UCDP Actor Dataset Codebook
Version 23.1

Uppsala Conflict Data Program

Department of Peace and Conflict Research
Uppsala University

This version compiled and updated by Therese Pettersson (2023)

When using this data, please always cite:

Davies, Shawn, Therese Pettersson & Magnus Öberg (2023). Organized violence 1989-2022 and the return of conflict between states. *Journal of Peace Research* 60(4).

When appropriate, also cite this codebook: Pettersson, Therese (2023) UCDP Actor Dataset Codebook v 23.1 (https://ucdp.uu.se/downloads/).

Always include the Version number in analyses using the dataset.

Contents

Introduction

This codebook presents the variables in the UCDP Actor Dataset, and clarifies the definitions.

Actor-specific information is also available at www.ucdp.uu.se. Questions regarding the definitions and the content of the dataset can be directed to ucdp@pcr.uu.se

The unit of analysis in the UCDP Actor Dataset is a unique actor. The dataset is constructed so as to be compatible with all UCDP's datasets on organized violence (the UCDP Geo-referenced Events Dataset, the UCDP/PRIO Armed Conflict Dataset, the UCDP Dyadic Dataset, the UCDP Battle-related Deaths Dataset, UCDP One-sided Violence Dataset and the UCDP Non-state Conflict Dataset) and thus contains all actors – both state – and non-state ones – included in these datasets. As such, it functions as an inventory of actors in UCDP data, but its scope is broader than that. The UCDP Actor Dataset also ascribes unique actor IDs to all actors, identifies which types of organized violence the actor has been involved in (state-based, non-state and/or one-sided) and, more particularly, in which conflicts and dyads of state-based, non-state and/or one-sided violence.

Furthermore, the dataset contains information about the name history of the actor, outlining the most current name (which is also the name used in the data), the original name (i.e. the name the group had when it entered UCDP data for the first time) and any possible name changes between the two. In addition, the dataset also includes information on the group's origins and alliances. First, information is given about whether the non-state actor was created by splintering off from another actor recorded in UCDP data, and, if so, which actor. Second, information is included on whether the actor is created by a temporary split in the original actor. Third, the dataset also includes data on whether the non-state actor ever entered into an alliance with another recorded non-state actor in UCDP data (which would beget a new non-state actor), as well as the name of said alliance. Finally, information is given about whether a non-state actor has joined a group already registered in UCDP data, not creating a new alliance but merely entering the other group. For non-state actors, information can also be found on temporary coalitions i.e. when non-state actors fight together in non-state conflict or one-sided violence without creating a formal alliance.

UCDP's three categories of organized violence

Below follows a brief introduction to the three types of organized violence that UCDP collects information on, and that make up the datasets to which the UCDP Actor dataset corresponds.

State-based armed conflict

An armed conflict is defined as a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths.

Data on this category of organized violence is found in the UCDP/PRIO Armed Conflict Dataset (presently covering 1946-2022), in the UCDP Dyadic Dataset (1946-2022), in the UCDP Geo-referenced Events Dataset (1989-2022), and in the UCDP Battle-related deaths Dataset (1989-2022). For more information, see the codebooks for these datasets, at https://ucdp.uu.se/downloads/).

One-sided violence

One-sided violence is defined as the use of armed force by the government of a state or by a formally organized group against civilians, which results in at least 25 deaths. Extrajudicial killings in custody are excluded.

Data on this category of organized violence is found in the UCDP One-sided Violence Dataset (1989-2022), and in the UCDP Geo-referenced Events Dataset (1989-2022). For more information, see the codebook for the dataset, at https://ucdp.uu.se/downloads/).

Non-state conflict

A non-state conflict is defined as the use of armed force between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year.

Data on this category of organized violence is found in the UCDP Non-state Conflict Dataset (1989-2022), and in the UCDP Geo-referenced Events Dataset (1989-2022). For more information, see the codebook for the dataset, at https://ucdp.uu.se/downloads/).

Definition of a state actor

A state actor is always the government of a state.

A state is an internationally recognized sovereign government controlling a specific territory or an internationally unrecognized government controlling a specified territory whose sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory.

When in doubt as to who is actually controlling the government of a state, the party controlling the capital of the state is regarded as the government.

