed discussion of the distinction between guerrilla warfare and terrorism, see for example, Ganor (2002).

**g2massacre**                    government massacres of civilians

0 = government did not massacre civilians

1 = government massacred civilians

**r2massacre**                    rebel group massacres of civilians

0 = rebel group did not massacre civilians

1 = rebel group massacred civilians

Uses the same definition of massacre described above under *g5massacre* and *r5massacre*. However, *g2massacre* and *r2massacre* use a lower threshold for the total number of massacres reported during the conflict; coding of “1” reflects cases in which the government or rebel group was responsible for more than two massacres during the course of the conflict.

***gscorched***                    government use of scorched earth strategies

0 = government did not use scorched earth strategies

1 = government used scorched earth strategies

***rscorched***                    rebel group use of scorched earth strategies

0 = rebel group did not use scorched earth strategies

1 = rebel group used scorched earth strategies

“Scorched earth strategies involve the intentional burning or destruction of villages and/or agricultural land.” (Stanton 2016, 68)

## V. *Variables Measuring Conflict Characteristics*

**troop** ratio of average government troop strength to average rebel group troop strength

**logtroop** ratio of average government troop strength to average rebel group troop strength, logged

**Source:** troop data constructed using annual data from *The Military Balance* and the *SIPRI Yearbooks* (Stockholm International Peace Research Institute (SIPRI) multiple years; The International Institute for Strategic Studies (IISS) multiple years). These annual data are averaged across all years of the conflict.

Government troop estimates include active armed forces, excluding paramilitary forces and police.

**rebstrength\_NSA** alternate measure of rebel group military strength

**Source:** Non-State Actor Data (Cunningham, Gleditsch, and Salehyan 2009), *rebstrength* variable.

This measure uses a five-point scale to code the strength of the rebel group relative to the government, ranging from much weaker to much stronger. *Rebstrength\_NSA* inverts the scale of the Cunningham et al. measure, to allow for comparison with *troop*.

**bdeath** total battle-related deaths for the conflict

**logbdeath** total battle-related deaths for the conflict, logged

**Source:** PRIO Battle Deaths Dataset, Version 3.0 (Lacina and Gleditsch 2005).

The PRIO Battle Deaths Dataset provides three battle-related death estimates for each conflict: low estimate, high estimate, and best estimate. *Bdeath* is coded using best estimate.

**duratn** duration of the conflict calculated in months, inclusive of the first and last month of the conflict

**logduratn** duration of the conflict, logged

**avdeath** battle-related deaths, per year of the conflict

**Source:** calculated using *bdeath* and *duratn*. For example, if the conflict lasted for 18 months or 2.5 years),  $avdeath = bdeath / (2.5)$ .

**logavdeath** battle-related deaths, per year of the conflict, logged

**pcgdp** per capita GDP

**logpcgdp** per capita GDP, logged

**Source:** data are from Gleditsch (2002), variable name “rgdp96pc”. Gleditsch uses the Penn World Tables to calculate per capita GDP; all numbers are in 1996 \$US. In the Penn World Tables, variable is called “Real GDP Per Capita (Constant Prices: Chain Series).”

Per capita GDP is coded for the year prior to the conflict start year. If GDP data are not available for the year prior to the conflict start year, per capita GDP is coded for the conflict start year. For conflicts that began prior to 1989, per capita GDP is coded for the year 1988 (one year prior to the first year of data collection).

Note: Gleditsch (2002) data cover years through 2000. For conflicts that began in 2002 or later (and would require GDP data for 2001 or later), per capita GDP is coded using data from the year 2000.

**separatist** type of conflict

0 = rebel group has revolutionary goals

1 = rebel group has separatist goals

Source: Fearon and Laitin (2003), updated to 2010.

**multiparty** multiparty conflict

0 = conflict is not a multiparty conflict

1 = conflict is a multiparty conflict

**Source:** coded using information on “incompatibility” (variable name: *incomp*) from the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al. 2002).

If multiple rebel groups are fighting against the same government, but they are fighting over different issues (different incompatibilities), then these rebel groups are coded as fighting in separate single-party conflicts. If multiple rebel groups are fighting against the same government and are fighting over the same issues (the same incompatibility), then these rebel groups are coded as fighting in a multiparty conflict. Conflicts are coded as multiparty conflicts even if the rival rebel groups do not engage in fighting that reaches 1,000 battle-related deaths.

