

Conflict-related Violence in Postwar Cities

CODEBOOK

Version 2024-10-31

Contents

Conflict-related Violence in Postwar Cities.....	1
Introduction	1
Definitions	2
Postwar cities	2
Conflict-related violence	4
Event and location	5
Coding rules and criteria	6
Selection of cities	6
Search procedure	7
Additional sources.....	8
Types of violence	8
Geocoding	9
Variables in the dataset	10
Limitations	13
Reference list	14

Introduction

This codebook sets out the operational procedures guiding the coding of events of conflict-related violence in postwar cities. The data collection is undertaken within the scope of the project *The continuation of conflict-related violence in postwar cities: Mapping violence at the street level* (Swedish Research Council 2019-02563). The aim is to generate a novel systematic dataset that will be relevant to a broad set of researchers and policymakers as it can be used to theorise preventive measures, explore intra-urban conflict dynamics, and assess the impact of conflict-related violence on the constitution and functioning of postwar cities. Events in the dataset will be coded according to maximum temporal and geographical precision (i.e. the exact date and if possible what time of day, and at what street or building) and will contain additional information on the actors involved, type of violence, and number of casualties. The dataset relies on similar coding procedure as the Uppsala Conflict Data Program (UCDP) and will be fully compatible with existing UCDP datasets.

Definitions

Postwar cities

In order to define the postwar city, we need to define its two components first. We subsequently define *postwar* as the period between war and the establishment of a mutually acceptable peace during which the socio-political ordering of society is fundamentally contested (Gusic, 2019; see also Richmond, 2016; Wallensteen, 2015). We understand war as the systematic armed force between groups with political and/or territorial ambition, resulting in large-scale destruction and loss of lives, while peace is the establishment of an unchallenged socio-political ordering of society. Postwar is consequently the period between these two endpoints. We acknowledge that the termination of war is not a dichotomous variable or singular event, but rather a long process that is not necessarily unidirectional. Operationally, we consider the postwar period to start when the large-scale loss of lives between the state army and one or more rebel groups ends, either through the signing of a peace agreement; through the military victory of one side over the other; or through a lasting halt to the fighting (if the conflict results in less than 25 battle-related deaths for at least five years, with or without a formal ceasefire agreement).¹

The city is in turn defined as a political entity that is heterogeneous, densely populated, and open/permeable; occupies a central position within its wider socio-political context; and functions through mixing, conflict, accommodation, creativity, and fragmentation (Brenner, 2014; Gusic, 2019; Hall, 1999; Lefebvre, 1996; Mumford, 1938; Sassen, 2013). The *postwar city* is subsequently – in strictly conceptual terms – a city which has experienced war, no longer does, but where the socio-political ordering of postwar society remains contested and where conflicts over peace(s) are ongoing (Gusic, 2019). To transform this conceptual definition into an operational one, we use the following steps.

We begin by identifying *wars* which have ended at any time between 1989 to 2020, either through the signing of a negotiated settlement, through the military victory of one side over the other, or through a de facto end to battle-related deaths (with or without a formal ceasefire agreement).² As mentioned above, for the latter, we consider the war to be terminated if battle-related deaths cease for a period of at least five years. To identify our population of cases, we rely on the Uppsala Conflict Data Program's (UCDP) definitions and the UCDP/PRIO Armed Conflict Dataset (Gleditsch, Wallensteen, Eriksson, Sollenberg, & Strand, 2002; Pettersson & Öberg, 2020).³ We focus on intrastate

¹ This operationalization is in line with for instance (Bara, 2020). We wish to distinguish between temporary lulls in the fighting, and situations where the war really did end for a substantial period. While the threshold may be seen as arbitrary, we contend that it is reasonable to consider a war to be terminated if there is a prolonged period with no or only a few battle-related fatalities.

² To exemplify: the war in Angola is considered terminated through the Memorandum of Understanding (reaffirming the Lusaka Protocol) in 2002; the war in Sri Lanka was ended through a military victory of the government over the LTTE in 2009; and the war over the territory Nagorno-Karabach was de facto terminated in 1994 as the large-scale fighting ceased but the conflict remained unresolved. The latter type of terminated wars is sometimes referred to as frozen conflicts (Smetana & Ludvík, 2019). For our purposes the end of large-scale fighting represents the end of war and start of a postwar period. Yet – given our theoretical understanding of postwar – this period does not mean that society is at peace.

