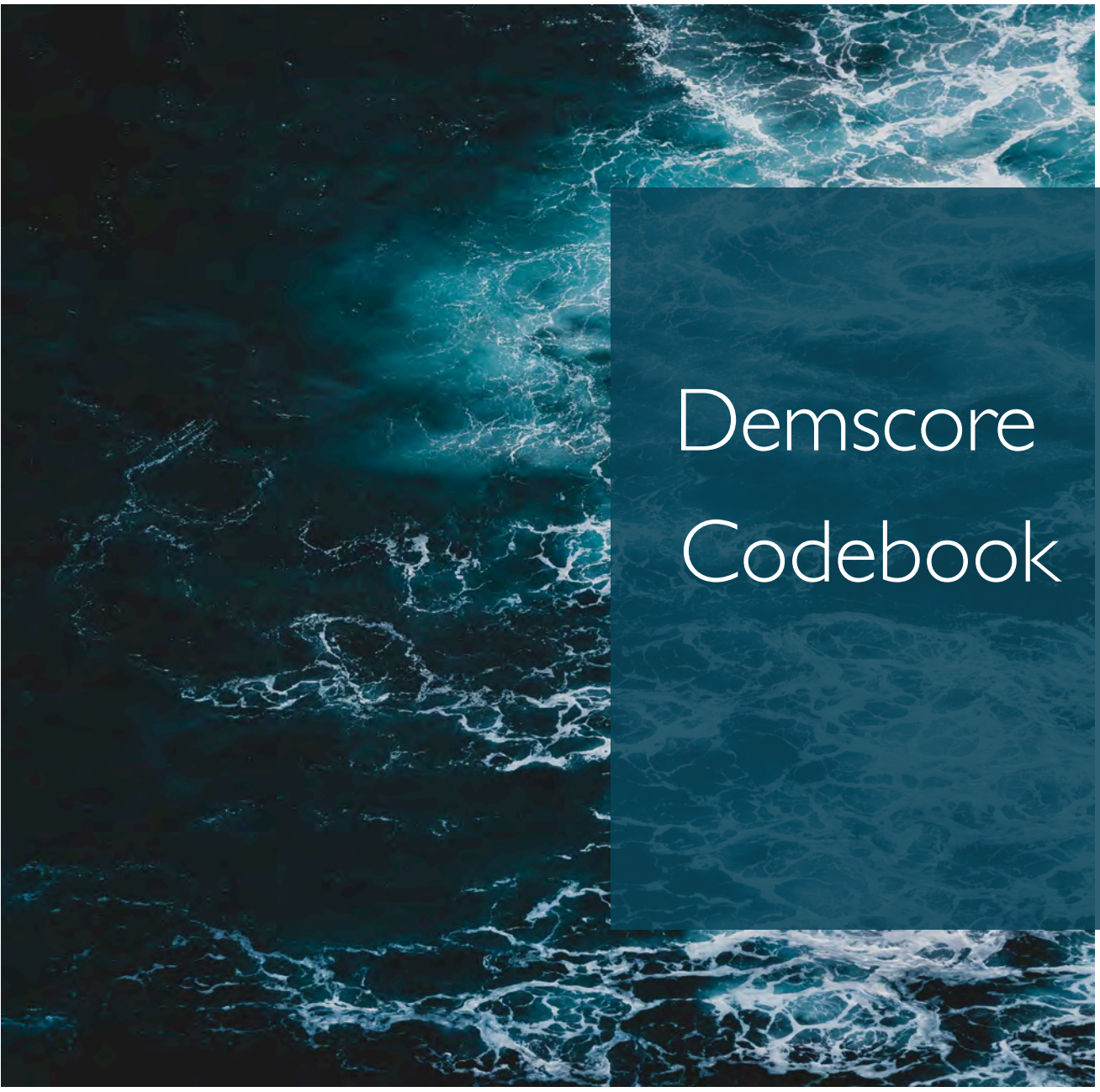
demscore



Demscore Codebook

October 2023

Copyright © Demscore
All rights reserved

Data version: 2.0
Document version: 2.0

Funders: Demscore is funded by the Swedish Research Council, University of Gothenburg, Stockholm University, Uppsala University, and Umeå University. For more information, please visit: `https://demscore.se`