In order to define the total population of states we use the Gleditsch & Ward (1999) systems membership definition, which is based on the Correlates of War project. Gleditsch & Ward include countries with a population of more than 250,000 that have 'a relatively autonomous administration over some territory', and is 'considered a distinct entity by local actors or the state it is dependent on' (Gleditsch & Ward, 1999: 398).

The Gleditsch & Ward definition differs from our original definition in two main ways. First, for Gleditsch & Ward a state is considered to be a new entity if it is 'considered a distinct entity by local actors' while the conflict definition only recognizes it as a new entity when the 'sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory'. Second, the Gleditsch & Ward states all have a population of more than 250,000 while our conflict definition does not prescribe a specific size of the population. In a few cases, countries with less than 250,000 inhabitants experience conflict or are active as secondary parties in a conflict. In those cases we use the Gleditsch & Ward tentative list of microstates. For example, the microstate Tonga was part of the multinational coalition fighting in Iraq in 2004.

Definition of a non-state actor

The UCDP Actor Dataset contains information on all actors/ groups/organisations that appear in UCDP's three categories of organized violence. Both the terminology and the type of organisations included in the three categories vary to some degree. The UCDP/PRIO Armed Conflict Dataset and the UCDP Dyadic Dataset includes "opposition actors" and "opposition organisations", defined as any non-governmental group of people having announced a name for their group and using armed force to influence the outcome of the stated incompatibility. The UCDP One-sided Violence Dataset contains "formally organized groups", similarly defined, apart from the requirement of an incompatibility. Lastly, the Non-state Conflict category in the UCDP Actor Dataset includes both "formally organized groups" and "informally organized groups". While the former is the same category as the one used in the One-sided Violence data, the latter is defined as any group without an announced name, but who uses armed force against another similarly organized group. To be included there has to be a clear pattern of violent incidents that are connected and in which both groups use armed force against the other.

The informally organized group are divided into two categories, members/supporters of political parties; and broader categories of identifications, such as ethnic groups and clans or religious groups.

The formally organized group category encompasses actors such as unrecognized states and rebel groups.

As outlined above, the UCDP Actor Dataset thus contains information on a broad spectrum of non-state actors. The smallest common denominators are that the non-state actor is an entity comprised of several individuals and that it has some form of organisational structure (be it formal, as in e.g. rebel groups or more traditional, as in e.g. ethnic groupings). Non-state actors included in the dataset hence range from small, informal ones such as the Dizi ethnic group in Ethiopia, numbering some 20 000 individuals, to large, state-like formations such as the self-declared Republic of Croatia (before it was recognised as an independent state).

Rules guiding the differentiation between non-state actors

To differentiate between non-state actors is at times problematic.[1] When is a non-state actor completely new and when is it simply a continuation of an already registered non-state actor, only slightly altered? Some actors tend to be fractious, splitting into different sub-groups while others join larger coalitions or umbrella groups. UCDP follows four main rules:

- However many times a group changes its name, it retains the same Actor ID (for exceptions, see below).
- When an actor splits into two or more actors, the splinter group/groups is considered a completely new non-state actor.
  Comment: It is sometimes difficult to determine which of the actors should be viewed as the splinter and which should be seen as the original one, as both factions tend to claim to be the "original" actor. This is most typically the case concerning rebel groups. UCDP deals with this by tracing the main part (i.e. the most numerous) of the actor, which is then coded as the original one. Should this be problematic to determine, a second option is to trace the original leader of the non-state actor in order to identify which actor to code as the original one.
- When two or more already registered non-state actors join together under a new name in a tight coalition with joint military operations, the new alliance is regarded as a new actor.
- When an already registered non-state actor is joined by another actor that has previously not been coded as active in the dataset, the new alliance is simply viewed as a continuation of the former, previously coded, actor. This applies even when the name of the actor is changed.
  Comment: It should be noted that while this rule is applied in most cases in the datasets conventional logic made it necessary to make some exceptions. Whereas the rule works well in e.g. the case of NRA in Uganda, which was originally called PRA but changed its name when it was joined by the much smaller and militarily insignificant UFF, the logic is less clear in the case of e.g. the Independent Nasserist Organisation in Lebanon (a.k.a. al- Mourabitoun), which was active in fighting the government of Lebanon in 1958 and also resurfaced as part of the Lebanese National Movement in 1975. In this case it is intuitively and factually incorrect to characterise the LNM as being a simple continuation of al- Mourabitoun, and the actor is thus characterised as being an entirely new one.

