# Non-State Actors in Civil Wars: A New Dataset

David Cunningham, University of Maryland & Peace Research Institute Oslo

Kristian Skrede Gleditsch, University of Essex & Peace Research Institute Oslo


Idean Salehyan, University of Texas

Abstract: This manuscript introduces the Non-State Actors (NSA) in Civil Wars dataset, which contains detailed information on the state-rebel group dyads included in the Uppsala Conflict Data Project Dyadic Dataset. Existing quantitative studies generally focus on characteristics on countries and conflicts to examine the duration, severity, outcome, and recurrence of civil wars, in ways which often ignore the actors in civil wars. The NSA data provide additional information on the organizations involved in conflict dynamics. We describe the structure of the NSA data and the variables included, provide descriptive statistics of the indicators, and discuss areas for future research on non-state actors to enhance our understanding of conflict processes.

For the past two decades or so, the quantitative study of civil war has seen a tremendous increase in scholarly output. Largely driven by the end of Cold War-era preoccupation with international conflict and devastating violence in countries such as Sudan, Sri Lanka, and Colombia, the field has seen significant progress in terms of well-articulated theories and sophisticated research designs. However, much of the literature seeks to understand *where* and *when* do conflicts break out to the exclusion of other important questions such as, who are the actors involved and how do they pursue their aims? In other words, scholars have generally employed country/years as their units of analysis in seeking to explain why some countries experience the outbreak of civil war while others do not. To use one example, understanding why Angola has experienced a bloody war while Zambia has not is indeed an important question. However, current empirical analyses have said little about the motives, tactics, and organizational capacity of the Angolan opposition.

Focusing on country/years as units of observation has yielded several insights, but also comes with limitations. For instance, Collier and Hoeffler (2004) and Fearon and Laitin (2003), two of the most widely cited articles in the civil war literature, focus on how macro-level characteristics of states predict to conflict onset or incidence. The former argues that abundant natural resources in the economy as a whole provide rebels opportunities to extract wealth and finance their operations. The latter study argues that factors that favor insurgency, such as rugged terrain and poor infrastructure, provide opportunities for rebel mobilization and conflict onset. Yet both of these studies make inferences about the behavior of rebel organizations using data at a high level of aggregation; namely, country-level GDP figures, percentage of territory covered by mountains, oil revenues, and so on. More recently, a new wave of research has disaggregated the study of civil war by examining where conflicts occur at the subnational level

(Buhaug and Rød 2006; Wucherpfennig et al. 2011; Buhaug and Gates 2002). Dissatisfied with the coarse resolution in country-level studies, this body of work looks at how local-level covariates such as population, terrain, and distance from borders, influence the location of conflict or individual battle events within the state.

While the question of *where* civil wars occur is indeed important, we argue that it is equally important to assess *who* participates in conflict. One approach to this question is to use individual-level survey data to examine who is inclined to join rebel groups (Humphreys and Weinstein 2008). In this article, we take a somewhat different approach and focus on the characteristics of rebel organizations. Recently, there has been growing attention to the behavior of opposition organizations and the relationships between them (Pearlman and Cunningham 2012), although empirical studies have been rather limited because of lack of comprehensive or high-quality data. We present a data resource on attributes of rebel groups—the Non-State Actors in Armed Conflict Dataset (NSA)—which contains variables on rebel strength, territorial control, organizational structures, and external support, among other factors. While previous studies have examined primarily the characteristics of governments or country-level economic and demographic factors, the NSA data allow scholars to incorporate critical information on the opposition side as well.[1]

In the following section, we discuss the motivation and rationale behind collecting these data. Then, we turn to the coding procedures and variables included in the analysis. Next, we present descriptive statistics, demonstrating a wide range of variation in the organizational

---

[1] Specifically, we describe Version 3.4 of the NSA data, which is updated through 2011 and lines up directly with the Uppsala Conflict Data Project Dyadic Dataset. Previous articles (Cunningham et al. 2009, Salehyan et al. 2011) have used earlier versions of the NSA data. Here, we present the new and updated dataset in detail, and discuss possibilities for combining data on non-state actors across various data projects.

characteristics of rebel organizations, and discuss potential applications of the data. The final section offers concluding remarks about how the NSA data can help advance scholarship on civil war.

**Motivation**

Data on civil war onset and duration has typically employed the country/year as the unit of analysis. For instance, the Correlates of War intrastate conflict data (COW, see Sarkees and Wayman 2010) and the Uppsala Conflict Data Project/Peace Research Institute Oslo Armed Conflict Dataset (ACD, see Gleditsch et al. 2002, Themner and Wallensteen 2012), define a civil war as an organized, violent contest between a state and a non-state actor over some political objective such as control over the government or secession that generates battle deaths exceeding some threshold. These two data resources focus on the violent conflicts as aggregate events with start and end dates, but use slightly different criteria for inclusion: the COW data require at least 1,000 battle deaths in a 12-month period for an incident to qualify as a war, while the ACD employs a much lower threshold of 25 battle deaths per year to identify events. Most studies of wars and internal armed conflicts rely on binary indicators (zero or one) of whether or not a country experiences a conflict in a given year. Studies of conflict onset use the initial outbreak of fighting as the dependent variable, while studies of incidence do not distinguish between the initial onset year and subsequent years of ongoing conflict, and studies of duration look at how long fighting lasts before some mode of termination (e.g. victory, settlement).