For data enquiries: contact@demscore.se

Summary of Table of Contents

1 Explanatory Notes **5**
1.1 Release Notes . . . 5
1.2 The Demscore Codebook . . . 5
1.3 Methodology . . . 6
1.4 Citations . . . 6
1.5 Missing Data . . . 6
1.6 Download ID . . . 7
1.7 Unit Identifier Variables . . . 7

2 UCDP and VIEWS **8**
2.1 UCDP Conflict Termination Dataset, Conflict Level Version 3-2021 . . . 8
2.1.1 Identifiers . . . 8
2.1.2 Incompatibility . . . 9
2.1.3 Actors and Identifiers . . . 11
2.1.4 Timely Dimension . . . 12
2.1.5 Dataset Version . . . 14
2.1.6 Geographical Information . . . 14

3 Bibliography **16**

Table of Contents

1 Explanatory Notes **5**

1.1 Release Notes . . . 5
1.2 The Demscore Codebook . . . 5
1.3 Methodology . . . 6
1.4 Citations . . . 6
1.5 Missing Data . . . 6
1.6 Download ID . . . 7
1.7 Unit Identifier Variables . . . 7

2 UCDP and VIEWS **8**

2.1 UCDP Conflict Termination Dataset, Conflict Level Version 3-2021 . . . 8
2.1.1 Identifiers . . . 8
2.1.1.1 Conflict Episode ID (conflictep_id) . . . 9
2.1.1.2 Conflict Episode (conflictepisode) . . . 9
2.1.2 Incompatibility . . . 9
2.1.2.1 Type of Conflict 2 (type_of_conflict2) . . . 9
2.1.2.2 Conflict Termination (confterm) . . . 9
2.1.2.3 Outcome (outcome) . . . 9
2.1.2.4 Recurrence (recur) . . . 10
2.1.2.5 Incompatibility (incompatibility) . . . 10
2.1.2.6 Intensity Level (intensity_level) . . . 10
2.1.2.7 Type of Conflict (type_of_conflict) . . . 10
2.1.3 Actors and Identifiers . . . 11
2.1.3.1 Dyad Count (dyadcount) . . . 11
2.1.3.2 Conflict ID (conflict_id) . . . 11
2.1.3.3 Side A (side_a) . . . 11
2.1.3.4 Side A ID (side_a_id) . . . 11
2.1.3.5 Supporters of Side A (side_a_2nd) . . . 11
2.1.3.6 Side B (side_b) . . . 11
2.1.3.7 Side B ID (side_b_id) . . . 12
2.1.3.8 Supporters of Side B (side_b_2nd) . . . 12
2.1.4 Timely Dimension . . . 12
2.1.4.1 Episode End Date (ependdate) . . . 12
2.1.4.2 Episode End Date Precision (ependprec) . . . 12
2.1.4.3 Year (year) . . . 13
2.1.4.4 Start Date (start_date) . . . 13
2.1.4.5 Start Date Precision (start_prec) . . . 13
2.1.4.6 Start Date 2 (start_date2) . . . 14
2.1.4.7 Start Date 2 Precision (start_prec2) . . . 14
2.1.5 Dataset Version . . . 14
2.1.5.1 Dataset Version (version) . . . 14
2.1.6 Geographical Information . . . 14
2.1.6.1 Location (location) . . . 14
2.1.6.2 Territory Name (territory_name) . . . 15
2.1.6.3 Gleditsch and Ward Location (gwno_loc) . . . 15
2.1.6.4 Region (region) . . . 15

3 Bibliography **16**

1 Explanatory Notes

1.1 Release Notes

Demscore provides worldwide free access to harmonized data on Democracy, Environment, Migration, Social Policy, Conflict and Representation from several of the world's most prominent social science research institutes. The interdisciplinary nature of Demscore data facilitates large-scale comparative analyses. This is essential to advance adequate policy responses to complex societal challenges associated with the Sustainable Development Goals (SDGs) and beyond, facing Sweden, Europe, and the world today.