---

[1] This problem is mainly associated with the formally organized non-state actors.

Variables in the Actor Dataset

Variable name	Content	Type
ActorId	The unique identifier of all actors	Integer
NameData	The name that is used for the actor in a UCDP dataset. The name used for an actor in UCDP data is either its current one or the last name it had when registered in a UCDP dataset. For formally organized non-state actors, such as militias or rebel groups, the names listed in this field are by default acronyms when applicable.	String
NameOrig	The name of the actor when it was first included in a UCDP dataset.	String
NameOrigFull	The full original name of the actor, in mother tongue.	String
NameOrigFullEng	The full original name of the actor, in English	String
NameChange	A binary variable that codes whether the actor has changed its name. If the actor has changed its name, this variable is coded as 1. If not, a 0 is coded.	Bolean
NewName	The new name of the actor	String
NewNameFullMotherTongue	The full new name in mother tongue	String
NewNameFullEng	The full new name in English	String
Org	This variable indicates the organizational level of the actor. The level of organization is determined according to the following categories: **Organizational level 1** (formally organized groups): Rebel groups and other organized groups that have a high enough level of organization so as to be possible to include in the state-based armed conflict category. These include rebel groups with an announced name, as well as military factions (Forces of…). **Organizational level 2** (informally organized groups): Groups composed of supporters and affiliates to political parties and candidates. These are commonly not groups that are permanently organized for combat, but who at times use their organizational structures for such purposes. In addition to supporters of political parties and candidates, included in this category is also fighting between groups composed of supporters of other organizations such as the supporters of al-Ahly football team fighting against the supporters of al-Masry football team in Egypt 2012.	Integer

Organizational level 3 (informally organized groups):

Groups that share a common identification along ethnic, clan, religious, national or tribal lines. These are not groups that are permanently organized for combat, but who at times organize themselves along said lines to engage in fighting.

Organizational level 4 (states):

The actor is the government of a state

ConflictId	The unique identifier of all state-based armed conflicts (i.e. conflicts included in the UCDP/PRIO Armed Conflict Dataset) in which the actor has been recorded. Comma-separated if multiple	String
DyadId	The unique identifier of all dyads active in state-based armed conflicts (i.e. dyads included in the UCDP Dyadic Dataset) in which the actor has been recorded. Comma-separated if multiple	String
PrimaryParty	A binary variable that codes whether the actor has been a primary party in a state-based armed conflict. If the actor has been a primary party, this variable is coded as 1. If not, a 0 is coded. This variable relates solely to actors active in state-based armed conflicts.	Bolean
OSID	The unique identifier of all dyads listed in the UCDP One-sided Violence Dataset.	Integer
OSCoalition	A binary variable listing whether the actor has been active in one-sided violence together with one or several other actors in the dataset. If the actor has been part of a temporary coalition in one-sided violence, this variable is coded as 1. If not, a 0 is coded.	Bolean
OSCoalitionID	The unique identifier of all temporary coalitions listed in the UCDP One-sided Violence Dataset. Comma-separated if multiple	String
NSID	The unique identifier of all non-state dyads (i.e. non-conflicts included in the UCDP Non-state Conflict Dataset) in which the actor has been recorded. Comma-separated if multiple	String
NSCoalition	A binary variable listing whether the actor has been active in non-state conflict together with one or several other actors in the dataset. If the actor has been part of a	Bolean

	temporary coalition in non-state conflict, this variable is coded as 1. If not, a 0 is coded.	
NSCoalitionID	The unique identifier of all temporary coalitions listed in the UCDP Non-state Conflict Dataset. Comma-separated if multiple	String
Splinter	A binary variable listing whether a non-state actor was created by breaking away from another actor listed in UCDP data.	Bolean
NamePrev	Name of the actor that a given non-state actor broke away from	String
ActorIdPrev	The unique identifier of the actor that a given group broke away from	Integer
SplitTemp	A binary variable listing whether a non-state actor was created by a temporary split in the original movement. Unlike the actors coded as created in the variable Splinter, those formed in SplitTemp are of a temporary nature. Often they may still view themselves as being part of the original group, but divisions within it has caused fighting between different factions or groupings. An example of this is fighting within the Taleban movement in Afghanistan. In 1996, two Taleban commanders and their followers fought each other over control of territory. The two factions, Taleban - Ali Dad faction and Taleban - Mola Khel faction, are registered in the UCDP Actor Dataset as separate actors created from a temporary split from the Taleban. The split is considered temporary as the commanders, after the fighting had ceased, continued to be part of the Taleban movement.	Bolean
NameSplitTemp	Name of the actor that a given non-state actor temporarily broke away from	String
ActorIdSplitTemp	The unique identifier of the actor that a given group temporarily broke away from	Integer
Alliance	A binary variable listing whether a non-state actor entered into an alliance with another non-state actor, also registered in UCDP data, thus creating a new non-state actor	Bolean
NameAlliance	Name of the alliance created by the merger of two or more non-state actors	String
ActorIdAlliance	The unique actor identifier of the new non-state actor (alliance) created by the merger of two or more non-state actors	Integer