**guerrilla** type of warfare

0 = conflict was not fought using guerrilla warfare

1 = conflict fought using guerrilla warfare

**Source:** Kalyvas and Balcells (2010).

**guerrilla\_VHB** alternate measure of type of warfare

0 = conflict was not fought using guerrilla warfare

1 = conflict fought using guerrilla warfare

**Source:** alternate measure of the use of guerrilla warfare, using data from Valentino, Huth, and Balch-Lindsay (2004).

**contraband** rebel group reliance on contraband for financing

0 = rebel group does not rely on contraband for financing

1 = rebel group relies on contraband for financing

**Source:** Fearon (2004), extended to 2010.

**rcontbaid** rebel group reliance on contraband or aid from a foreign government

0 = rebel group does not rely on contraband or aid from a foreign government

1 = rebel group relies on contraband or aid from a foreign government

**Source:** contraband data from Fearon (2004), extended to 2010. Data on external aid to rebel groups measures whether the rebel group received material assistance (funding, weapons, or troops) from a foreign government in the first year of the conflict. Data are from the UCDP External Support data set (Hogbladh, Pettersson, and Themner 2011).

**rdemaid** rebel group received material aid from a democratic foreign government

0 = rebel group did not receive material aid from a democratic foreign government

1 = rebel group received material aid from a democratic foreign government

**Source:** data on external aid to rebel groups measures whether the rebel group received material assistance (funding, weapons, or troops) from a foreign democratic government in the first year of the conflict. Data on external aid are from the UCDP External Support data set (Hogbladh, Pettersson, and Themner 2011). Data on government level of democracy are from the Polity IV data set; governments with polity scores of +6 or greater in the year that aid was provided are coded as democratic governments (Marshall and Jaggers 2009).

## VI. Variables Measuring Rebel Group Characteristics and Behavior

**rinclusive** inclusive rebel group

0 = rebel group with exclusionary political objectives

1 = rebel group with inclusive political objectives

**Source:** Ethnic Power Relations (EPR) data set (Wimmer, Cederman, and Min 2009).

The EPR data set provides data on the aims of each of the rebel groups included in the UCDP/PRIO Armed Conflict Dataset. Drawing on the EPR coding of rebel group aims, *Inclusive Rebel Group* measures whether the rebel group has inclusive political objectives that “would pertain to the country’s entire population” or exclusionary political objectives that aim “to provide selective benefits for particular ethnic groups” (Wucherpfennig et al. 2012, 95).

**rgovern** rebel group possess governance institutions

0 = rebel group does not possess governance institutions

1 = rebel group possesses governance institutions

**Source:** secondary sources; full list of sources available from author. This variable “measures whether a rebel group created political structures to govern territory or populations under its control – structures aimed at providing public services such as adjudication of disputes, provision of humanitarian relief, maintenance of sanitation or transportation infrastructure, or certification of marriage or other contracts” (Stanton 2016, 75). See pp. 75-76 of the book for additional discussion of this variable.

**rconcentrated** rebel group’s civilian constituency is regionally concentrated

0 = rebel group’s civilian constituency is not regionally concentrated

1 = rebel group’s civilian constituency is regionally concentrated

**Source:** Minorities at Risk (MAR) Project (2007, 8), variable *GC2*.

A rebel group’s civilian constituency is coded as regionally concentrated (*rconcentrated* = 1) if the “rebel group’s ethnic base of support is located in one region of the country; in addition, 25 percent or more of the group’s population must live in the region and the group must constitute ‘the predominant portion of the population’ in the region” (Stanton 2016, 78).

The Ethnic Power Relations (EPR) data set was used to identify the “politically relevant ethnic category” for each rebel group (Wimmer, Cederman, and Min 2009, 326). Some rebel groups are not tied to any particular ethnic group; these groups are coded as *rconcentrated* = 0. For rebel groups receiving support from a particular ethnic group, *rconcentrated* captures the degree to which this civilian constituency is regionally concentrated. Data on the geographic concentration of ethnic groups are from the Minorities at Risk (MAR) Project (2007). The MAR data set includes a variable, *GC2*, measuring whether the ethnic group is located in: “A spatially contiguous region larger than an urban area that is part of the country, in which 25% or more of the minority resides and in which the minority constitutes the predominant proportion of the population.”