With a firm commitment to transparency and openness, Demscore enables users to gain comprehensive insights into various topics across the social sciences. The joint infrastructure ensures data integrity and quality at the highest international standards and maximizes usability in the measurement of contextual data with more than 23,000 variables across nearly all countries in the world, from 1750 to the present.

This creates critical time- and cost saving advantages in data collection, management, distribution, and not the least for end-users in the scientific community. This collaborative effort between leading Swedish universities pushes the scale of social science data to a new level and offers unprecedented possibilities for interdisciplinary research and knowledge advancement.
These are the key features of Demscore:

1. **Customized Download:** A fully normalized, joint PostgreSQL database, sophisticated programming, and a user-friendly web-based interface for users to generate custom-designed datasets and codebooks for download.
2. **Translations and Data Merges:** Demscore offers 800+ merges for dataset and Output Unit combinations.
3. **Metadata:** Demscore takes information on and organization of metadata to new heights with the inclusion of customized codebooks, a detailed methodology document, and a comprehensive handbook.
4. **Handling of Missing Data:** Demscore pioneers in developing an innovative approach to tackle missing data. Researchers can now account for missing values with increased precision, leading to more robust and reliable analyses.
5. **Merge Scores:** Demscore introduces a unique merge score mechanism. This powerful tool enables researchers to combine datasets effortlessly, uncovering connections and patterns that were previously hidden in isolated data silos.
6. **Thematic Datasets:** Demscore provides researchers with curated thematic datasets, each focused on a specific topic. These datasets bring together relevant variables from across the Demscore partners, facilitating in-depth investigations and comprehensive analyses of specific domains.
7. **Interactive Web Portal:** In addition to all the above, Demscore's web portal offers interactive visualization tools, user support and additional information on all partners and data sources.

For more information, please visit `https://www.demscore.se/` or contact contact@demscore.se.

1.2 The Demscore Codebook

The autogenerated Demscore Codebook lists variable entries for those variables chosen by the user along with citation guidelines and license per variable.

The meta data is extracted from the codebooks per dataset stored in a table in The Demscore PostgreSQL database with one row per variable for all datasets. This table includes codebook entries, variable tags, labels and other variable information in LaTeX format used for the generating of an automated codebook.
For all meta data across all datasets, Demscore uses one set of standard entries that all members

projects provided information on. In addition, variables from the different datasets have or require different sets of additional information as meta data specific to each dataset. These entries are also included, but as variable-specific meta data below the standard entries.

To avoid name clashes of variables and because variables entries have different naming schemes across datasets and projects, Demscore uses a hybrid form between original variable entries and new Demscore variable entries:

To begin with, all variable names in Demscore were cleaned in the beginning of the harmonization process. This cleaning included for instance that spaces or dots in the variable names were replaced with underscores and all letters were converted to lower case. The original tag is however still available and stored in the PostgreSQL table. Each variable in Demscore is available in a short- and long form. The short form is the cleaned name of the priginal variable tag, the long form begins with the name of the dataset from which it is retrieved, followed by the cleaned variable name.

For instance, the original name of the variable *MinisterPersonalID* from the H-DATA Foreign Minister Dataset is included as *ministerpersonalid* (short form) *hdata_fomin_ministerpersonalid* (long form) in Demscore.

In addition, each dataset includes Demscore unit-identifier variables which are named according to the following naming scheme: Beginning with u_, followed by the name of the primary unit and finally the variable tag. The *year*- variable from the COMPLAB SPIN The Out-of-Work Benefits Dataset (OUTWB), which is part of the primary unit u_complab_country_year has the Demscore unit identifier name *u_complab_country_year_year*.

1.3 Methodology

For details on our methodology please see the Demscore Methodology document which is included in the zip file retrieved from the download interface.

1.4 Citations

The Demscore project does not have a formal citation of its own. Hence, when using Demscore, we suggest that you cite the respective projects and datasets. We indicate how every dataset is to be cited in the autogenerated codebook you retreive with your data download, both in the dataset description and the codebook entry for each variable. Most often it is sufficient to cite the dataset a variable originates from, but sometimes there is a variable specific citation listed in the codebook entry in addition to that. For these cases, please also add the variable specific citation to the reference list of your publication. Full references are linked in the codebook entries of the variables and listed in the codebook's bibliography. We suggest you to also cite the Demscore Methodology Document when using data retrieved through Demscore.