As useful as these data resources are for empirical analysis, theories of violent civil conflict usually focus on how government and opposition actors strategically interact with one another and with the population at large. Borrowing insights from studies of international

conflict, civil war scholars have argued that factors such as the relative capabilities of actors, their resolve, the informational environment, and ability to make credible commitments, should affect bargaining and resort to violence (Lake and Rothchild 1996; Walter 2009). Mobilization theories analyze how rebels secure resources and support among the wider population and overcome the collective action problem, all within the shadow of government violence (Lichbach 1995; Weinstein 2007; Wood 2000). Kalyvas (2006) argues that patterns of rebel and government control over territory affect the intensity and type of violence during civil wars. Finally, some have begun to look at civil wars involving more than two parties and how strategic interactions between rebels and the government, as well as between various rebel factions, shape conflict outcomes (Cunningham 2006, 2011; Cunningham et al. 2012; Kydd and Walter 2002; Nilsson 2008; Stedman 1997). Each of these theoretical perspectives brings the focus squarely on the characteristics of rebel organizations, their strategies, and the constellation of actors during civil war. Thus, there is often a mismatch between empirical studies of civil conflict—which focus on where and when civil wars occur—and theoretical perspectives which focus on the nature of the combatants involved.

For this reason, we believe that research into the dynamics of civil war has been seriously hampered by the lack of systematic data on rebel organizations. As opposed to data on country-level attributes and features of the state—for example, government spending, size of military forces, and political institutions—quality data on dissident groups is inherently more difficult to gather. Rebel organizations are often clandestine and do not report official statistics. Moreover, data collection projects such as Correlates of War and the Polity IV Project on democratic institutions have provided a wealth of information about government resources and institutions, but little information about non-state actors.

This lack of data has seriously impeded progress in the field. For instance, studies of civil war duration and outcome distinguish between rebel victory, government victory, and negotiated settlements. Mason, Weingarten and Fett (1999) look at the covariates of various conflict outcomes, while DeRouen and Sobek (2004) employ a competing risks analysis of termination type. While both of these studies have been quite influential, neither includes variables accounting for the attributes of rebel organizations. They have relied entirely upon data on governments and the national-level context of the war. Similarly, studies of civil war recurrence (e.g. Quinn, Gurses, and Mason 2007; Walter 2004), examine the mode of conflict termination and the broader political/economic environment, without taking into account who actually fought during the war. Given data constraints, this shortcoming is understandable, but we argue that a more complete picture includes information on factors such as rebel strength, territorial control, and capacity to procure arms and other resources (see Cunningham, Gleditsch, and Salehyan 2009). Attempting to understand why rebels or governments win in civil wars is indeed difficult without data on who the rebels are.

To address this gap in the current literature, we set out to gather systematic information on rebel organizations. Important information such as rebel strength, arms procurement capacity, leadership characteristics, and external sponsorship were gathered from primary and secondary sources. Rather than characterize civil wars as two-party contests between rebels and governments, this data collection allows us to distinguish between various competing factions in a conflict (see Cunningham 2011). Multiple rebel groups fighting over the same incompatibility may be aligned or collaborate against the government, but such organizations are often competitors and may spend as much time fighting one another as contesting the government. For instance, rather than a single rebellion in the Philippine province of Mindanao, our data

5

collection allows us to parse out the different characteristics of the Moro National Liberation Front, the Moro Islamic Liberation Front, and Abu Sayyaf. While these groups may share some common goals, they vary on several important dimensions such as size and transnational support. In the following section, we describe the data collection process in more detail.

**Structure of the Dataset**

The NSA data extends the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al. 2002, Themner and Wallensteen 2012) and encompasses all internal armed conflicts over the period 1945-2011. In order for a conflict to be coded as an internal armed conflict in the ACD, it must meet five general criteria—the conflict must (1) involve the government of the state, (2) take place primarily within the state, (3) involve organized opposition forces, (4) be fought over either control of the government or territory, and (5) generate 25 battle deaths in a calendar year. The ACD itself is not dyadic, but identifies all rebel organizations operating in opposition to the government in each civil war in each calendar year. The UCDP Dyadic Dataset divides all of the observations in the ACD into state-rebel group dyad years. For a dyad to be coded as active in a year there must be at least 25 battle-related deaths in the dyad.

The NSA data uses all of the dyads from the UCDP Dyadic Dataset. As we describe below, we offer a wider set of informational variables for each group than the UCDP data. While the UCDP is annual, the NSA data is structured so that the unit of analysis is state-rebel group dyad spell. The data reflect changes in rebel characteristics over time, and may be easily converted to other forms such as annual observations. We identify a new dyadic conflict spell if there is a period exceeding two calendar years in which the level of conflict falls below 25 battle deaths in a year. To illustrate the data structure, Table 1 presents the UCDP representation of the

actors in the Burundian civil war of 1991-2008 and Table 2 presents the NSA data for this

period.