³ A state-based armed conflict is “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 in one calendar year” (UCDP, 2020). We focus on wars and major armed conflicts, i.e., conflicts that cause at least 1000 battle-related deaths in one year and/or have reached a cumulative intensity of at least 1000 battle-related deaths (Pettersson, 2019). We include cases where this threshold was reached only according to UCDP's high estimate (6 cases) – this given that UCDP best estimates are very conservative and only includes clear cases of direct fighting between the two sides in the conflict; in

wars rather than wars between states, as the former are more likely to result in continued intrastate contestation because the parties involved must coexist within the same political system. However, we include wars with external involvement (i.e. internationalized wars). The starting focus on wars rather than cities is motivated by our desire to exclude cases situated within a context of ongoing war but where the fighting no longer affects the city – such as Damascus in the still-ongoing Syrian war, or Mostar between 1994 and 1995 (i.e. after the local ceasefire but before the end of war in its entirety across the state).⁴

Having established the population of terminated wars, we proceed to identify *cities* affected by these wars. As no readily available systematic definition of cities is readily available,⁵ we make the following operational decisions, based on our theoretical definition of cities above.⁶ First, we exclude all cities below a certain population size (50 000 inhabitants).⁷ This is because we do not expect cities below that size to exhibit the type of density and heterogeneity we associate with cities. Second, in order to focus on cities that play a central role in their wider region,⁸ we identify the most populous cities, the most economically important ones,⁹ and the politically most important ones,¹⁰ for each postwar context. More specifically, we include the five most populous cities for each state; the five most economically important cities for each state (as reflected in local economic production); national capitals (for conflicts over government) and regional capitals (in territorial conflicts); and the five cities most affected by violence during the war (as reflected by fatal conflict events). This results in a

many cases, the high estimate may better reflect the true level of destructiveness. Such cases include the conflict over Bihaćka Krajina (Bosnia-Herzegovina) which caused 971 fatalities according to the UCDP best estimate (REF GED), but 1634 according to the high estimate, and the conflict over government in Guinea which caused 649 fatalities according to the UCDP best estimate but 1300 according to the high estimate.

⁴ We do include cases where the war formally ended *across the state* but then restarted at a later point. Concrete examples include Baghdad between 1996 and early 2001, and Gaza between 1996 and 2000.

⁵ For some context, each state has its own definition of what constitutes a city or an urban locality. Each state also has different time intervals, methods, as well as levels of precision when measuring these idiosyncratically defined cities. This means e.g. the UNHABITAT list of cities – which is based on the information of states – includes multiple city definitions as well as presents cities that are not in temporal sync, have been measured differently, and whose underpinning numbers are not equally reliable (see e.g. Satterthwaite, 2010).

⁶ See further Elfversson, Gusic & Murtagh, 2023.

⁷ We argue that it is unlikely that any socio-political entity will either have the constitutive aspects of the city or function through urban dynamics if its population is below this threshold. We rely on population figures from a number of sources in order to triangulate the information: we begin by consulting the UN statistics division (UNSD, 2019) and the Africapolis database (OECD/SWAC, 2023), and triangulate using other databases such as the World Cities Dataset (SimpleMaps, 2018) which is based on US governmental sources and the City Population database (<http://www.citypopulation.de>), original census reports, and other country-specific reports.

⁸ We also expect that cities which are not central to their particular postwar societies are not likely to be contested to the point that they affect the wider postwar society in any substantial way, nor will they concentrate conflict-related violence in any significant numbers due to not having enough symbolic or political importance, enough resources, or enough anonymity in the aftermath of war to attract and/or retain those engaging in conflict-related violence. To test expectations about where and how such violence emerges and concentrates, a focus on the most central cities is thus warranted.

⁹ No systematic data source for city GDP is readily available. Instead, we proceeded as follows: We rely on the PRIO GRID (Tollefsen, Strand, & Buhaug, 2012) which contains data on local economic productivity (Nordhaus, 2006). We matched the grid to the location of cities (using the Open Street Map plugin in QGIS) and, for each country, identified the cities contained in the cells with the highest economic output (using 2005 values, i.e. the middle of the time period). We acknowledge that this procedure has important limitations – for instance, it does not capture the informal economy – but given that population is another measure we expect that the economically most central cities will not be excluded even if we are unable to measure informal economies.