**rlobby** rebel group lobbied Western actors for support

0 = rebel group did not lobby Western actors for support

1 = rebel group lobbied Western actors for support

**Source:** U.S. Department of Justice database of public documents on file with the Foreign Agents Registration Act (FARA) Registration Unit, available at: <https://efile.fara.gov/pls/apex/f?p=145:10> and Unrepresented Nations and Peoples Organization (UNPO) Yearbooks (1995, 1996, 1997). *Rebel Group Lobbying (rlobby)* measures “whether the rebel group lobbied the U.S. government or joined the UNPO while the conflict was ongoing” (Stanton 2016, 102).

**rlobbyUS** rebel group engaged in political lobbying in the United States

0 = rebel group did not engage in political lobbying in the United States

1 = rebel group engaged in political lobbying in the United States

**Source:** U.S. Department of Justice database of public documents on file with the Foreign Agents Registration Act (FARA) Registration Unit, available at: <https://efile.fara.gov/pls/apex/f?p=145:10>. The Foreign Agents Registration Act (FARA), passed in 1938, requires that any representative of a foreign individual or group engaging in political or quasi-political activities in the United States register with the Department of Justice. Registration documents, publicly available in the database linked above, include the name of the foreign individual or group (the principal), the name of the individual or organization representing them in the United States (the agent), as well as details about the agent’s planned activities in the United States. See pp. 101-102 of the book for additional details on this variable, as well as an example from the FARA registration documents for the Patriotic Union of Kurdistan (PUK).

Variable measures whether the rebel group engaged in lobbying in the United States during any of the years that the civil war was active, using the war start and end dates recorded in this data set.

**runpo** rebel group membership in UNPO

0 = rebel group is not a member of UNPO

1 = rebel group is a member of UNPO

**Source:** historical data on membership in the Unrepresented Nations and Peoples Organization (UNPO) is available in UNPO Yearbooks, published annually (Unrepresented Nations and Peoples Organization 1995, 1996, 1997). These yearbooks were first published in 1995 and provide detailed entries for each group member, including information on the year in which the group joined UNPO.

**rebcentcontrol** rebel group central command structure, dichotomous

0 = rebel group does not possess a clear central command structure

1 = rebel group possesses a clear central command structure

**Source:** Non-State Actor Data (Cunningham, Gleditsch, and Salehyan 2009), *centcontrol* variable. According to the Non-State Actor Codebook, the *centcontrol* variable measures “whether the rebels have a clear central command.” (Cunningham, Gleditsch, and Salehyan 2012, 4).

**reborgcontrol**                      rebel group organizational control

0 = rebel group does not possess a clear central command structure

1 = rebel group central command exercises a low level of control

2 = rebel group central command exercises a moderate level of control

3 = rebel group central command exercises a high level of control

**Source:** Non-State Actor Data (Cunningham, Gleditsch, and Salehyan 2009), *strengthcent* variable. According to the Non-State Actor Codebook, the *strengthcent* variable measures “the extent to which a central command exercises control over the constituent groups of an insurgent movement.” (Cunningham, Gleditsch, and Salehyan 2012, 4).

**leftist**                                      rebel group has a leftist political objectives

0 = rebel group does not have leftist political objectives

1 = rebel group has leftist political objectives

**Source:** information on rebel group political objectives is from the conflict database created by the Uppsala Conflict Data Program (UCDP) in the Department of Peace and Conflict Research at Uppsala University, as well as from Jane’s World Insurgency and Terrorism database, available online at [www.janes.com](http://www.janes.com).

**religious**                                      conflict has a religious dimension

0 = conflict does not have a religious dimension

1 = conflict has a religious dimension

**Source:** information on rebel group political objectives is from the conflict database created by the Uppsala Conflict Data Program (UCDP) in the Department of Peace and Conflict Research at Uppsala University, as well as from Jane’s World Insurgency and Terrorism database, available online at [www.janes.com](http://www.janes.com).