1.5 Missing Data

Demscore indicates different types of missingness for observations in the customized datasets:
Missing in original data = Whenever an observation in the original variable is a missing (NA, missing code such as 7777, blank cell), we preserve this missing value. When the original source has special codes for various types of missing, those are preserved.
Missing code: -11111 = Demscore code for observation is missing due to the translation/merge, i.e., missing data due to no data being included for this combination of identifiers in the end Output Unit.
Missing code: -22222 = No observation is merged/translated, but the original data contains information for these identifier combinations elsewhere. For these cases, we use a different code. The user needs to consult the reference documents (Methodology Document Section 5.1. or the Demscore Handbook) to clarify why the translation to the identifier combinations in the end Output Unit was not possible.

1.6 Download ID

The download ID allows the user to share the ID with other users for replication purposes. A user can type the download ID into the Demscore website and retrieve the same download selection and files as the original user. This ID is autogenerated for each download from the Demscore website and will always retrieve the same data, even if the Demscore version was updated in the meantime.

Download ID:

1.7 Unit Identifier Variables

An Output Unit is defined as an output format in which variables can be retrieved from one or more datasets through a strictly defined output grid. A unit table defining this output grid contains unit identifier columns with u_ prefixes and the table is sorted based on these unit identifier columns and has a fixed number of rows. Unit columns are based on the columns that constitute a the unit of analysis in a dataset. They are added to the original dataset and marked by a unit prefix (consisting of a u_ and the dataset unit name) before the original variable name. Unit columns can contain slightly modified data, e.g., missing values are replaced by a default value. Sometimes we add additional columns to the unit table, for instance if a dataset includes both a country_id column with a numeric country code, we add the variable storing the full country name to the unit table as well for better readability.

2 UCDP and VIEWS

The Uppsala Conflict Data Program (UCDP) is the world's main provider of data on organized violence and the oldest ongoing data collection project for civil war, with a history of almost 40 years. Its definition of armed conflict has become the global standard of how conflicts are systematically defined and studied. UCDP produces high-quality data, which are systematically collected, have global coverage, are comparable across cases and countries, and have long time series which are updated annually. Furthermore, the program is a unique source of information for practitioners and policymakers. UCDP also operates and continuously updates its online database (UCDP Conflict Encyclopedia) on armed conflicts and organised violence, in which information on several aspects of armed conflict such as conflict dynamics and conflict resolution is available. This interactive database offers a web-based system for visualising, handling and downloading data, including ready-made datasets on organized violence and peacemaking, all free of charge. Data on armed conflicts have been published yearly in the Journal of Peace Research since 1993, in the Human Security Reports since 2005, in the SIPRI Yearbook since 1988, and in the report series States in Armed Conflict (1987-2012). In addition, UCDP researchers regularly publish research on organized violence, its causes, escalation, spread, prevention and resolution, in top scientific journals and books. More information is available on the project's website: `https://ucdp.uu.se/`

The Violence Impacts Early-Warning System (VIEWS) is a research consortium between Uppsala University and Peace Research Institute Oslo that brings together the political Violence Early-Warning System (ViEWS) and the interdisciplinary conflict impacts projects Societies at Risk and ANTICIPATE. At the core of the consortium lies a prediction model that systematically monitors hundreds of structural drivers and more complex conflict dynamics related to each country and sub-national grid cell (PRIO-GRID cell) in its geographic scope, and generates uniform predictions for impending conflict therein for each month in a rolling three-year window. New forecasts are released each month. The VIEWS data provided via the Demscore database is currently limited to forecasts for impending state-based conflict. They are provided as point predictions for the (log) number of fatalities in a given month and location. In the near future, they will be accompanied by corresponding predictions for armed conflict between non-state actors, as well as for violence against civilians. As the conflict impacts projects progress, the forecasting system will also be expanded with models that predict the impact of armed conflict on human development. For more information, please visit the consortium website: `https://viewsforecasting.org/`

2.1 UCDP Conflict Termination Dataset, Conflict Level Version 3-2021

Dataset tag: ucdp_term_conflict

Output Unit: UCDP Conflict-Year, i.e., data is collected per conflict and year.