¹⁰ I.e. national capitals (for conflicts over government) and regional ones (in territorial conflicts). These often, but not always, overlap with the most populous cities.

potential of up to 16 cities per state/conflict, but usually there is a lot of overlap between the different categories.

In the next step, we limit ourselves further by excluding cities which were never directly affected by the war. Here, we retain only those cities that they were either located in a war zone and/or experienced battle-related deaths during the war in which they were embroiled.¹¹ While some wars embroil the entire country, others are more localized; in particular, conflicts over territory often concentrate in the contested region of a state. We expect postwar dynamics to be different in a city which was affected by the war directly – such as Kissidougou in the civil war in Guinea – than in a city outside of the conflict zone, such as Nzérékoré in the same country. The same goes for Pristina (in Kosovo) and Belgrade (outside of Kosovo) during the Kosovo War as well as for Beirut and Tripoli in the civil war in Lebanon. For countries that have experienced multiple armed conflicts, we focus on cities within the area each specific war was fought over; for instance, in Ethiopia we include cities within the areas affected by the wars over government and over the status of Oromiya, Ogaden and Eritrea, but not those cities that were only affected by the less intense armed conflicts (i.e. armed conflicts that do not qualify as wars) over other territories.

Summing up, our list of postwar cities includes cases that:

- Are located within a postwar context
- Have at least 50,000 inhabitants
- Have been directly affected by the war
- Are the most populous **and/or** economically **and/or** politically important

Conflict-related violence

Within this data collection, we define *violence* as the use of direct physical force exerted by one actor against another (or by two or more actors against each other). This includes beatings, shootings, torture, rape, and landmines, but not discrimination, threats of violence, or denial of healthcare. Violence inflicted upon oneself, such as suicide, is not included. We also exclude violence that is clearly unintentional: e.g., when people get injured by falling bullets after people fire in the air to celebrate something. We include only *serious violent events*: fatal violence, and events of serious non-fatal violence. For the latter category, we include events of violence levelled with the presumed intent of causing major bodily harm to the other actor, and which result in physical injury¹² (including incidents where injuries are not explicitly reported, but where circumstances imply that injuries were sustained). In effect, this means that we include e.g. shootings (even if they only result in minor injuries), but we do not include injuries (even if major) if sustained from violence which appear to not have been levelled with the intent to cause major bodily harm (e.g. teargassing).

Conflict-related violence is in turn defined as violence directly connected to the preceding war because it is perpetrated by those previously active in the war or is manifested across conflict lines – e.g. residual fighting between former warring parties, revenge killings or instrumentally employed sexual violence, violence perpetrated by unintegrated former fighters, and violence along identity lines polarized during war (Suhrke & Berdal, 2013). Conflict-related violence, by our definition, includes violence against peacekeepers if it meets the operational criteria below (i.e. it is perpetrated by the

¹¹ For this purpose, we rely on UCDP events data (Sundberg & Melander, 2013) as well as conflict polygons (Bara, 2020).

¹² This is in line with the operationalization used by the Peacemakers at Risk project (PAR) which also collects systematic data on both fatal and non-fatal violence. See (Bromley, 2018)

former warring parties or affiliated actors, and/or explicitly related to the conflict). We also include violence committed by peacekeepers during their deployment.

The explicit focus on violence directly related to the war means that other forms of violence such as social violence (e.g. rape or sexual harassment, assault, and domestic violence) or economic violence (bank robberies, muggings) are not considered conflict-related – unless perpetrated by unintegrated former fighters or explicitly politically motivated. This also means that violence that is *indirectly* conflict-related – e.g. violence spurred by economic hardships caused by war or enabled by the lack of rule of law in its aftermath – will not be included in the dataset.

To systematically identify events of conflict-related violence, even with the narrower scope adopted here, is a difficult task. Automated searches always run the risk of bias where e.g. formal ties to warring parties may be openly acknowledged in some case but not in others while the framing of news reports risks identifying violence as conflict-related in some cases but not others. Within this project, therefore, research assistants make an assessment based on qualitative information. More specifically, events are coded as conflict-related based on the following information:

Formal warring parties – events are perpetrated by at least one of the former warring parties or groups explicitly and clearly connected to them (e.g. pro-government militia). Special rules apply to the government

Identity lines – events of violence across the same identity lines that were central in the war (e.g. Serbs vs Croats in postwar Bosnia-Herzegovina) – unless contextual information clearly indicates the event is *not* conflict-related, e.g. the actors were unaware of the other's identity or the event is connected to dispute clearly unrelated to their identities.