Conflicts in which the rebel group’s stated political objectives have a religious dimension are coded as *religious* = 1, including groups with the following aims: to establish a government based on religious ideals; to gain independence or autonomy for a particular territory because of ethnic or national differences that include religious differences; or to gain independence or autonomy for a particular territory, at least in part, because of differences in opinion over the proper role of religion in the national government. Conflicts in which the rebel group represents an ethnic or national community whose primary religion differs from the primary religion of the government are also coded as *religious* = 1.

**Islamic**                      conflict involves differences over Islam

0 = conflict does not involve differences over Islam

1 = conflict involves differences over Islam

**Source:** information on rebel group political objectives is from the conflict database created by the Uppsala Conflict Data Program (UCDP) in the Department of Peace and Conflict Research at Uppsala University, as well as from Jane's World Insurgency and Terrorism database, available online at [www.janes.com](http://www.janes.com).

Conflicts in which the rebel group's stated political objectives involve Islam are coded as *Islamic* = 1 – for example, groups seeking to establish an Islamic government, or groups seeking independence or autonomy for a particular territory, at least in part, because of a desire to establish an Islamic government in the disputed territory. Also coded as *Islamic* = 1 are conflicts in which the rebel group represents an ethnic or national community whose primary religion is Islam and the government represents a community whose primary religion is not Islam.

## ***VII. Variables Measuring Government Characteristics and Behavior***

**ginclusive**                      inclusive government regime

0 = government is not politically inclusive

1 = government is politically inclusive

**Source:** Ethnic Power Relations (EPR) data set (Wimmer, Cederman, and Min 2009).

The EPR data set provides data on the extent to which different ethnic groups have access to political power in a given country over time. A government is coded as having an exclusionary regime (*ginclusive* = 0) if the EPR data set indicates that one or more ethnic groups has “dominant” or “monopoly” control. A government is coded as having an inclusive regime (*ginclusive* = 1) if the EPR data set indicates that no ethnic group has “dominant” or “monopoly” control.

**gconcentrated**                government has a concentrated constituency within the separatist region

0 = government does not have a concentrated civilian constituency within the separatist region

1 = government has a concentrated constituency within the separatist region

**Source:** primary and secondary sources, including national census data; full list of sources available from author. Variable is coded as *gconcentrated* = 1 if the government has a base of civilian support that constitutes between 5 and 50 percent of the population in the separatist region.

**polity** government's polity score in the first year of the conflict

**Source:** Polity IV Dataset (Marshall and Jaggers 2009).

If the conflict began prior to 1989, the polity score from 1989 is used.

If the polity score in the first year of the conflict is not available because the government is in a period of “interregnum” (−77), the polity score is coded as 0, according to the recommendation of the Polity IV data set creators.

The following cases are coded in this way:

Afghanistan (Rabbani government) v. Taliban, 1994-1996

Bosnia v. Croats, 1992-1994

Bosnia v. Serbs, 1992-1995

Congo v. RCD, 1998-2001

Congo v. MLC, 1998-2001

Côte D'Ivoire v. MPCI, MPIGO, MJP, FN, 2002-2004

Somalia v. USC faction led by Aideed, 1991-2002 (\*\*see additional notes on this case below\*\*)

Somalia v. ARS/UIC, 2006-ongoing

Somalia v. Al-Shabaab, 2008-ongoing

If the polity score in the first year of the conflict is not available because the government is in a period of “interruption” (−66), the polity score is coded as 0. The Polity IV data set recommends coding these cases as missing data since an “interruption” is intended to indicate that a foreign power is occupying the country and a new government has not yet been established. The two cases in the data set coded as “interruption” are Afghanistan and Iraq; while both of these countries still had significant foreign presences in the country at the start of their civil wars, in 2003 and 2004, respectively, both countries did have functioning national governments and national police and military forces. Thus, coding for these countries follows the procedure recommended for “interregnum” states, using a polity score of 0.

The following cases are coded in this way:

Afghanistan v. Taliban, 2003-ongoing

Iraq v. Al-Mahdi Army, 2004-2008

Iraq v. Ansar al-Islam, 2004-ongoing

Iraq v. ISI, 2004-ongoing

See additional notes below on the polity coding for specific cases.