Description: This dataset provides information on specific start- and end- dates for conflict activity and means of termination for each conflict episode. The data is available as a conflict-level dataset which corresponds with the UCDP/PRIO Armed Conflict Dataset v 21.1, and a dyad-level dataset which corresponds with the UCDP Dyadic Dataset v. 21.1.

Dataset citation:
Kreutz, Joakim (2010) How and When Armed Conflicts End: Introducing the UCDP Conflict Termination Dataset. *Journal of Peace Research*, 47(2).

License: UCDP offers a web-based system for visualising, handling and downloading data, including ready-made datasets on organized violence and peacemaking. All UCDP data are free of charge.

More detailed information on the dataset can be found at the following web page: `https://ucdp.uu.se/downloads/index.html`

2.1.1 Identifiers

Variables in this section identify rows in the dataset.

2.1.1.1 Conflict Episode ID (conflictep_id)

Long tag: ucdp_term_conflict_conflictep_id

Original tag: conflictep_id

Dataset citation: Kreutz (2010)

Description:

The unique identifier for each conflict episode. It is constructed by the Conflict ID*100 + 1, 2, 3, etc...

2.1.1.2 Conflict Episode (conflictepisode)

Long tag: ucdp_term_conflict_conflictepisode

Original tag: conflictepisode

Dataset citation: Kreutz (2010)

Description:

The unique identifier for each conflict episode. It is constructed by the Conflict ID*100 + 1, 2, 3, etc...

2.1.2 Incompatibility

These variables indicate the cause(s) of the conflict, i.e. the stated (in writing or verbally) generally incompatible positions.

2.1.2.1 Type of Conflict 2 (type_of_conflict2)

Long tag: ucdp_term_conflict_type_of_conflict2

Original tag: type_of_conflict2

Dataset citation: Kreutz (2010)

Description:

The same conflict episode, or dyadic conflict episode, may include both years where neither side receive secondary support and years when they do. Type 2 thus combine the categories of internal armed conflict and internationalized armed conflict described above.

1. Extrasystemic armed conflict.
2. Interstate armed conflict.
3. Intrastate armed conflict.

2.1.2.2 Conflict Termination (confterm)

Long tag: ucdp_term_conflict_confterm

Original tag: confterm

Dataset citation: Kreutz (2010)

Description:

Confterm is a dummy variable that codes whether the conflict is inactive the following year and an episode of the conflict thus ends. If the conflict is inactive the following year(s), this variable is coded as 1. If not, a 0 is coded.

2.1.2.3 Outcome (outcome)

Long tag: ucdp_term_conflict_outcome

Original tag: outcome

Dataset citation: Kreutz (2010)

Description:

The coding of outcomes are based on the final year of activity and first year of non-activity. While the dataset include some information (i.e. ceasefires and peace agreements) outside this window, it does not follow warring party devel- opment beyond this time period.

1= Peace agreement

2= Ceasefire
3= Victory for Side A /Government Side 4= Victory for Side B /Rebel Side
5= Low activity (less than 25 battle-deaths) 6= Actor ceases to exist

2.1.2.4 Recurrence (recur)

Long tag: ucdp_term_conflict_recur

Original tag: recur

Dataset citation: Kreutz (2010)

Description:

A dichotomous measure that this observation is a recurrence of a conflict or dyad which have experienced an spell of non-conflict.

2.1.2.5 Incompatibility (incompatibility)

Long tag: ucdp_term_conflict_incompatibility

Original tag: incompatibility

Dataset citation: Kreutz (2010)

Description:

The incompatibility for the conflict, taken from the UCDP/PRIO Armed Con- flict Dataset. The stated incompatibility is what the parties claim to be fighting over.
1= Territory
2= Government
3= Government and Territory

2.1.2.6 Intensity Level (intensity_level)

Long tag: ucdp_term_conflict_intensity_level

Original tag: intensity_level

Dataset citation: Kreutz (2010)

Description:

The intensity variable is coded in two categories:
1. Minor: between 25 and 999 battle-related deaths in a given year. 2. War: at least 1,000 battle-related deaths in a given year.