**Table 1: UCDP Dyadic Dataset Coding for Burundi Civil War (1991-2008)**

| DyadID | ConflictID | Location | SideA | SideB | YEAR |
|---|---|---|---|---|---|
| 11 | 90 | Burundi | Burundi | Palipehutu | 1991 |
| 11 | 90 | Burundi | Burundi | Palipehutu | 1992 |
| 12 | 90 | Burundi | Burundi | CNDD | 1994 |
| 12 | 90 | Burundi | Burundi | CNDD | 1995 |
| 12 | 90 | Burundi | Burundi | CNDD | 1996 |
| 12 | 90 | Burundi | Burundi | CNDD | 1997 |
| 12 | 90 | Burundi | Burundi | CNDD | 1998 |
| 13 | 90 | Burundi | Burundi | Frolina | 1997 |
| 14 | 90 | Burundi | Burundi | CNDD–FDD | 1998 |
| 14 | 90 | Burundi | Burundi | CNDD–FDD | 1999 |
| 14 | 90 | Burundi | Burundi | CNDD–FDD | 2000 |
| 14 | 90 | Burundi | Burundi | CNDD–FDD | 2001 |
| 14 | 90 | Burundi | Burundi | CNDD–FDD | 2002 |
| 14 | 90 | Burundi | Burundi | CNDD–FDD | 2003 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 1997 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 1998 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 1999 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2000 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2001 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2002 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2003 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2004 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2005 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2006 |
| 15 | 90 | Burundi | Burundi | Palipehutu–FNL | 2008 |

**Table 2: NSA Coding of Burundi Conflict**

| obsid | ucdpid | dyadid | side_a | side_b | startdate | enddate |
|---|---|---|---|---|---|---|
| NSA.3.4-649 | 90 | 11 | Burundi | Palipehutu | 11/26/91 | 12/31/92 |
| NSA.3.4-652 | 90 | 12 | Burundi | CNDD | 11/26/94 | 5/8/98 |
| NSA.3.4-655 | 90 | 13 | Burundi | Frolina | 1/1/97 | 12/31/97 |
| NSA.3.4-658 | 90 | 14 | Burundi | CNDD-FDD | 6/7/98 | 12/31/03 |
| NSA.3.4-661 | 90 | 15 | Burundi | Palipehutu-FNL | 1/1/97 | 12/31/08 |

We use the state-rebel group dyad spell (as opposed to the state-rebel group dyad year) as observations may extend for periods both longer and shorter than a calendar year and many features that we measure do not change on an annual basis. However, there are attributes in our data that do change over time and when one of these variables does change we code a new spell for the group. As an example, if a rebel group was receiving support from an external state and then lost that support, there would be a change in the coding of the external state variables and so a new period would be created. Hence, each row in the NSA data represents a spell or period of conflict between a state and rebel group dyad where the coding of the variables remain the same. If a user is interested, this approach can easily be converted to annualized records for combatant groups, allowing the researcher to combine the information with other data collected on an annual basis.

The current version of the NSA data is coded to line up with the 2012 version of the UCDP Dyadic dataset. The first version of the NSA data was created prior to the first release (v1.0) of the UCDP Dyadic Dataset in 2008, and was based simply on the list in the main conflict table in the original UCDP ACD data. The UCDP Dyadic Data involved extensive changes in the actors identified, which have now been adopted in the NSA data. The current version NSA data also include the UCDP Conflict ID and Dyad ID code, which make other

datasets that code information on these dyads easily compatible with the NSA data. One of the barriers to aggregation of data on civil war as compared to interstate war has been the lack of standard ID codes to allow for merging of data across different data collection efforts. By assuring that our data line up with the UCDP Dyadic Data, we address this problem and allow for greater aggregation of data on civil war and non-state actors.

**Indicators**

For each dyad in the NSA data, we code a series of variables. The first set of indicators measures the attributes of rebel groups (often coded relative to the government) and the second set of indicators measures the transnational dimension of the conflict—whether the government or rebels receive external support, from whom, and of what type. We discuss each variable in the dataset here briefly, with reference to some examples, and there is a larger codebook which discusses the coding of each in depth. In addition, we have compiled extensive documentation which describes the rational for coding the variables in the data and identifies the sources used. We make this documentation available to researchers.

While the coding notes give much more detail about the sources used to code specific conflicts, we briefly discuss the major sources used here. For several of the variables included in the NSA data, the UCDP Conflict Encyclopedia was an invaluable resource—it contains information on the number of troops possessed by rebel groups and the government on an annual basis as well as external support to rebels and states. The database only has information on conflicts from 1975, however, and so we used other sources to code dyads prior to 1975 as well as other variables for which there was not information in the database. We primarily used Keesing's Record of World Events, news reports found through searches in Lexis-Nexis

academic, and secondary academic sources about individual conflicts and countries to code this information.

The first two variables in the NSA measure the extent to which rebel groups are affiliated with political organizations. While rebel groups are primarily military organizations, many of them cooperate either directly or indirectly with political actors to push for policy change. For example, it was widely known that the Irish Republican Army had a close relationship with Sinn Fein. The first variable (*Rebpolwing*) measures whether the rebel group has a political wing. It takes the following values—"explicit link," "acknowledged link," "alleged link," and "none." Among rebel groups that do affiliate with a political organization, there is variation in how the government responds to those political wings. Governments often ban those organizations, while others allow them to operate. For those groups that have some sort of political wing, we code whether or not that political wing is legal in the state in which the group operates. That variable (*Rebpolwinglegal*) takes the values of "yes" or "no."