Statements – the primary actors involved make explicit statements connecting the violence to the preceding war.

For each event, research assistants also code the degree of certainty that the event is indeed conflict-related (see further under “Variables in the dataset”).

Event and location

We define an event as an individual incident of physical violence, exerted by one actor against another (or by two or more actors against each other), and which occurs at a specific place and at a specific time.¹³ Both the temporal and spatial dimension will be coded with maximum precision, in order to enable analysis of microdynamics of conflict-related violence within postwar cities.

In line with the above, the data covers events of conflict-related violence that occur in postwar cities. A further specification required concerns the geographic extent of cities – i.e., where we draw the line for the city's spatial domain: for instance, should events taking place in suburbs or satellite cities be included? Here, we note our theoretical definition of cities - defined as being heterogeneous, densely

¹³ In some cases, information available is of an aggregated nature – e.g., the fighting started at the town square on Tuesday evening and then spread to encompass the city over the next two days. In such instances, the events are disaggregated as much as possible. For instance, a triggering event may be possible to identify and code with high precision, and reports may identify precise death tolls for some neighborhoods. In such cases the summary event is broken down into all clearly identifiable components whereas the remaining fatalities are coded as a separate summary event.

populated, open/permeable, and functioning through mixing, conflict, accommodation, creativity, and fragmentation.

In order to translate this into an operational guideline, we expect these characteristics to be present within the contiguous built-up area with dense communication networks to the rest of the city – this is where people and ideas are in close proximity and contact. This means that suburbs that are well-integrated into the city fabric are included, but not satellite cities and peri-urban areas. In terms of terminology used in urban studies (e.g. population statistics), this operationalization lies between the “city proper” (often a rather small core overlapping with the “old city”) and the “metropolitan area” (which often covers multiple well-connected cities within the same administrative area) (see further [UNHABITAT 2020](#)). Instead, it aligns with what is often called the “urban area” – the city including suburbs, stopping where the dense built-up area and infrastructure stops.

Coding rules and criteria

In line with the definitions above, research assistants will – for each city and time-period included in the coding – rely on a pre-set search string in order to identify relevant events in online news databases covering a vast number of local and international news sources. They will identify events of conflict-related violence within each city, and code their exact location assisted by software such as Google Earth and online gazetteers. For each event, the dataset will list the date, type of conflict-related violence (e.g. assassination, clash, riot, rape; including non-fatal violence) and number of casualties, and the gender and group identity of perpetrators and victims if available. For each event, geo-coordinates will be coded at maximum precision (i.e. at the street level).

Selection of cities

From the total population of postwar cities (see operationalization above), we selected a representative sample of 21 cities for which data on conflict-related violence is collected. We select postwar cities from all regions (Africa, Americas, Asia, Middle East, Europe) weighted by frequency of cases per region, and bounded such that 3-6 per region (resulting in Asia and the Americas being over-represented by one case, and Africa under-represented by two).

Table 1. Case selection criteria

Region	Conflicts	Countries	Cities	Ave. share	Cases included
Africa	20 (31%)	18 (38%)	73 (42%)	37%	6
Americas	6 (9%)	6 (13%)	20 (12%)	11%	3
Asia	15 (23%)	10 (21%)	26 (15%)	20%	5
Europe	14 (22%)	8 (17%)	28 (16%)	18%	4
Middle East	9 (14%)	6 (13%)	25 (15%)	14%	3
<i>Total</i>	<i>64</i>	<i>48</i>	<i>172</i>	<i>100</i>	<i>21</i>

In addition to regional distribution, we strove to include cases such that our final sample reflected variation in the following dimensions, in order to generate a sample representative of different postwar dynamics:

- Capitals and not-capitals
- Large and small cities (in global comparison as well as within the postwar country)

- Conflict issue (government vs territory)
- Intensity and duration of the war, and form of conflict termination (e.g. peace agreement, victory, or stalemate/'frozen conflict')