**democracy** democratic government

0 = polity < 6

1 = polity ≥ 6

**autocracy** autocratic government

0 = polity > −6

1 = polity ≤ −6

**anocracy**                      anocratic government

0 = polity  $\geq$  6 OR polity  $\leq$  -6

1 = -6 < polity < 6

**avgpolity**                      government's average polity score for all years of the conflict, including the start and end years

**polityyr1**                      government's polity score in the first year of the conflict, including for conflicts that began prior to 1989

**xpolity**                          government's xpolity score in the first year of the conflict

**Source:** Vreeland (2008). In the Polity IV data, the incidence of political violence is factored into a country's polity score, which can be problematic when using a country's polity score as a predictor of civil war onset, as Vreeland (2008) points out. Thus, Vreeland constructs an alternate measure of democracy, *xpolity*, which removes the problematic political violence portions of the polity coding.

The main analyses in the book use the polity measure from the Polity IV data set. This is unlikely to present a problem for the analyses in the book, as all of the countries in the data set are experiencing civil wars. Any decline in polity score as a result of ongoing violence is likely to be similar across cases. However, robustness tests shown in the Appendix rerun the analyses using Vreeland's *xpolity* measure.

**govunstable**                      unstable government regime

0 = no government regime transition or coup attempt within five years prior to start of civil war

1 = government regime transition or coup attempt within five years prior to start of civil war

**Source:** data on regime transitions are from Marshall and Jagers (2009); data on coups are from Powell and Thyne (2011) and Marshall and Marshall (2014). This variable measures "whether the government experienced a change in regime or an attempted coup within the five years prior to the start of the conflict" (Stanton 2016, 76). The definition of regime change is from the Polity IV data set: a three or more point shift in polity score within a three-year period (Marshall and Jagers 2009). This variable includes attempted coups as well as successful coups; both types of coup incidents should affect the government's assessment regarding the stability of its regime.

**govunstable\_alt**                      unstable government regime, alternate measure

0 = no government regime transition or successful coup within five years prior to start of civil war

1 = government regime transition or successful coup within five years prior to start of civil war

**Source:** data on regime transitions are from Marshall and Jagers (2009); data on coups are from Powell and Thyne (2011) and Marshall and Marshall (2014). This variable is identical to the govunstable variable, except that it includes only *successful* coups. Thus, if a government experienced one or more coup attempts within five years prior to the start of the conflict, but none of these attempts was unsuccessful and the regime did not otherwise undergo a political transition, the case is coded as *govunstable\_alt* = 0.

**regimetransition**      government regime transition

0 = no regime transition within the five years prior to the start of the conflict

1 = regime transition within five years prior to the start of the conflict

**Source:** data on regime transitions are from Marshall and Jagers (2009). The definition of regime change is from the Polity IV data set: a three or more point shift in polity score within a three-year period (Marshall and Jagers 2009). Unlike the *govunstable* variables, the *regimetransition* variable relies solely on the polity coding to identify regime transitions. It does not incorporate attempted or successful coups.

**postransition**      positive regime transition

0 = no positive regime transition within the five years prior to the start of the conflict

1 = positive regime transition within five years prior to the start of the conflict

**Source:** data on regime transitions are from Marshall and Jagers (2009). The definition of regime change is from the Polity IV data set: a three or more point shift in polity score within a three-year period (Marshall and Jagers 2009). “Positive” regime transition indicates a transition in the direction of democracy on the polity scale, which ranges from -10 (most autocratic) to +10 (most democratic). New states – states that came into existence during the five years preceding the conflict – are also coded as *postransition* = 1 (e.g., Georgia, Moldova). Governments installed as part of an agreement to create a transition government following a prior civil war are also coded *postransition* = 1 if this transition took place within the five years prior to the start of the current conflict.

**pressfreedom**      degree of press freedom

0 = not free

1 = partly free

2 = free

**Source:** annual data on press freedom, published by Freedom House, available at [www.freedomhouse.org](http://www.freedomhouse.org).

**trade**      involvement in international trade

**logtrade**      involvement in international trade, logged

**Source:** World Bank, *World Development Indicators*, available online at: [data.worldbank.org](http://data.worldbank.org). Measure is the sum of total exports and total imports of goods and services as a percentage of GDP. Coded similarly to per capita GDP, for the year prior to the start of the conflict.