One of the most crucial tasks for rebel groups is to acquire and maintain the ability to challenge the government militarily. We code a number of variables that refer broadly to the ability of the rebels to do so. The first set of these variables measure the number of troops possessed by the rebel group. The variable *Rebestimate* provides the best estimate of the number of troops under the command of the organization. Because troop estimates can vary dramatically across sources, we have two additional variables (*Rebestlow* and *Rebesthigh*) that measure the lowest and highest estimates found, respectively. Having these different estimates allows researchers to see whether any impact of rebel troop strength is robust to variation in the numbers reported.

These measures of troop strength are absolute, in that they measure the total number of troops commanded by the rebels. However, in civil wars, the military capacity of rebels is best thought of in relation to that of the government—a group with a relatively small number of troops might be a serious threat to a weak government while a comparably sized group would be outmatched by a much more powerful state. To measure the relative comparison of the groups to the government, we include a composite ordinal measure (*Rebstrength*) which indicates whether the rebel group is much weaker, weaker, at parity, stronger, or much stronger than the government.

In addition to this composite measure, we code three additional relative indicators. The first (*Mobcap*) measures the ability of the rebel group to mobilize popular support relative to the government. Rebel groups that can mobilize or draft greater numbers of fighters should be better able to present a clear threat to governments and engage in direct attacks, and this ability varies across rebel organizations. Many anti-colonial movements, for example, could rely on large levels of popular support (as compared to the colonial power) and had the potential to mobilize significant numbers of fighters even though they were out-matched militarily.

The second (*Armsproc*) indicates how capable the rebels are at procuring arms, relative to the government. While some rebel organizations have access to high-tech weaponry, others are hampered in their ability to secure weapons, and this variation should have a direct effect on their ability to challenge the government militarily.

The third variable (*Fightcap*) indicates how capable the rebels are at fighting, relative to the government. Some rebel organizations may be quite small in terms of number of troops but quite effective in their fighting ability. Other rebel groups, meanwhile, have large numbers of troops but lack the fighting capacity to effectively coordinate and target the state.

Each of these three variables is ordinal, and measured as low, moderate, and high—indicating the ability of the rebels, relative to the government. These variables measure different components of the ability of rebels to effectively target the state. As an example, consider the Rwandan Patriotic Front (RPF) during the Rwandan civil war of 1990-1994. The RPF was made up primarily of Rwandan Tutsi refugees who had been living in Uganda for decades, and had little popular support within Rwanda, where the Hutu represented a large majority. As such, the organization is coded as "low" on *mobcap*. The RPF had access to military equipment, as did the Rwandan government, and so the group is coded as "moderate" on *armsproc*. Many members of the RPF, however, had spent years engaged in guerrilla warfare in Uganda and had helped Yoweri Museveni to overthrow the government of Milton Obote. Partly as a result of this experience, and partly due to clear organization and a command structure, the RPF was a very effective fighting force, and proved more than a match for the Rwandan army. In fact, with a similar number of troops, the RPF was much more able to conduct effective military operations than the Rwandan army. The RPF is coded as "high" on the *fightcap* variable.

The next two variables measure the extent to which the rebel group has a command structure that controls its troops. The first variable (*Centcontrol*) is a dichotomous measure of whether the organization in fact has a central leadership at all (coded as "yes" or "no"). While most rebel organizations have some formal command structure, there are a handful of groups that really exist as disparate factions with very little leadership. The Armed Conflict Dataset, for example, identifies a handful of groups simply as "insurgents" (examples include "Kashmiri Insurgents," "Patani Insurgents," and "Sikh Insurgents") because it is difficult to identify actual organizations in these conflicts. We code these rebels as not having clear leadership.

The second variable (*Strengthcent*) is coded for those groups that are identified as having central control. It is an ordinal measure of the extent of control exercised by the leadership, measured as low, medium, and high. Among those groups with a central command there is considerable variation in how much control the leadership actually exercises over the day-to-day activities of the organization. In the Eritrean conflict in Ethiopia, for example, the leadership of the Eritrean People's Liberation Front (EPLF) clearly controlled and coordinated the activities of the organization. This level of control was evident as the organization relatively easily transitioned into the government of Eritrea once it gained independence from Ethiopia. Other insurgent organizations in Eritrea, meanwhile, such as the Eritrean Liberation Front (ELF) had command structures that had considerably less direct control over the workings of the organization.

Another important type of variation among rebel groups is that some rebel groups control territory in the country and others do not. The NSA data codes three variables for each dyad indicating rebel territorial control. The first variable (*Terrcont*) is a dichotomous measure of whether the rebel organization controls territory at all. The second and third variables are only coded for cases where a rebel organization controls territory. *Terrname* gives the name of the territory to the degree of specificity that we could identify[2] and *Effterrcont* is an ordinal measure indicating whether the extent of control exercised by the organization over the territory is low, medium, or high. Even among those organizations that control territory, the degree of control they exercise varies. Some rebel groups, such as the Movement for the Liberation of Congo during the Congolese civil war of 1998-2002, have been able to establish administrative control

---

[2] In some cases, this variable indicates the name of a specific province or piece of territory. In other cases, it indicates a general area such as "parts of northern Iraq."

and collect taxes. Others, such as the Rally for Congolese Democracy during that same war, have nominal control over some territory, but that control is heavily contested.