Table 2. List of cases

City	Region	City status	Population	Conflict issue
Abidjan	Africa	Largest city	Ca 4,1 million	Government
Bujumbura	Africa	Capital	Ca 1 million	Government
Goma	Africa	Regional center	Ca 500 000	Government
Dire Dawa	Africa	Chartered city	Ca 300 000	Territory
Jijiga	Africa	Regional center	Ca 200 000	Territory
Benghazi	Africa	Former capital	Ca 1 million	Government
Bogotá	Americas	Capital	Ca 7 million	Government
Ayacucho	Americas	Regional center	Ca 150 000	Government
San Salvador	Americas	Capital	Ca 2 million	Government
Colombo	Asia	Capital	Ca 600 000	Government
Jaffna	Asia	Regional center	Ca 150 000	Territory
Kathmandu	Asia	Capital	Ca 1 million	Government
Imphal	Asia	State capital	Ca 500 000	Territory
Sittwe	Asia	Regional center	Ca 200 000	Territory
Belfast	Europe	Regional center	Ca 300 000	Territory
Mitrovica	Europe		Ca 100 000	Territory
Mostar	Europe	Regional center	Ca 100 000	Territory
Sukhumi	Europe	Regional center	Ca 70 000	Territory
Beirut	Middle East	Capital	Ca 400 000	Government
Kirkuk	Middle East	Provincial capital	Ca 1 million	Territory
Jerusalem	Middle East	Religious center	Ca 900 000	Territory

Search procedure

The project follows data collection procedures developed by the UCDP. The core of the coding is based on a systematic search in Factiva, a database that contains reports from a wide range of news sources. It includes the major news outlets such as BBC monitoring, Reuters and AFP, but also local news media and regional sources such as Africa Research Bulletin, Africa Confidential, Xinhua News as well as reports from International Crisis Group.

For each case, research assistants rely on a pre-set search string, which is further adapted to the specific postwar city in focus. The search string contains three segments. The first contains the city name (including alternate spellings). The second contains different keywords relating to violence, and is based on the UCDP core search methods, with some modification to fit the project scope. The third

Search string: (citynameA or citynameB or citynameC) and

(kill* or murder* or die* or injur* or dead* or death* or wounded or massacre* or clash* or riot* or violent or violence or rape* or stab* or bomb* or beat* or kick* or shot or shooting or hospital*) and

(([conflict + actor keywords] or war or conflict)

segment contains keywords relating to the preceding war and the actors (incl abbreviations, English and original full names) and identity dimensions involved.

The time period for the search is set according to the identification of “postwar period(s)” in the separate Postwar Cities list. The relevant country is selected in the Factiva category “Region”.

To minimize the risk that events are overlooked, the basic search contains no further restriction – i.e. the search is conducted in all sources, keywords are searched for in the full text of each article, and the search is not restricted to particular subjects. However, for some cases and time periods the search may result in an overwhelming number of hits, and the search procedure needs to be limited. Strategies for narrowing down the search include restricting the sources; using the “not”-function to exclude topics; selecting specific subjects; and searching only in “head and lead paragraph”. In such situations, research assistants compare the results from different searches and the search is discussed with the broader team in the weekly coding meeting to minimize the risk that important events are overlooked. The procedure for narrowing down the search are clearly documented in the case workbook.

For each case, the specific search procedure and the number of hits it generated, along with Factiva basic analytics (which illustrates hits over time and most frequent keywords) are documented in a workbook. The hits generated in the search are saved and clearly marked (e.g. Beirut_ 19901013-20191231_0001-0100).

Additional sources

In addition to the main search presented above, research assistants rely on additional secondary sources of relevance to the specific cases under study in order to triangulate information. This becomes particularly important for cases where news reports are scarce, in order to improve reliability of the data. Such sources include reports from global and local NGOs and independent research institutes. The additional sources consulted are documented in the separate workbook for each case.

Types of violence

As noted above, we define an event as an individual incident of physical violence, exerted by one actor against another (or by two or more actors against each other), and which occurs at a specific place and at a specific time. In order to be able to analyze dynamics of violence across time and space, we code each event according to which broad category violence it falls within:

- Assassination/murder: the killing of a person or persons where circumstances suggest a high degree of organization/ planning. May include remote violence (e.g. bombing).
- Attack: cases where clearly one side/actor attacks another actor/group or groups, with death or injury ensuing, but where the violence is less planned or organized; the victim is not necessarily specifically targeted, but can be attacked on basis of group identity etc. May include remote violence (e.g. bombing).
- Clash: violence where two sides are mutually engaged in fighting each other – a violent interaction between two sides
- Riot/violent protest: Captures mass protests or demonstrations that turn violent, for instance where protestors engage in stone-throwing, arson or other forms of violence, either against the state or against another actor.