**igo** involvement in intergovernmental organizations (IGOs), total number

**Source:** Mansfield and Pevehouse (2006). Variable measures the total number of inter-governmental organizations (IGOs) of which the country was a member. Coded for the year the conflict began. Data runs through 2000, so for conflicts beginning after 2000, number of IGO memberships from 2000 is used.

**igosat** involvement in intergovernmental organizations (IGOs), ratio

**Source:** Mansfield and Pevehouse (2006), *saturation* variable. Variable measures the ratio of the number of inter-governmental organizations (IGOs) of which the country was a member to the number of IGOs of which the country *could* have been a member. Coded for the year the conflict began. Mansfield and Pevehouse data set runs through 2000, so for conflicts beginning after 2000, number of IGO memberships from 2000 is used.

**pro2ratification** ratification of Second Additional Protocol to the Geneva Conventions

0 = government had not ratified the Second Additional Protocol to the Geneva Conventions prior to the start of the conflict

1 = government ratified the Second Additional Protocol to the Geneva Conventions prior to the start of the conflict

**Source:** International Committee of the Red Cross (ICRC). A list providing the dates of ratification or accession for all states parties to the Second Additional Protocol is available on the ICRC website at: <https://ihl-databases.icrc.org/ihl/INTRO/475?OpenDocument>. Date of ratification or accession must be before the conflict start date (see discussion of *startyr* and *startmo* above) for a coding of *pro2ratification* = 1.

**USally** formal alliance with the United States

0 = government did not have a formal alliance with the United States while conflict was ongoing

1 = government had a formal alliance with the United States while conflict was ongoing

**Source:** Alliance Treaty Obligations and Provisions (ATOP) data set (Leeds et al. 2002).

## ***VIII. Additional Coding Notes***

### **Polity Coding Notes:**

#### ***Burundi v. CNDD, 1994-2003***

***Coded as polity = -5***

The Polity IV Data Set codes Burundi as -88 (“transition”) for March 1992-October 1993 and -77 (“interregnum”) for October 1993-July 1995. This period is coded as interregnum because of the extensive armed resistance to Buyoya’s government during this period. Although there was armed resistance to Buyoya’s rule, Buyoya remained in power during this period and was the clear government actor during the 1994-2003 conflict with the CNDD opposition. Thus, the Buyoya government during the 1994-2003 conflict is coded as having a polity score of -5, using the score assigned to Buyoya’s government in the year following the “interregnum” period (1996). The polity score for the year following the “interregnum” is used instead of the year prior to the “interregnum” because Buyoya made significant changes to the government in 1992, by introducing multiparty politics. The government’s polity score for 1991 reflects the government prior to the introduction of multiparty politics (polity= -7). The Buyoya government’s polity score for 1996 and 1997 more accurately reflects the character of the government during the 1992-1995 period.

#### ***Cambodia v. Khmer Rouge, KPNLF, and FUNCINPEC, 1978-1991***

***Coded as polity = -7***

The Polity IV Data Set codes Cambodia as -7 in 1978, the year that Vietnam invaded Cambodia and installed a government under Hun Sen. From 1979-1987, the period of civil war involving the most intense combat, the Polity IV data set codes Cambodia as -66 (“interruption”). From 1988-1993, the Polity IV data set codes Cambodia as -88 (“transition”). Although peace negotiations began in December 1987 (presumably why Polity IV codes the period from 1988-1993 as a “transition”), peace negotiations broke down several times and it was not until October 1991 that the parties to the conflict signed a peace agreement to resolve the conflict. Hun Sen remained in power throughout this period and continued to fight against the rebel groups when negotiations collapsed. While Vietnam did withdraw its troops in 1989, the character of Hun Sen’s government did not change; Freedom House gave Cambodia its lowest possible score (7) on both political rights and civil liberties throughout the conflict and did not change its ratings for the Cambodian government until 1991 with the signing of the peace agreement. The polity coding for the Cambodian government thus retains the polity score assigned to the Hun Sen government in the first year of its rule.