Rebel groups vary greatly in terms of their ability to challenge the government militarily, whether they operate through political means or exclusively through fighting, the degree to which the leadership of the organization actually coordinates the activity of its members, and whether and how much control over territory within the civil war state the group exercises. Each of these is likely to affect the dynamics of conflict between the government and the rebels. The variables present in the NSA data allow researchers to evaluate the effect of these various factors.

In addition to these measures, the NSA data also include a series of variables that capture the transnational dimension to the conflict. The focus is on whether and to what degree the rebel group and the government receive support from external actors. While the term "civil war" would seem to imply some sort of domestic process, there is a significant literature that has demonstrated that external influences on civil war are profound (Gleditsch 2007, Salehyan 2009, Salehyan and Gleditsch 2007). Much of the empirical work in this literature, however, focuses on the role of external actions in the conflict generally, and does not examine directly the relationship between external actors and the government and rebels involved (for a discussion, see Salehyan et al. 2011).

The NSA data allow for examining these relationships more directly. For the rebels, we include variables measuring different aspects of the transnational dimension. The first (*Transconstsupp*) measures whether the group had a transnational constituency—in other words, it makes appeals to ethnic or religious groups in other states—and whether this constituency supplied support. It is an ordinal measure, coded as "no," "tacit," or "explicit." The second

variable (*Rebextpart*) measures whether the rebels received support or foreign fighters from external non-state actors and takes the following values—"no," "minor," or "major." For instance, organizations that were part of the Afghan Mujahedin received material support from private individuals in wealthy Arab states and the Irish Republican Army received funds from the Irish community in the United States.

The next set of variables measures whether the rebel organization receives assistance from external states. The first variable (*Rebpresosts*) measures whether the rebel group operates to at least some extent out of another country's territory. While we generally think about rebel groups as internal actors, many rebels—such as the Ugandan rebel group Lord's Resistance Army which is based primarily in southern Sudan and eastern DRC—operate out of another country's territory. This variable is trichotomous, taking values of 'no," "some" (minor and periodic use of external territory) and "extensive" (major, long-term presence in external territory). Another variable (*Presname*) is only coded for cases where *Rebpresosts* is coded as "some," or "extensive," and lists the countries from which the organization operates.

In addition to allowing rebels sanctuary, external states can provide more direct assistance to rebels. For instance, the Pakistani government is known to provide material support to several rebel organizations operating in India. The variable *Rebsuport* measures whether the organization receives support from one or more external states and is coded as "no," "alleged," and "explicit." The variable *Rtypesuport* indicates, for those groups receiving support, what type of support they receive—"none," "endorsement," "non-military," "military," or "troops." Because groups may receive multiple different types of support, only the highest level of support is listed. Finally, the variable *Rsupname* lists the states identified as providing support to the rebel organization. Afghan groups fighting the USSR during the 1980s, for example, are coded

as receiving "explicit," "military" support from the U.S., and in some cases, Iran, Pakistan, and Saudi Arabia.

While our data focuses on the rebel organizations, we also recognize that governments in civil wars receive external support and that this external support can have significant effects on conflict dynamics. In the Rwandan civil war, for example, the RPF would have almost certainly defeated the government in February 1993 if French troops had not intervened to block the RPF from taking Kigali. As such, we have parallel measures for the government with many of the rebel support variables. The variable *Govsuport* measures whether the government receives support from another state, again coded as "no," "alleged," and "explicit." The variables *Gtypesup* and *Gsupname* indicate the type of support and the name of the supporting states, respectively. The variable *Govextpart* measures whether the government receives support from non-state actors, coded as "no," "alleged," or "explicit." The Rwandan government is coded as receiving "explicit" "military" support from "France and Belgium" during the Rwandan civil war.

These variables, then, give detailed information about the transnational dimension of civil conflicts, which has been shown to have significant influence on conflict dynamics. Combined with the indicators of the attributes of rebel groups, they provide a wealth of information on the characteristics of state-rebel group dyads in conflict. Each of these variables is potentially time-varying, in that these attributes can change over the course of conflicts. Rebels can lose (or gain) control over territory, external support, leadership can falter, these organizations can expand or contract in size, and so on. As such, we code these variables as time-varying where possible and when characteristics of the state-rebel group dyad change we code a new period in the dyad.

In addition to these variables measuring attributes of rebel groups and their and the government's transnational support, we also have two variables that indicate the relationship of each dyad to other dyads in the same conflict. In some cases, the UCDP identifies new dyads because there has been a period of inactivity followed by a resumption in fighting. In others (such as the URNG in Guatemala and the FMLN in El Salvador), a new dyad is identified when previously independent rebel organizations join together to form a new organization. In the opposite case, there are examples of rebel groups splintering so that multiple factions continue as separate dyads (as happened with the Rally for Congolese Democracy in the conflict in the DRC). Because we are aware that researchers may want to consider the previous history of a conflict dyad in analyzing the impact that characteristics of rebel organizations have on its dynamics, we code a variable (*Prevactive*) that is a dichotomous measure of whether this dyad has been previously active and an additional variable (*Prevact.ref*) that indicates the observation ID code for all other periods that the dyad was active.

The NSA, then, has a wealth of information on various characteristics of all rebel groups active in civil wars for the last 65 years. In the next section, we provide descriptive statistics for some of these characteristics, which help illustrate the depth and utility of our data.