JoinGroup	A binary variable listing whether a non-state actor joined a group already registered in UCDP data.	Bolean
	While this variable is closely related to the Alliance variable, there is a fundamental difference between the two. When an actor is coded in Alliance, it ceases to exist on its own, as it together with another registered actor creates a new organisation.	
	When an actor is coded in Join_group on the other hand, it ceases to exist because it joins, or is overtaken by another already existing group. An example of this is the case of the two Congolese rebel groups Ninjas and Ntsiloulous. When the Cobra militia, together with Angolan support, managed to oust the sitting president Pascal Lissouba (supported by the Cocoyes) and his Prime Minister Bernard Koleleas (supported by the Ninjas), the Ninjas hid and regrouped in the Pool region. In a parallel development, a new group surfaced in Pool; the Ntsiloulous. The Ntsiloulou leader Pasteur Ntumi was able to attract a large following amongst the Ninjas who thought that by following Ntoumi's orders they were carrying out the will of God, and subsequently began increasing his military force. Over time more and more Ninjas joined the Ntsiloulous – some by their own free will and others under the threat of force – until, eventually, the Ninjas had ceased to exist.	
GroupName	Name of the non-state group that the given non-state actor joined	String
ActorIdGroup	The unique actor identifier of the group that a given non-state actor joined	Integer
<mark>Location</mark>	The countries in which the actor has been active. Comma-separated if multiple	String
GWNOLoc	The Gleditsch and Ward code for the countries in which the actor has been active. Comma-separated if multiple	String
Region	Identifies the region of the location 1 = Europe (GWNo: 200-399) 2= Middle East (GWNo: 630-699) 3= Asia (GWNo: 700-999) 4= Africa (GWNo: 400-626) 5= Americas (GWNo: 2-199).	String
Version	The version of the dataset: 23.1	Float

Version name convention

In 2017, the version name convention was changed, giving all UCDP datasets the same version number across the board. This was done so that users more easily can see which UCDP dataset corresponds with which.

This codebook corresponds to Version 23.1 of the UCDP Actor Dataset. We recommend that whenever this dataset is used, the version number should be cited.

The version number is a combination of a year and a number. The year refers to when the dataset is updated with new observations. If there are changes in the data between yearly updates, or if there are substantial changes in the structure of the dataset, the number behind the year is incremented.

A note on UCDP ID changes

In version 17.1 of all UCDP datasets, the ID system for conflicts, actors and dyads was changed in order to make them unique across all UCDP core datasets and all UCDP types of violence. This allows easier aggregation and disaggregation of data as well as simplify data management for users, especially when combining multiple UCDP products together.

For example, a non-state conflict cannot have the same ID as a (different) state-based conflict or as a case of one-sided violence (which was possible before).

Further, actor/side IDs representing a government were decoupled from their corresponding Gleditsch and Ward country codes (GWNo). Gleditsch and Ward country codes for state sides/actors are provided in separate columns from their sides ID (i.e. side_a_id is different from gwno_a but represents the same entity).

This means all IDs used in the dataset from version 17.1 and onwards are no longer compatible with those in older versions of UCDP products. Further, this means that external products relying on UCDP IDs for data management tasks will have to be adapted to work with the new ID systems.

A conversion table between the new and old ID systems is available here: http://ucdp.uu.se/downloads/.

References

Gleditsch, Kristian S. & Michael D. Ward. 1999. "Interstate System Membership: A Revised List of the Independent States since 1816." International Interactions 25: 393-413

Davies, Shawn, Therese Pettersson & Magnus Öberg (2023). Organized violence 1989-2022 and the return of conflict between states. *Journal of Peace Research* 60(4).