#### ***Congo-Kinshasa v. AFDL, 1996-1997***

***Coded as polity = -8***

The Polity IV Data Set codes Congo-Kinshasa as -77 (“interregnum”) for 1992-2003 because of the intense armed conflict during this period. Although there was armed resistance to Mobutu’s rule, Mobutu remained in power until being ousted by Laurent Kabila’s AFDL in 1997. Mobutu’s government was the clear government actor during the 1996-1997 conflict with the AFDL opposition. Mobutu had been in power since the 1960s, with polity scores of -9 for most of his years in power and polity scores of -8 in 1990-1992. The polity coding for the Congolese government during the 1996-1997 conflict thus uses the polity score assigned to Mobutu’s government in the year prior to the “interregnum” period (1991).

***Ethiopia v. Oromo/OLF, 1992-ongoing***  
***Ethiopia v. Ogaden/ONLF, 1994-ongoing***

***Coded as polity = 1***  
***Coded as polity = 1***

The Polity IV Data Set codes Ethiopia as -88 (“transition”) for 1993 and 1994 because in 1991 Mengistu’s government was overthrown by a coalition of rebel groups, which had been fighting throughout the 1980s; a period of transition followed the rebel victory. Although the OLF and ONLF had fought alongside other rebel groups to overthrow Mengistu, they continued to fight after the new government was installed because they felt the new government did not allow adequate representation (and autonomy) for particular ethnic groups. The government that was in place in 1992-1994 is the same government that remained in power after 1994. The polity coding for the Ethiopian government thus uses the polity score assigned to the government in 1995, the first year following the period of transition.

***Guinea Bissau v. military faction, 1998-1999***

***Coded as polity = 5***

The Polity IV Data Set codes Guinea Bissau as -77 (“interregnum”) for 1998 and -88 (“transition”) for 1999. The polity score for the years preceding the conflict is 5 and for the years following the conflict is 5. During the conflict, the Viera government, which had been in power since 1980, remained in power until Viera was ousted by the military junta at the end of the conflict. The polity coding for the Guinea Bissau government thus retains the polity score assigned to the government in the four years preceding the conflict and in the years immediately following the conflict.

***Liberia v. INPFL, 1990-1992***

***Coded as polity = -6***

The Polity IV Data Set codes Liberia as -6 for 1989 and -77 (“interregnum”) from 1990-1995. The INPFL began fighting in July 1990, while Samuel Doe was still president of Liberia. In September 1990, the INPFL killed Samuel Doe, leading to the installation of an interim government under Amos Sawyer and backed by ECOMOG troops. Polity IV codes the year 1990 as “interregnum” even though Doe was the acting president for much of the year. Because Doe was in power at the start of the conflict and during the period of greatest INPFL activity (from July-September 1990), the polity coding for the Liberian government thus retains the polity score assigned to the government in 1989.

***Somalia v. USC faction led by Aideed, 1991-2002***

***Coded as polity = 0***

The Polity IV Data Set codes Somalia as -77 (“interregnum”) from 1991-2010. In Somalia, after the fall of Siad Barre’s government in 1991, the central government collapsed and left multiple rebel groups fighting for control; no clear government actor existed. For this reason, I do not code government behavior toward civilians for this case and do not include this case as a government actor in the data set (the total number of government actors in the data set is thus 102). However, I do code the behavior of the main rebel group in this conflict, Mohammed Farrah Aideed’s faction of the United Somali Congress (USC). To allow for the inclusion of this case in analyses of rebel group behavior, I use a polity score of 0 for the period from 1991-2010, following the recommendation of the Polity IV data set creators for addressing cases of “interregnum”.