2.1.2.7 Type of Conflict (type_of_conflict)

Long tag: ucdp_term_conflict_type_of_conflict

Original tag: type_of_conflict

Dataset citation: Kreutz (2010)

Description:

UCDP define four types of conflict:

1. Extrasystemic armed conflict occurs between a state and a non-state group outside its own territory. (In the COW project, extrasystemic war is subdivided into colonial war and imperial war, but this distinction is not used here.) These conflicts are by definition territorial, since the government side is fighting to retain control of a territory outside the state system.
2. Interstate armed conflict occurs between two or more states.
3.Internal armed conflict occurs between the government of a state and one or more internal opposition group(s) without intervention from other states.
4. Internationalized internal armed conflict occurs between the government of a state and one or more internal opposition group(s) with intervention from other states (secondary parties) on one or both sides.

2.1.3 Actors and Identifiers

These variables identify the conflicting parties using the UCDP ID system for conflicts, actors and dyads.

2.1.3.1 Dyad Count (dyadcount)

Long tag: ucdp_term_conflict_dyadcount

Original tag: dyadcount

Dataset citation: Kreutz (2010)

Description:

This variable provide information about how many different dyads are active in the conflict this year.

2.1.3.2 Conflict ID (conflict_id)

Long tag: ucdp_term_conflict_conflict_id

Original tag: conflict_id

Dataset citation: Kreutz (2010)

Description:

The unique conflict ID, taken from the UCDP/PRIO Armed Conflict Dataset.

2.1.3.3 Side A (side_a)

Long tag: ucdp_term_conflict_side_a

Original tag: side_a

Dataset citation: Kreutz (2010)

Description:

The first primary party to the conflict, taken from the UCDP/PRIO Armed Conflict Dataset. Side A is by definition always a primary party to the conflict. In internal conflicts, side A is always the government side, it is one of the sides in interstate conflicts and the colonial state in extrastate conflicts.

2.1.3.4 Side A ID (side_a_id)

Long tag: ucdp_term_conflict_side_a_id

Original tag: side_a_id

Dataset citation: Kreutz (2010)

Description:

ID for Side A

2.1.3.5 Supporters of Side A (side_a_2nd)

Long tag: ucdp_term_conflict_side_a_2nd

Original tag: side_a_2nd

Dataset citation: Kreutz (2010)

Description:

The state which contribute with troops to actively support Side A in the conflict.

2.1.3.6 Side B (side_b)

Long tag: ucdp_term_conflict_side_b

Original tag: side_b

Dataset citation: Kreutz (2010)

Description:

The second primary party to the conflict during the conflict episode, taken from the UCDP/PRIO Armed Conflict Dataset. Like Side A, Side B is by definition a primary party

to the conflict. Side B is the opposition side of all internal and extrastate conflicts and the second side in an interstate conflict. Thus, side B can include both states and non-governmental opposition groups, depending on the type of conflict.
When the primary party listed on Side B is an opposition group, the column lists the group name in abbreviated form. Even if the group changes its name during the course of the conflict we record them under the same name for all years. See the UCDP Actor Dataset for the full name and name history of opposition groups.

2.1.3.7 Side B ID (side_b_id)

Long tag: ucdp_term_conflict_side_b_id

Original tag: side_b_id

Dataset citation: Kreutz (2010)

Description:

Side B ID is the unique identifier of the actor on side B in the conflict. For government actors, the Gleditsch and Ward (2007) country codes are used. For non-state actors, the ID is taken from the UCDP Actor Dataset (UCDP 2015a).

2.1.3.8 Supporters of Side B (side_b_2nd)

Long tag: ucdp_term_conflict_side_b_2nd

Original tag: side_b_2nd

Dataset citation: Kreutz (2010)

Description:

The state which contribute with troops to actively support Side B in the conflict.

