

UCDP Battle-Related Deaths Dataset Codebook:

Definitions, sources and methods for the UCDP Battle-related death estimates

Version 5.0
1 August 2011

Uppsala Conflict Data Program (UCDP)

Department of Peace and Conflict Research, Uppsala University

This codebook presents the variables in the UCDP Battle-related Deaths Dataset, and clarifies the definitions and methods, and provides explanations concerning the treatment of the sources consulted. The document focuses specifically on the coding of estimates for battle-related deaths, published since 2005 in the *Human Security Report* (Oxford: Oxford University Press). Funding for the UCDP Battle-related Deaths Dataset was provided by the Human Security Centre at the University of British Columbia.

When using the data, please refer to it with its correct name: the UCDP Battle-related Deaths Dataset with the correct version number, and provide the URL for the UCDP website.

(This version finalised by Ralph Sundberg)

The UCDP Armed Conflict Battle-Deaths Dataset

The definition of conflict inclusion in this dataset follows the UCDP definition of armed conflict. Variables A-I, M-S and U in this dataset are extracted directly from the UCDP-PRIO Armed Conflict Dataset: more information about the definitions and coding rules is available in the UCDP-PRIO Armed Conflict Dataset 1946-2010 Codebook (version 4-2011)¹

A.1 Conflict ID

The unique identifier of all conflicts, as given in the UCDP-PRIO Armed Conflict Dataset.

A.2 Dyad ID

The unique identifier of a dyad in a conflict, as given in the UCDP Dyadic Dataset.

B. Location_INC

The name of the country/countries whose government(s) has a primary claim to the object in dispute (the incompatibility), as given in the UCDP-PRIO Armed Conflict Dataset.

C. SideA

The name of the country/countries of Side A in a conflict. Always the government side in internal conflicts, as given in the UCDP-PRIO Armed Conflict Dataset.

D. SideA2nd

Side A Secondary lists all states that enter a conflict with troops to actively support side A in the dyad. By definition, only independent states can be a secondary party in conflict. A secondary party on side A shares the position in the incompatibility with Side A in the conflict. Side A Secondary does not need to meet the 25 battle-related deaths criterion to be included in the dataset; an active troop participation is enough.

Side A Secondary is a string variable, where the different names are separated by a comma (',').

E. SideB

The name of the country or opposition organizations of side B in a conflict, as given in the UCDP-PRIO Armed Conflict Dataset.

¹ For more information about the UCDP-PRIO Dataset, see Gleditsch et al. (2002)

F. SideBID

Side B ID is the unique identifier of the actor on side B in the conflict. For government actors, the Gleditsch & Ward (2007) country codes are used. For non-state actors, the ID is taken from the UCDP Actor Dataset (<http://www.pcr.uu.se/research/ucdp/datasets/>).

IDs for the state(s) on side A, as well as on side A secondary and side B secondary are found in GWNOA, GWNOA2nd and GWNOB2nd.

G. SideB2nd

Side B Secondary lists all states that enter a conflict with troops to actively support side B in the dyad. Only states are included as Side B Secondary. Furthermore, the states listed share the position in the incompatibility with Side B in the conflict. Side B Secondary does not need to meet the 25 battle-related deaths criterion to be included in the dataset; active troop participation is enough.

Side B Secondary is a string variable, where the different names are separated by a comma (',').

H. Incomp

A general coding of the conflict issue, if the conflict is an incompatibility over government or over territory, as given in the UCDP-PRIO Armed Conflict Dataset.

Incompatibility over government is coded as 1, and incompatibility over territory as 2.

I. Terr

The name of the territory over which the conflict is fought, provided that the incompatibility is over territory, as given in the UCDP-PRIO Armed Conflict Dataset.

J. Year

The year of observation.

K. bdBest

The UCDP Best estimate for battle-related deaths in the conflict/dyad in the given year.

L. bdLow

The UCDP Low estimate for battle-related deaths in the conflict/dyad in the given year.

M. bdHigh

The UCDP High estimate for battle-related deaths in the conflict/dyad in the given year.

N. Type

The UCDP-PRIO Armed Conflict Dataset identifies four different types of conflict:

extrasystemic (1), interstate (2), internal (3) and internationalized internal (4). See the UCDP-PRIO Armed Conflict Dataset codebook for specific definitions of the four types.

O. Battlelocation

The name of the country/countries in which battle-related deaths that can be related to the incompatibility have been recorded.

P. GWNoA

The Gleditsch and Ward country code of SideA (see UCDP-PRIO Armed Conflict Dataset codebook for list of countries and codes).

Q. GWNoA2nd

The Gleditsch and Ward country code of the state/states supporting SideA (see UCDP-PRIO Armed Conflict Dataset codebook for list of countries and codes).

R. GWNoB

The Gleditsch and Ward country code of SideB (see UCDP-PRIO Armed Conflict Dataset codebook for list of countries and codes).

S. GWNoB2nd

The Gleditsch and Ward country code of the state/states supporting SideB (see UCDP-PRIO Armed Conflict Dataset codebook for list of countries and codes).