**Attributes of State-Rebel Group Dyads**

The NSA data contain information on 477 state-rebel group dyads over 578 distinct observation periods or spells.[3] The data reveal substantial variation among these groups on the indicators included in the dataset. Because of the number of variables in the NSA data, we do not

---

[3] Periods change as important attributes of the rebel organization change; thus, one "period" refers to a phase in the group's organizational variables.

have the space to provide descriptive statistics of all of them here; rather, we focus on some of the more interesting variation.

One of the key variables in the NSA data is the dyadic measure of the strength of rebel groups, relative to the state. Much of the literature on inter-state conflict has focused on how differences in material capabilities between pairs of states affect their likelihood of engaging in violence or the course that violence takes. Studies of civil war often proxy the balance of power in the conflict with a measure of government military strength (such as military personnel, military expenditure, or the Composite Indicator of National Capabilities), but such a monadic measure does not actually measure the concept of relative strength. Table 3 shows the distribution of cases on our rebel strength measure, divided into 5 categories.

**Table 3: Rebel Group Strength (Composite Measure)**

|  | **Number of Dyads** | **Percentage** |
|---|---|---|
| **Much Weaker** | 174 | 36.71% |
| **Weaker** | 231 | 48.73% |
| **At Parity** | 45 | 9.49% |
| **Stronger** | 18 | 3.8% |
| **Much Stronger** | 6 | 1.27% |
| **Total** | 474 | 100% |

Note: Some observations have missing values, and the percentages refer to cases with valid observations.

As can be seen from Table 3, the overwhelming majority of rebel groups in our data are weaker than their governments, with over 85% being identified as either "much weaker," or "weaker." This should not be surprising given that, particularly in the post-World War II period, civil war has been primarily a form of "asymmetric conflict" between states and rebels. Despite this, however, we do find a noticeable set of cases (9.5% of the total) where rebel groups are at parity with the government, and more than 5% of the time we code individual rebel organizations

as being stronger or much stronger than the state. This variation allows for exploring the effect that rebel group strength has on various aspects of the conflict process.

The rebel strength measure is primarily a composite indicator, but, as described previously, the NSA data contain a number of component measures as well. In particular, we include measures that identify the ability of rebels to mobilize support, procure arms, and fight, relative to the government. Table 4 presents descriptive statistics for these measures.

**Table 4: Rebel Group Strength (Component Measures)**

|  | Mobilization Capacity | Arms Procurement | Fighting Capacity |
|---|---|---|---|
| **Low** | 253 (55.6%) | 288 (64.6%) | 289 (63.2%) |
| **Moderate** | 168 (36.9%) | 145 (32.5%) | 149 (32.6%) |
| **High** | 34 (7.5%) | 13 (2.9%) | 19 (4.2%) |
| **Total** | 415 (100%) | 410 (100%) | 407 (100%) |

Note: Some observations have missing values, and the percentages refer to cases with valid observations.

Table 4 shows that, as would be expected, the majority of rebel groups are weaker than the government on each of these components. However, it again shows significant variation, with approximately 1/3rd of rebel groups being measured as equivalent to the government in their ability to mobilize support, procure arms, and to fight, and a small percentage higher. These results are interesting when we consider that the overall measure of strength described above (which is based to a significant extent on the troop ratio between the rebel group and the government) indicates that the vast majority of rebel groups are weaker than the government, but these composite measures show that almost half are at least equal to the government in their ability to mobilize, procure arms, and fight. This suggests that, to a significant extent, the balance-of-power in favor of the government in many of these conflicts is due to the

government's ability to field a larger fighting force (due to possessing a standing army) more than to its advantage in these individual measures.

A third element of the state-rebel group interaction that we consider here is the degree to which rebel groups control territory in the country.  Our data show that 271 310 rebel groups (approximately 2/3) do not control territory, while 131 158 (approximately 1/3) do. Among those groups that do control territory, the most common level of control exercised is "moderate," with 85 92 groups (60.7%) coded that way. Nineteen 25 rebel groups are found to have a "high" degree of control (13.6%) and 373 (23.64%) have a "low" degree of control.

The NSA data show, then, that rebel groups differ in terms of their strength, relative to the government, and whether and to what degree they control territory within the civil war state. Each of these factors could have a profound effect on the dynamics of conflict. Strong rebel groups that have the ability to mobilize support, procure arms, and control territory pose a direct threat to the government, while weaker groups with little ability to fight and who do not control territory do not. This does not mean that these weaker groups will be irrelevant, but the degree to which the government sees rebels as a threat and is willing to grant concessions to them, for example, is likely to be affected by the relative strength between them.

In addition to these internal characteristics, a significant body of literature suggests that the external dimensions of civil war matter as well. While many studies of international and transnational dimensions of civil war have been conducted, most of these studies look at international interaction at the conflict level. However, the degree to which external states support specific actors in conflict should influence the dynamics of those conflicts.

Table 5 presents statistics on the degree to which rebels and governments receive support from external states and the type of support received.

**Table 5: External State Support to Rebels and Governments**

| | Rebels | Governments |
|---|---|---|
| **External State Support** | | |
| None | 229 (51.69%) | 223 (50%) |
| Alleged | 41 (9.26%) | 5 (1.12%) |
| Explicit | 173 (39.05%) | 218 (48.88%) |
| Total | 443 (100%) | 446 (100%) |
| **Type of Support** | | |
| non-military | 28 (13.4%) | 21 (9.42%) |
| military | 158 (75.6%) | 152 (68.16%) |
| troops | 23 (11%) | 50 (22.42%) |
| Total | 209 (100%) | 223 (100%) |

Note: Percentages refer to cases with valid observations, and cases with external support for percentages on type.