2.1.4 Timely Dimension

These variables provide information on when the conflict takes place.

2.1.4.1 Episode End Date (ependdate)

Long tag: ucdp_term_conflict_ependdate

Original tag: ependdate

Dataset citation: Kreutz (2010)

Description:

The date, as precise as possible, when the conflict violence stopped. If detailed information is lacking the Conflict Termination Dataset sets the date to 31 December.

2.1.4.2 Episode End Date Precision (ependprec)

Long tag: ucdp_term_conflict_ependprec

Original tag: ependprec

Dataset citation: Kreutz (2010)

Description:

The enddate is coded as precisely as possible. For certain conflicts we can pinpoint the termination of the armed conflict down to a single event, taking place on a specific day. For other conflicts, this is not possible, due to lack of precise information. The Endprec (end precision) is coded to highlight the level of certainty for the date set in the Enddate variable.

1= Day, month and year are precisely coded; we have good information on the event.
2= Day is assigned; month and year are precisely coded. The assigned date can either be one of several events that can be classified as the last; it can be the last day in a period when several fatalities have been reported jointly or it can be an event that different sources claim occurred on different dates.
3= Day is unknown; month and year are precisely coded. The day is known to be in a given month, but we are missing information on an exact date. Day is then set to the last day of

the month.
4= Month is assigned; year is coded precisely.
5= Day and month are unknown, year is coded precisely.
6= Year is assigned. There is a wide disagreement between different sources, so that not even year can be coded precisely. The end year is assigned based on subjective judgment.
7= Year is missing. No information on the end date is available; Enddate is set to 31 December of the last year recorded in the conflict.

2.1.4.3 Year (year)

Long tag: ucdp_term_conflict_year

Original tag: year

Dataset citation: Kreutz (2010)

Description:

The year of the observation.

2.1.4.4 Start Date (start_date)

Long tag: ucdp_term_conflict_start_date

Original tag: start_date

Dataset citation: Kreutz (2010)

Description:

The date of the first battle-related death recorded in the conflict is coded as the Startdate in the dataset. The date is set after the conflict fulfils all criteria required in the definition of an armed conflict, except for the number of deaths. In some cases, the initial fatality occurs in a year prior to the first year of activity. For instance, in the conflict in Ethiopia over the territory Eritrea, the first battle-related deaths occurred in September 1961. During the remaining months of 1961, the conflict did not reach the required total of 25 battle-related deaths and the conflict is thus coded as inactive in 1961. 25 battle-related deaths in a year were not recorded until three years later.

2.1.4.5 Start Date Precision (start_prec)

Long tag: ucdp_term_conflict_start_prec

Original tag: start_prec

Dataset citation: Kreutz (2010)

Description:

The Startdate is coded as precisely as possible. For certain conflicts we can pinpoint the start of the armed conflict down to a single event, taking place on a specific day. For other conflicts, this is not possible, due to lack of precise in- formation. Startprec (start precision) is coded to highlight the level of certainty for the date set in the Startdate variable.
1= Day, month and year are precisely coded; we have good information on the event.
2= Day is assigned; month and year are precisely coded. The assigned date can either be one of several events that can be classified as the first; it can be the last day in a period when several fatalities have been reported jointly or it can be an event that different sources claim occurred on different dates.
3= Day is unknown; month and year are precisely coded. The day is known to be in a given month, but we are missing information on an exact date. Day is then set to the first day of the month.
4= Month is assigned; year is coded precisely. Day is set as the first day of the assigned month.
5= Day and month are unknown, year is coded precisely. Day and month are set as the 1 January of the coded year.
6= Year is assigned. There is a wide disagreement between different sources, so that not even year can be coded precisely. The start year is assigned based on subjective judgment.
7= Year is missing. No information on the start date is available; Startdate is set to 1 January of the first year recorded in the conflict.

2.1.4.6 Start Date 2 (start_date2)

Long tag: ucdp_term_conflict_start_date2

Original tag: start_date2

Dataset citation: Kreutz (2010)

Description:

Startdate2 provides information about the date when a conflict episode reach 25 battle-related deaths in a calendar year, thus indicating the date that all criteria required in the definition of armed conflict are fulfilled.

