

UCDP Non-State Conflict Codebook¹

Version 1.0 – September 28, 2005

*Joakim Kreutz
& Kristine Eck*

in collaboration with

*Peter Wallensteen,
Lotta Harbom,
Stina Högladh,
& Margareta Sollenberg*

Uppsala Conflict Data Program
Department of Peace and Conflict Research, Uppsala University

Introduction

This document describes the Non-State Conflict Dataset, a project within the Uppsala Conflict Data Program (UCDP) at the Department of Peace and Conflict Research, Uppsala University. Funding for the UCDP Non-State conflict project is provided by the Human Security Centre at the University of British Columbia.

Definition of Non-State conflict

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

The separate elements of the definition are operationalized as follows:

¹ The UCDP Non-State conflict project was started in 2002 by the Uppsala Conflict Data Program. This is the first version of the Codebook. When using the data, please cite UCDP Non-State conflict Dataset, the address of the UCDP website, and (when appropriate) this codebook. Please always include the Version number in analyses using the dataset.

- (1) *Use of armed force*: use of arms in order to promote the parties' general position in the conflict, resulting in deaths.
 - (1.1) *Arms*: any material means, e.g. manufactured weapons but also sticks, stones, fire, water, etc.
- (2) *25 deaths*: a minimum of 25 battle-related deaths per year and per warring dyad
 - (2.1) *battle-related deaths*: deaths directly related to combat between the warring parties
- (3) *Organized groups*: must be either
 - (3.1) a formally organized group: any non-governmental group of people having announced a name for their group and using armed force, *or*
 - (3.2) any group who does not have an announced name, but who uses armed force and whose violent activity meets at least one of the following organizational requirements:
 - (3.2.a.) there must be a clear pattern of incidents which are connected, *or*
 - (3.2.b.) there must be evidence that violence was planned in advance
- (4) *State*: a state is
 - (4.1) an internationally recognized sovereign government controlling a specified territory, *or*
 - (4.2) an internationally unrecognized government controlling a specified territory whose sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory.
- (5) *Government*: the party controlling the capital of the state

Summary of variables in the Dataset

No	Variable	Label	Description
1	Dyad ID	Dyad identifier	The unique identifier of all active non-state dyads.
2	Location	Country name	The name of the country where the actor responsible for the one-sided violence is based.
3	Side A	Name of party	The name of the Side A in the conflict.
4	Side B	Name of party	The name of the Side B in the conflict.
5	Year	Year of observation	
6	Fat_best	Estimate of deaths	UCDP Best estimate of deaths caused by fighting between the two parties.
7	Fat_low	Estimate of deaths	UCDP Low estimate of deaths caused by fighting between the two parties.
8	Fat_high	Estimate of deaths	UCDP High estimate of deaths caused by fighting between the two parties.
9	Cnfl_active Side A	(dummy)	Whether Side A also was involved in an armed conflict in the same year according to the UCDP-PRIO dataset
10	Cnfl_active Side B	(dummy)	Whether Side B also was involved in an armed conflict in the same year according to the UCDP-PRIO dataset
10	COW_location ²	COW numbers of location	COW numbers of location country
11	Region	Region of Location	The region of the location; UCDP classifies the world into 5 regions.
12	Version	Version number	The current version of the dataset.

² COW= Correlates of War project

Non-State conflict Dataset

Dyad ID. This variable makes it possible to follow the same dyad, that is, the conflict activity between two parties over several years. Thus, each dyadic relationship in the dataset is given a unique Actor ID.

Location. The country where the parties involved in the conflict is based. While some incidents may occur in an adjoining country, the majority of conflict activity is reported in the location given.

Side A. The formally organized group that constitute Side A in the conflict. For each dyad the parties are listed in alphabetical order. For a directory of abbreviations of group names, see http://www.pcr.uu.se/research/UCDP/HumSec_index1.htm.

Side B. The formally organized group that constitute Side B in the conflict. For each dyad the parties are listed in alphabetical order. For a directory of abbreviations of group names, see http://www.pcr.uu.se/research/UCDP/HumSec_index1.htm.

Year. The calendar year of observation. UCDP employs a threshold for inclusion of at least 25 battle-related deaths for each dyad in a calendar year. Cases which do not meet the threshold are coded as absence of non-state conflict. This does not mean that the dataset is free from uncertainties; for a discussion of unclear cases, see http://www.pcr.uu.se/research/UCDP/HumSec_index1.htm

Best, Low and High fatality estimates.