The descriptive statistics in Table 5 reveal that almost half of rebel groups and exactly half of governments in our data receive either alleged or explicit support. However, it is not the case that governments and rebels always receive support in the same conflicts. Indeed, a cross tab comparing two dichotomous variables measuring whether the government received alleged or explicit support and whether the rebels did shows that in ~~195~~ 146 dyads neither received support, in ~~163~~ 135 both received support, in ~~115~~ 84 dyads only the government received support, and in ~~97~~ 78 only the rebels did. Thus, while it is quite common for both (or neither) the government or rebels to have support from an external state, there are a significant number of cases where only one of the two sides receives support.

Table 5 also shows that the vast majority of cases of support in our data are military support (either military assistance or the actual deployment of troops by the external state). We do identify cases of non-military support (e.g. finances, logistics, medical supplies), but it is much rarer. Remember, however, that we code the highest level of support for each dyad, so that

in many of the cases where we identify rebels and governments as receiving military support from an external state they are likely receiving other types of support as well.

The tables presented above show significant variation across these measures of state and rebel group characteristics. One advantage of our data over much of the work that has been conducted is that our data allow for variation within conflicts in the features of rebel groups. While civil war is often treated as a process involving only two actors—a state and a rebel group—in reality, internal conflicts show significant variation in the number of combatants involved (Cunningham 2006, 2011). Our data allow for examining how variation in the characteristics of different rebel groups involved in the same conflict affect the dynamics of those dyadic conflicts, as compared to other groups in the same conflict.

As an example, Table 6 returns to the Burundian example from Table 2 and shows variation in some of the indicators from the NSA data among the groups in that conflict. For some of the variables described above (mobilization capacity, arms procurement, fighting capacity, territorial control, and government support) we do not show variation among the rebel groups in the Burundian civil war. In others, however, there is variation.

**Table 6: Variation among rebels in the Burundian Civil War**

| Rebels | Start | End | Rebel Political Wing? | Legal Wing? | Rebel Strength | Rebel External Support | Type of Support |
|--------|-------|-----|------------------------|-------------|----------------|-------------------------|-----------------|
| Palipehutu | 26/11/1991 | 31/12/1992 | explicit link | no | weaker | no | no |
| CNDD | 26/11/1994 | 8/5/1998 | explicit link | no | weaker | explicit | military |
| Frolina | 1/1/1997 | 31/12/1997 | explicit link | no | much weaker | no | no |
| CNDD-FDD | 7/6/1998 | 31/12/2003 | no | NA | weaker | explicit | military |
| Palipehutu-FNL | 1/1/97 | 31/12/2008 | no | NA | weaker | no | no |

Indeed, Table 6 shows that, while all of the rebels in the Burundian civil war were weaker than the government, one of the groups (Frolina) was much weaker. Additionally, three groups (Palipehutu, Frolina, and Palipehutu-FNL) did not receive support from external states, while two others (CNDD and CNDD-FDD) did receive military support. Finally, we see that while three of the rebel groups had political wings, two of them—CNDD-FDD and Palipehutu-FNL, which were the military wings of CNDD and Palipehutu, respectively, and broke away from their political organizations during a peace process in Arusha, Tanzania in 1997—did not.

This intra-conflict variation is important because we know empirically that the dynamics of conflict can be quite different among different rebel groups fighting over the same incompatibility. Even within the same conflict, some rebel groups fight much longer than others, some sign negotiated agreements and others do not, and so on. Identifying variation in the attributes of groups within conflicts allows for analyzing their dynamics more closely.


**Conclusion**

The study of civil war has made significant progress in recent years, but our understanding of conflict processes has been hindered by insufficient attention to the attributes of the actors involved. While the cross-national study of interstate war generally is based on dyadic data and examines how characteristics of states affect whether they go to war with each other, how long these wars last, how they end, and so on, studies of civil war generally do not examine how attributes of states and rebels affect similar dynamics in those conflicts. The NSA data help address this gap by providing a variety of data on the actors involved in civil wars.

While the data we present here are important, and allow for testing the effect of a number of characteristics of these actors on civil war dynamics, they also open up the opportunity for

further data collection and standardization across data projects. One of the ways that the study of

interstate war has advanced significantly is through the aggregation of data—different

researchers working on different data projects have compiled information on various

characteristics of states, and combining information from multiple datasets is quite easy.[4] While

little data has been collected on non-state actors, the various datasets that do exist are not easily

combined because there is not a unified list of these actors in conflicts.

To address this issue, at least partly, the NSA data uses the Conflict ID and the Dyad ID

from the UCDP Dyadic Dataset. This means that if other researchers collect data on the actors in

the Dyadic Dataset, and include the Conflict ID or Dyad ID in their data, different data projects

on state-rebel groups can be combined. In particular, other data projects such as the Ethnic

Power Relations (EPR) data (Cederman et al.) can be combined with the ACD data through the

so-called ACD2EPR docking project, which links the specific non-state actor organizations to

EPR's list of ethnic groups (see Wucherpfenning et al. 2012). This aggregation allows for the

advancement of research on the dynamics of civil war in a way that to date has been very

difficult to do consistently.