- Sexual violence: events of rape or violent sexual assault, perpetrated against women, men, girls or boys.
- Violent abduction: cases of kidnapping resulting in injury or death
- Other: Violence that does not fall into any of these categories (specified in a separate comment).

The type of violence is coded based on observable dimensions of the violence. While we do not specifically code incidents as constituting e.g. one-sided violence against civilians, the user can combine the “type of violence” with information on the sides involved to identify events along these lines. One-sided violence is generally understood as a repertoire used by organized actors for strategic purposes such as to terrorize, force concessions, or cleanse civilians of a certain identity from an area (cf. Eck & Hultman, 2007; Joshi, 2020; Sundberg, 2009). While we are unable to directly observe strategic intentionality in the coding, the user could operationalise one-sided violence as cases of murder, attack, sexual violence and violent abduction, where Side A is an organised actor and Side B is civilian(s).

Geocoding

Each event is geocoded at maximum precision within the city. Latitude and longitude coordinates are provided for each event (in decimal degrees format (six decimals)). Some locations – such as public landmarks and government buildings – are relatively easy to geocode whereas others require more triangulation. Research assistants rely on a range of general and case-specific sources to triangulate location coordinates (including to trace different versions of place names). General sources include Google Maps and Google Earth, Mapcarta (<https://mapcarta.com/>) and Wikimapia (<https://wikimapia.org>). Other sources include static city- and neighbourhood maps (e.g. provided in reports by peacekeeping missions or humanitarian actors), place directories, history forums, and city government sources.

Maximum precision is always strived for. A few special cases can be mentioned. For cases of abduction, we code the location where the abduction took place, if possible. If that location is unknown, we code the location where the victim was found (only violent abductions are included, e.g. where a victim is later found dead, or is beaten/tortured and then released).

Sometimes there is very little geographic information to go on. Some rules of thumb for coding such cases:

- “the outskirts of the capital/city” (if no more information available) – interpreted as in the city (logic being that otherwise the news text would have stated “just outside the city” or similar).
- *More examples to be added.*

Variables in the dataset

List and description of the included variables.

Variable name	Content
id	<p>Unique identifier for each event</p> <p>2020_11_17: for the time being, this variable uses the following construction:</p> <p><i>(CITY CODE – 3 letters)YYMMDDxx_yy</i> wherein</p> <p>CITY CODE is a three-letter abbreviation for the city in which the event took place;</p> <p>YYMMDD is the event date (date_start, see further below);</p> <p>xx is a two-digit number that identifies a unique event. If a given date has only one event, then it should be numbered as “01”. It does not matter in which order the event should be numbered, so long as the xx ID is unique for that date.</p> <p>yy is a two-digit number that identifies a sub-incident within a main incident. This is because the dataset aims to be as precise as possible in terms of geocoding and types of violence. As such, an event may contain a series of sub-incidents that took place at different unique locations or involved different types of violence.</p> <ul style="list-style-type: none"> • For an event with no sub-incident, yy will take the value of “00”. • For an event with multiple sub-incidents, yy will start at “01”, “02”, and so on. If there are leftover injuries and fatalities than cannot be attributed to those sub-incident, the leftover figures will be coded in an overarching observation with yy=“00”. If there are no leftover, then the overarching event is not needed. • For cases where an overarching event spans several days, and sub-events take place on different dates within that time span, each sub-event still retains the date of the overarching event (i.e. the date in the ID and date_start may differ for these events).
UCDP_id	If relevant: ID of the event in the UCDP-GED dataset
city	Name of the city in which the event takes place
country	Name of the country in which the event takes place.
country_id	Gleditsch & Ward country id number
type_of_violence	Categorical variable describing the type of violence (see “Types of violence” above)
Weapon/method	<ul style="list-style-type: none"> • Blunt instrument • Explosives • Fire • Gunfire • Sharp instrument • Without weapon/melee