## References

- Cunningham, David E., Kristian S. Gleditsch, and Idean Salehyan. 2012. "Codebook for Non-State Actor Data," January. [http://privatewww.essex.ac.uk/~ksg/data/NSAEX\\_codebook.pdf](http://privatewww.essex.ac.uk/~ksg/data/NSAEX_codebook.pdf).
- Cunningham, David E., Kristian Skrede Gleditsch, and Idean Salehyan. 2009. "It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution* 53 (4): 570–97.
- Doyle, Michael W., and Nicholas Sambanis. 2006. *Making War and Building Peace: United Nations Peace Operations*. Princeton: Princeton University Press.
- Eck, Kristine, and Lisa Hultman. 2007. "One-Sided Violence Against Civilians in War: Insights from New Fatality Data." *Journal of Peace Research* 44 (2): 233–46.
- Fearon, James D. 2004. "Why Do Some Civil Wars Last So Much Longer than Others?" *Journal of Peace Research* 41 (3): 275–301.
- Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (01): 75–90.
- Gates, Scott, and Havard Strand. 2006. "Modeling the Duration of Civil Wars: Measurement and Estimation Issues." Working Paper. PRIO: Centre for the Study of Civil War.
- Gleditsch, Kristian S., and Michael D. Ward. 2001. "Measuring Space: A Minimum-Distance Database and Applications to International Studies." *Journal of Peace Research* 38 (6): 739–58.
- Gleditsch, Kristian Skrede. 2002. "Expanded Trade and GDP Data." *Journal of Conflict Resolution* 46 (5): 712–24.
- Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand. 2002. "Armed Conflict 1946-2001: A New Dataset." *Journal of Peace Research* 39 (5): 615–37.
- Heger, Lindsay, and Idean Salehyan. 2007. "Ruthless Rulers: Coalition Size and the Severity of Civil Conflict." *International Studies Quarterly* 51 (2): 385–403.
- Hogbladh, Stina, Therese Pettersson, and Lotte Themner. 2011. "External Support in Armed Conflict, 1975-2009: Presenting New Data" paper presented at International Studies Association Annual Meeting, Montreal, Canada.
- Kalyvas, Stathis N., and Laia Balcells. 2010. "International System and Technologies of Rebellion: How the End of the Cold War Shaped Internal Conflict." *American Political Science Review* 104 (03): 415–29.
- Lacina, Bethany, and Nils Petter Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population* 21 (2): 145–66.
- Leeds, Brett Ashley, Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long. 2002. "Alliance Treaty Obligations and Provisions, 1815-1944." *International Interactions* 28: 237–60.
- Mansfield, Edward D., and Jon C. Pevehouse. 2006. "Democratization and International Organizations." *International Organization* 60 (01): 137–67.
- Marshall, Monty G., and Keith Jagers. 2009. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2007." Center for Systemic Peace.
- Marshall, Monty G., and Donna Ramsey Marshall. 2014. "Coup D'Etat Events, 1946-2013: Codebook." Center for Systemic Peace. [www.systemicpeace.org](http://www.systemicpeace.org).
- Minorities at Risk Project. 2007. "Minorities at Risk (MAR) Codebook Version 2/2009." College Park, MD: Center for International Development and Conflict Management.
- Powell, Jonathan M., and Clayton L. Thyne. 2011. "Global Instances of Coups from 1950 to 2010 A New Dataset." *Journal of Peace Research* 48 (2): 249–59.

- Raleigh, Clionadh. 2012. "Violence Against Civilians: A Disaggregated Analysis." *International Interactions* 38 (4): 462–81.
- Stanton, Jessica A. 2009. "Strategies of Violence and Restraint in Civil War." New York: Columbia University.
- . 2016. *Violence and Restraint in Civil War: Civilian Targeting in the Shadow of International Law*. New York: Cambridge University Press.
- Stockholm International Peace Research Institute (SIPRI). multiple years. *SIPRI Yearbook: Armaments, Disarmament and International Security*. New York: Oxford University Press.
- The International Institute for Strategic Studies (IISS). multiple years. *The Military Balance*. The International Institute for Strategic Studies.
- Unrepresented Nations and Peoples Organization. 1997. *Unrepresented Nations and Peoples Organization: Yearbook 1996*. The Hague: Kluwer Law International.
- Valentino, Benjamin, Paul Huth, and Dylan Balch-Lindsay. 2004. "'Draining the Sea': Mass Killing and Guerrilla Warfare." *International Organization* 58 (02): 375–407.
- Vreeland, James Raymond. 2008. "The Effect of Political Regime on Civil War." *Journal of Conflict Resolution*, 52 (3): 401–25.
- Wimmer, Andreas, Lars-Erik Cederman, and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." *American Sociological Review* 74: 316–37.
- Wucherpfennig, Julian, Nils W. Metternich, Lars-Erik Cederman, and Kristian Skrede Gleditsch. 2012. "Ethnicity, the State, and the Duration of Civil War." *World Politics* 64 (1): 79–115.