The general rule for UCDP's estimation of fatalities is moderation. All incidents are ideally verified by two independent sources, and all estimates reported are grounded in UCDP's expertise of each particular country. Due to the varying certainty of fatality reports, the project provides three estimates concerning battle-related deaths for each year.

- (a) **Best estimate.** The best estimate is constructed by aggregating the best estimates for each individual event. If reports provide conflicting estimates, an examination is made as to what source is most reliable. If no such distinction can be made, UCDP employs the lower figure given.
- (b) **Low estimate.** The low estimate is constructed by aggregating the low estimates for each individual event. If different reports provide different estimates and a higher estimate is considered more reliable, the low estimate is also reported if deemed reasonable.
- (c) **High estimate.** The high estimate is constructed by aggregating the high estimates for each individual event. If different reports provide different estimates and a lower estimate is considered more or equally reliable, the high estimate is also reported if deemed reasonable. If there is uncertainty about exactly which party was involved in a given incident, it may also be included in the high estimate.

Conflict Active Side A. This variable records if Side A is involved in a state-based armed conflict – according to the UCDP-PRIO dataset – in the same year.

Conflict Active Side B. This variable records if Side B is involved in a state-based armed conflict – according to the UCDP-PRIO dataset – in the same year.

COW location. The Correlates of War code for the location of the non-state conflict.

Region. The geographical region of the location. UCDP use the following regions:

- (1) **Europe.** Geographic definition, including Russia and the states in the Caucasus, corresponding to the COW numbers [200-395]
- (2) **Middle East.** Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Syria, Turkey, and the states of the Arabian peninsula, corresponding to the COW numbers [630-698]
- (3) **Asia.** Geographic definition, including Oceania, Australia, and New Zealand, and excluding the states in the Middle East, corresponding to the COW numbers [700-990]
- (4) **Africa.** Geographic definition, excluding Egypt, corresponding to the COW numbers [400-625]
- (5) **Americas.** Geographic definition, including the states in the Caribbean, corresponding to the COW numbers [2-165]

Sources

The data presented by UCDP is based on information taken from a selection of publicly available sources, printed as well as electronic. The sources include news agencies, journals, research reports, and documents of international and multinational organizations and NGOs. Global, regional and country-specific sources are used for all countries.

Both the independence and the transparency of the sources are deemed crucial. Each source is judged according to the context in which it is published, that is, according to the potential interests of the source in misrepresenting political or violent events. Since most sources are secondary sources, UCDP attempts to trace reports back to the primary source in order to decide whether they are reliable. In cases of biased sources or in situations where there is unreliable information, the events are normally only included in the high estimate.

Little information on the exact number of battle-related deaths is usually available, and media coverage varies considerably from country to country. It is important to emphasize that the fatality estimates given by UCDP is based on publicly accessible sources. Due to the lack of available information, it is possible that there are more fatalities than the UCDP high estimate, but it is very unlikely that there is fewer than the UCDP best estimate. One important caveat is that the precision of the numbers belies the uncertainty of the estimates; while UCDP estimates are based on aggregated totals of all incidents recorded in our sources, we acknowledge that numerous events is not reported in the media. Figures are revised retroactively each year as new information becomes available.

Methods

The data on non-state conflict is compiled by calendar year. The Factiva news database (previously known as the Reuters Business Briefing)³ is indispensable for the collection of general news reports. To extract the relevant articles from Factiva, the project uses an automated events data search using VRA® technology.⁴ This software makes it possible to retrieve all reports which contain information about individuals killed or injured. Each news report is then read by UCDP staff, and every event that contains information about individuals killed is coded manually into an events dataset. For every event, the following information is recorded: the date of the event, the reporting source, the primary source, the actors involved, where the event took place, what occurred, and an estimate of fatalities. Ideally, these individual figures are corroborated by two or more independent sources. These fatalities are later aggregated into a low, high, and best estimate for every calendar year.

The aggregated figures are also compared to total figures that appear in official documents, special reports and the news media. This serves as a contribution to clarify the context in which events occur, thus facilitating proper interpretation of the reporting in published sources.

Further information

More information about the non-state conflicts can be found at the UCDP website: <http://www.ucdp.uu.se> The UCDP intends to publish yearly summaries of the active non-state conflicts accordingly.

References

Bond, Doug; Joe Bond; Churl Oh; J. Craig Jenkins & Charles Lewis Taylor, 2003. “Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development”, *Journal of Peace Research* 40 (6): 733-745.

³ Factiva is a news and information service database which contains over 8,000 sources.

⁴ The VRA software system automatically generates specific events data and displays them in summary form. For more information on automated events data and coding, see Bond et al. (2003).