	<ul style="list-style-type: none"> • Other • Unknown
dyad	E.g. hindus vs muslims, ex-rebels vs civilians, etc. This variable is context-dependent and different dyads can be identified in different cities; research assistants strive for standardization within each postwar city.
UCDP_dyad	If relevant: ID of the dyad in the UCDP-GED dataset
side_a	The name (or a description) of Side A in the event (the attacker if the violence is unidirectional).
side_a_identity	group identity of side A, if available
side_a_UCDP_id	If relevant: ID of the side A actor in the UCDP actor dataset.
side_a_gender	Comment on gender of side A, if available
side_a_age	Comment on age of side A (incl. e.g. “youth”), if available
side_b	The name (or a description) of Side B in the event (the target if the violence is unidirectional).
side_b_identity	group identity of side B, if available
side_b_UCDP_id	If relevant: ID of the side B actor in the UCDP actor dataset
side_b_gender	Comment on gender of side B
side_b_age	Comment on age of side B
Additional_actor	If relevant, additional/“third” actors involved in the fighting – e.g. if a clash begins between two sides and peacekeepers intervene and become part of the violent event.
description	A text description of the event (typically cut & paste from news source – if so, place in quotation marks and note which article, if several) <i>[for internal use only – in public version remove this column, or standardise]</i>
cr_clarity	<p>Categorical variable reflecting the amount of certainty that the event is conflict-related.</p> <p>1: high – reliable/credible information that the event involves the formal former warring parties or clearly takes place across the identity lines of the war; and/or a clear statement is made by those involved tying it to the war</p> <p>2: medium – circumstances strongly suggest it is conflict-related, and/or unverified statements by other actors than those primarily involved in the violence, and/or information in line with clarity 1 but the information is less clear or reliable</p>

	3: low – there is some reason to believe it is conflict-related, e.g. based on group identities, but no clear statements or evidence, or a biased party alleges their opponents were responsible
clarity_comment	Text comment on basis for the cr_clarity code.
sources	Names, dates and titles of the source material from which information on the event is gathered
sources_original	Note the sources of the most credible/"best estimate" information (e.g. eyewitnesses, government spokesperson, medical staff etc)
date_start	The earliest possible date when the event took place
date_end	The last possible date when the event took place
exact_date	Binary variable coding whether the date_start and date_end information were exact (1) or not (0).
time	If possible, the time of day that the event took place
where	A text description of the location of the event (e.g., "in Tahrir square", "on XX street"
Neighbourhood/area	In what neighbourhood of the city (may be the same as "where" for imprecise events)
admin_area	The name of the relevant subcity administrative unit ¹⁴
longitude	longitude coordinates for the event (as precise as possible – see "geoprecision")
latitude	latitude coordinates for the event (as precise as possible)
geoprecision	<p>Categorical variable denoting how precisely the geocoordinates are coded.</p> <p>1: exact location (e.g. street, building, square)</p> <p>2: within a small radius of an exact location (e.g. "near the town hall" or "near" another location we know the exact location, and provide the coordinates of that exact location)</p> <p>3: neighbourhood (nb: use a middle point of the neighbourhood provide the coordinates of that point. Use the same point for all events that are coded for level 3 and the same neighbourhood.)</p> <p>4: sub-city admin area (same procedure as for level 3)</p> <p>5: city-level precision only</p>
deaths_a	The best estimate of deaths sustained by side a

¹⁴ structure varies by country/city – a separate word file for each city in the data collection presents the relevant structure and units

	For imprecise accounts (e.g. “several”, “over 100”) please refer to the UCDP vague number translator (available in the Box).
deaths_b	The best estimate of deaths sustained by side b
deaths_unknown	The best estimate of deaths of persons of unknown side affiliation
deaths_best	The best (most likely) estimate of total fatalities resulting from an event
deaths_high	The high estimate of total fatalities resulting from an event
deaths_low	The low estimate of total fatalities resulting from an event
deaths_gender	The gender of those killed, if available (e.g., “1 male”; “2 male, 1 female, 2 unknown”
Injuries_total	Number of non-fatal casualties (if possible)
Injuries_a	
Injuries_b	
Injuries_other	
Injuries_unknown	
fatalities_comment	If needed, a comment on how deaths are coded, e.g. the basis for coding deaths as side a or side b For non-fatal events, a comment on whether the violence was near-fatal (e.g. stabbing with intent to kill vs. fist fight)
Injuries_comment	If needed, a comment on how injuries are coded, e.g. if it is our estimate based on imprecise information

Limitations

It is well-documented that coding of violent events based on secondary sources is subject to different forms of reporting bias (see e.g. Croicu & Kreutz, 2017; Dietrich & Eck, 2020; Weidmann, 2015). On a general level, we seek to mitigate news bias by complementing news reports with NGO reports and other forms of non-news sources (Dietrich & Eck, 2020). The fact that cities tend to be relatively well covered by news reports (Weidmann, 2015) also increases our confidence in the data.