2.1.4.7 Start Date 2 Precision (start_prec2)

Long tag: ucdp_term_conflict_start_prec2

Original tag: start_prec2

Dataset citation: Kreutz (2010)

Description:

The level of certainty for the coding of StartDate2.

1= Day, month and year are precisely coded; we have good information on the event.
2= Day is assigned; month and year are precisely coded. The assigned date can either be one of several events that can be classified as the first; it can be the last day in a period when several fatalities have been reported jointly or it can be an event that different sources claim occurred on different dates.
3= Day is unknown; month and year are precisely coded. The day is known to be in a given month, but we are missing information on an exact date. Day is then set to the first day of the month.
4= Month is assigned; year is coded precisely. Day is set as the first day of the assigned month.
5= Day and month are unknown, year is coded precisely. Day and month are set as the 1 January of the coded year.
6= Year is assigned. There is a wide disagreement between different sources, so that not even year can be coded precisely. The start year is assigned based on subjective judgment.
7= Year is missing. No information on the start date is available; Startdate is set to 1 January of the first year recorded in the conflict.

2.1.5 Dataset Version

The version of the dataset.

2.1.5.1 Dataset Version (version)

Long tag: ucdp_term_conflict_version

Original tag: version

Dataset citation: Kreutz (2010)

Description:

The version of the dataset. Note that this most recent version (v2-2015) include a different coding scheme for outcomes compared with earlier versions.

2.1.6 Geographical Information

These variables provide information on where the conflict takes place.

2.1.6.1 Location (location)

Long tag: ucdp_term_conflict_location

Original tag: location

Dataset citation: Kreutz (2010)

Description:

The location of the conflict, taken from the UCDP/PRIO Armed Conflict Dataset. Location is defined as the government side of a conflict, and should not be interpreted as the geographical location of the conflict.

For internal and internationalized internal conflicts only one country name is listed. This is the country whose government or territory is disputed. For certain conflicts, such as Kurdistan, the disputed territory is divided between different countries. Following our definition, we have coded separate conflicts for each country.

For interstate conflict, both primary parties are listed in the Location field. Even if several governments are involved in the conflict, only countries that fulfill the inclusion criteria for primary actors are listed here. This normally means that two countries are listed, but there are three notable exceptions: In the Arab-Israeli war of 1948–49 as well as the Suez war of 1956 and the war in Iraq in 2003, there are more than two primary parties to the conflict.

For extrastate conflicts, Location is set to be the disputed area, not the govern- ment of the colonial power. Location is a string variable, listing the names of the countries involved. These might be fighting together or against each other. The string is split in two ways, hyphen ('-') splits the different sides in an in- terstate war, and comma (',') splits different countries fighting together on the same side.

2.1.6.2 Territory Name (territory_name)

Long tag: ucdp_term_conflict_territory_name

Original tag: territory_name

Dataset citation: Kreutz (2010)

Description:

The specified contested territory for conflicts over territory, taken from the UCDP/PRIO Armed Conflict Dataset. In case the two sides use different names for the disputed territory, the name listed is the one used by the opposition organization.

2.1.6.3 Gleditsch and Ward Location (gwno_loc)

Long tag: ucdp_term_conflict_gwno_loc

Original tag: gwno_loc

Dataset citation: Kreutz (2010)

Description:

This field contains the country code(s) for the state(s) listed in the Location variable. Thus, it lists the country codes for the primary party/parties in the conflict. The country codes are taken from Gleditsch and Ward (2007).

2.1.6.4 Region (region)

Long tag: ucdp_term_conflict_region

Original tag: region

Dataset citation: Kreutz (2010)

Description:

The geographic region of the conflict, taken from the UCDP/Prio Armed Con- flict Dataset. This variable groups the various conflicts into five geographical categories, dependent on the location of the conflict.

1= Europe
2= Middle East
3= Asia
4= Africa
5= Americas

3 Bibliography

Kreutz, J. (2010), 'How and when armed conflicts end: Introducing the ucdp conflict termination dataset', *Journal of Peace Research* **47**(2), 243–250.