T. GWNoLoc

The Gleditsch and Ward country code of the country in which the incompatibility is located (see UCDP-PRIO Armed Conflict Dataset codebook for list of countries and codes).

U. GWNoBattle

The Gleditsch and Ward country code of the country/countries in which battle-related deaths relating to the incompatibility have been recorded (see UCDP-PRIO Armed Conflict Dataset codebook for list of countries and codes).

V. Region

Identifies the region of the location. The following regions are given in the UCDP-PRIO Armed Conflict Dataset; Europe (1), Middle East (2), Asia (3), Africa (4), and the Americas (5).

W. Version

Identifies the version of the dataset.

Definition of Armed Conflict

The Uppsala Conflict Data Program (UCDP) defines an armed conflict as a contested incompatibility that concerns government and/or territory over which the use of armed force between the military forces of two parties, of which at least one is the government of a state, has resulted in at least 25 battle-related deaths each year.

The separate elements of this definition are defined as follows.

- (a) *Incompatibility that concerns government and/or territory.* The incompatibility must concern government and/or territory and it refers to the stated generally incompatible positions of the parties. An *incompatibility that concerns government* refers to the type of political system, the replacement of the central government or the change of the composition of the current government. An *incompatibility that concerns territory* refers to the status of a territory, for example, the change of the state in control of a certain territory (interstate conflict), secession or autonomy (intrastate conflict).
- (b) *Use of armed force.* This refers to the use of arms by the military forces of the parties in order to promote the parties' general position in the conflict, resulting in at least 25 deaths in a year. Arms are defined as any material means of combat, for example, manufactured weapons as well as sticks, stones, fire, water, and so on.
- (c) *Party.* This refers to the government of a state or an opposition organization or alliance of opposition organizations. *The government of a state* is that party which is generally regarded as being in central control, even by those organizations seeking to take over power. If this criterion is not applicable, the government is the party controlling the capital of the state. In most cases where there is a government, the two criteria coincide. An *opposition organization* is any non-governmental group which has announced a name for the group and its political goals and has used armed force to achieve them.
- (d) *State.* A state is an internationally recognized sovereign government controlling a specified territory or an internationally non-recognized government controlling a specific territory which is not disputed by an internationally recognized sovereign government, which previously controlled the same territory.

Definition of Battle-related Deaths

Battle-related deaths refer to those deaths caused by the warring parties that can be directly related to combat over the contested incompatibility². This includes traditional battlefield fighting, guerrilla activities (e.g. hit-and-run attacks/ambushes) and all kinds of bombardments of military bases, cities and villages etc. Urban warfare (bombs, explosions, and assassinations) does not resemble what happens on a battlefield, but such deaths are considered to be battle-related. The target for the attacks is either the military forces or representatives for the parties, though there is often substantial collateral damage in the form of civilians being killed in the crossfire, indiscriminate bombings, etc. All fatalities – military as well as civilian – incurred in such situations are counted as battle-related deaths.

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. This includes documents of the warring parties (governments and opposition organizations) when such sources are available, since they serve as a crucial complement when identifying statements about the parties' incompatible positions. 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 deaths in conflict 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,

² Battle-related deaths, which concern direct deaths, are not the same as *war-related deaths*, which includes both direct as well as indirect deaths due to disease and starvation, criminality, or attacks deliberately directed against civilians only (one-sided violence).

we acknowledge that numerous events are not reported in the media. Figures are revised retroactively each year as new information becomes available.

In addition to deciding the level of reliability of available sources, the project strives to identify the existence of censorship. Thus, other sources than regular news sources must be used to establish what is occurring in a country. Documents and reports issued by international and multinational organizations and NGOs are consulted for this purpose.

Methods

The data on armed conflicts are compiled by calendar year. They include data on conflict location, type of incompatibility, year the incompatibility was formed, year the warring party began its use of armed force, warring parties, secondary support, negotiations, and battle-related deaths during the year.

The data on battle-related deaths constitute the largest part of the data collection. 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.

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).

Best, Low and High Estimate of Battle-related Deaths

The general rule for UCDP's estimation of battle-related deaths is moderation. All battle-related deaths have to be verified in one way or another, and all estimates reported are based on UCDP expertise of each particular conflict. As a general rule, all figures are disaggregated as far as possible and any figures that are not trustworthy are disregarded in the coding process. Due to the great uncertainty of reports from conflict areas, the project provides three estimates concerning battle-related deaths for each year.

(a) *Best estimate.* The UCDP Best estimate consist of the aggregated most reliable numbers for all battle-related incidents during a year. If different reports provide different estimates, an examination is made as to what source is most reliable. If no such distinction can be made, UCDP as a rule include the lower figure given.

(b) *Low estimate.* The UCDP Low estimate consists of the aggregated low estimates for all battle-related incidents during a year. 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 UCDP High estimate consists of the aggregated high estimates for all battle-related incidents during a year. 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 are incidents when there is some uncertainty about what parties have been involved, these are also included in the high estimate.

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.

Gleditsch, Nils Petter; Peter Wallensteen; Mikael Eriksson; Margareta Sollenberg & Håvard Strand, 2002. "Armed Conflict 1946-2002, A New Dataset", *Journal of Peace Research* 39 (5): 615-637.