We encourage other researchers to add to these data by collecting additional information

on rebel group characteristics.  For instance, given that many rebel groups split internally, it

would be useful to collect information on internal divisions and leadership disputes within

groups.  Additionally, it would be useful to collect basic demographic information on

participation in rebel movements, including the use of child soldiers, the recruitment of women,

and the class or occupational background of fighters.  Finally, additional external variables could

---

[4] This is particularly true due to the availability of software such as Eugene (Bennett and Stam 2004)

be included such as indictments by the International Criminal Court against government and

rebel leaders as well as scrutiny by non-governmental organizations.  By providing a list of

groups and a common, integrated platform, extensions to the NSA can be incorporated easily to

the existing data.

## Works Cited

Bennett, D. Scott and Allan Stam (2004). *The Behavioral Origins of War*. Ann Arbor: University of Michigan Press.

Buhaug, Halvard and Jan Ketil Rød (2006). Local determinants of African civil wars, 1970-2001. *Political Geography* 25(3): 315-335.

Buhaug, Halvard and Scott Gates (2002). The Geography of Civil War. *Journal of Peace Research* 39(4): 417-433.

Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics* 62 (1): 87-119.

Collier, Paul and Anke Hoeffler (2004). Greed and Grievance in Civil War. *Oxford Economic Papers* 56(4): 563-95.

Cunningham, David E. (2006). Veto Players and Civil War Duration. *American Journal of Political Science* 50(4): 875-92.

Cunningham, David E. (2011). *Barriers to Peace in Civil Wars*. Cambridge: Cambridge University Press.

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): 570-597.

Cunningham, Kathleen Gallagher, Kristin Bakke and Lee Seymour (2012). Shirts Today, Skins Tomorrow: Dual Contests and the Effects of Fragmentation in Self-Determination Disputes. *Journal of Conflict Resolution* 56(1): 67-93.

Derouen, Karl R. and David Sobek (2004). The Dynamics of Civil War Duration and Outcome. *Journal of Peace Research* 41(3): 303-320.

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

Gleditsch, Kristian Skrede. (2007). Transnational Dimensions of Civil War. *Journal of Peace Research* 44(3): 293-309.

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.

Humphries, Macartan and Jeremy M. Weinstein (2008). Who Fights? The Determinants of Participation in Civil War. *American Journal of Political Science* 52(2): 436-455.

Kalyvas, Stathis N. (2006). *The Logic of Violence in Civil Wars*. Cambridge: Cambridge University Press.

Kydd, Andrew and Barbara F. Walter (2002). Sabotaging the peace: The politics of extremist violence. *International Organization* 56(2): 263-296.

Lake, David A. and Donald Rothchild (1996). Containing Fear: The Origins and Management of Ethnic Conflict. *International Security* 21(2): 41-75.

Lichbach, Mark Irving (1995). *The Rebel's Dilemma*. Ann Arbor: University of Michigan Press.

Mason, T. David, Joseph P. Weingarten Jr., Patrick J. Fett (1999). Win, Lose, or Draw: Predicting the Outcome of Civil Wars. *Political Research Quarterly* 52(2): 239-268.

Pearlman, Wendy and Kathleen Gallagher Cunningham. (2012). Nonstate Actors, Fragmentation, and Conflict Processes. *Journal of Conflict Resolution.* 56(1): 3-15.

Quinn, J. Michael, T. David Mason & Mehmet Gurses (2007). Sustaining the Peace: Determinants of Civil War Recurrence. *International Interactions* 33(2): 167-193.

Salehyan, Idean and Kristian Skrede Gleditsch. (2006). Refugees and the spread of civil war. *International Organization* 60(2): 335-66.

Salehyan, Idean (2009). *Rebels without Borders*. Ithaca, NY: Cornell University Press.

Salehyan, Idean, Kristian Skrede Gleditsch and David E. Cunningham (2011). Explaining External Support for Insurgent Groups. *International Organization* 65(4): 709-744.

Sarkees, Meredith Reid and Frank Wayman (2010). *Resort to War: 1816-2007*. Washington, DC: CQ Press.

Stedman, Stephen J. (1997). Spoiler Problems in Peace Processes. *International Security* 22(2): 5-53.

Themner, Lotta and Peter Wallensteen (2012). Armed Conflicts, 1946-2011. *Journal of Peace Research* 49(4): 565-575.

Walter, Barbara F. (2004). Does conflict beget conflict? Explaining recurring civil war. *Journal of Peace Research* 41(3): 371-388.

Walter, Barbara F. (2009). Bargaining Theories of Civil War. *Annual Review of Political Science* 12: 243-261.

Weinstein, Jeremy M. (2007). *Inside rebellion: The politics of insurgent violence*. Cambridge: Cambridge University Press.

Wood, Elizabeth (2000). *Forging Democracy from Below: Insurgent Transitions in South Africa and El Salvador*. Cambridge: Cambridge University Press.

Wucherpfennig, Julian, Nils Metternich, Lars-Erik Cederman, and Kristian Skrede Gleditsch. (2012). Ethnicity, the State, and the Duration of Civil Wars. *World Politics* 64 (1):79-115.

Wucherpfennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman and Andreas

  Wimmer (2011). Politically Relevant Ethnic Groups across Space and Time: Introducing

  the GeoEPR Dataset. *Conflict Management and Peace Science* 28(5): 423-437.