However, we acknowledge that a number of factors may affect the degree of quality in the data across cases, and over time. For instance, legacies of war may impede the reporting of violent events due to insecurity or repression in the first few years after the war. Case-specific dynamics may also mean that reporting from some cities is more detailed than from other cities. Finally, patterns of reporting may vary across cities: Events taking place in the city center and highly public places are more likely to be reported than events in informal settlements on the city’s outskirts.

To maximize transparency about the case-specific confidence in the data and potential factors that may affect patterns of reporting, we comment on these dynamics in the case workbook associated with each case. This way, the user can take note of the potential threats to inference or causes for uncertainty in the given case, and analyze the data with these caveats in mind.

Reference list

- Bara, C. (2020). Shifting targets: the effect of peacekeeping on postwar violence. *European Journal of International Relations*, 26(4), 979-1003.
- Brenner, N. (2014). *Implosions/explosions: Towards a study of planetary urbanization*. Berlin: Jovis.
- Bromley, S. L. (2018). Introducing the UCDP Peacemakers at Risk dataset, sub-Saharan Africa, 1989–2009. *Journal of Peace Research*, 55(1), 122-131.
- Croicu, M., & Kreutz, J. (2017). Communication technology and reports on political violence: Cross-national evidence using African events data. *Political research quarterly*, 70(1), 19-31.
- Dietrich, N., & Eck, K. (2020). Known unknowns: Media bias in the reporting of political violence. *International Interactions*, Online first, 26 Oct 2020.
- Eck, K., & Hultman, L. (2007). One-sided violence against civilians in war: Insights from new fatality data. *Journal of Peace Research*, 44(2), 233-246.
- Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research*, 39(5), 615-637.
- Gusic, I. (2019). *Contesting Peace in the Postwar City: Belfast, Mitrovica and Mostar*: Springer Nature.
- Hall, P. (1999). *Cities in civilisation* New York: Phoenix Giant.
- Joshi, M. (2020). A comparative analysis of one-sided violence and civil war peace agreement implementation. *Stability: International Journal of Security and Development*, 9(1).
- Lefebvre, H. (1996). *Writings on cities* (Vol. 63). Oxford: Blackwell.
- Mumford, L. (1938). *The culture of cities*. London: Secker & Warburg.
- Nordhaus, W. D. (2006). Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences of the USA*, 103(10), 3510-3517.
- OECD/SWAC. (2023). Africapolis: Country report Kenya Retrieved from <https://africapolis.org/en/country-report/Kenya> (accessed 17 March 2023).
- <https://africapolis.org/en/country-report/Kenya> (accessed 17 March 2023)
- Pettersson, T. (2019). *UCDP/PRIO Armed Conflict Dataset Codebook v 19.1*. Uppsala Conflict Data Program. Uppsala. Retrieved from <https://ucdp.uu.se/downloads/>
- Pettersson, T., & Öberg, M. (2020). Organized violence, 1989–2019. *Journal of Peace Research*, 57(4), 597-613.
- Richmond, O. P. (2016). *Peace formation and political order in conflict affected societies*: Oxford University Press.
- Sassen, S. (2013). Cities as one site for religion and violence. In M. Juergensmeyer, M. Kitts, & M. Jerryson (Eds.), *The Oxford Handbook of Religion and Violence* (pp. 455-466). Oxford: Oxford University Press.
- SimpleMaps. (2018). *World Cities Dataset*.
- Suhrke, A., & Berdal, M. (2013). *The Peace in Between: Post-war violence and peacebuilding*: Routledge.
- Sundberg, R. (2009). *Revisiting one-sided violence: A global and regional analysis*: Department of Peace and Conflict Research, Uppsala University.
- Sundberg, R., & Melander, E. (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, 50(4), 523-532. doi:10.1177/0022343313484347
- Tollefsen, A. F., Strand, H., & Buhaug, H. (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2), 363-374. doi:doi:10.1177/0022343311431287
- UCDP. (2020). Definitions. Retrieved 2020-04-29, from Uppsala Conflict Data Program
- UNSD. (2019). *City Population by Sex, City and City Type*. Retrieved from
- Wallensteen, P. (2015). *Quality Peace: Peacebuilding, victory and world order*: Oxford University Press.
- Weidmann, N. B. (2015). On the accuracy of media-based conflict event data. *Journal of Conflict Resolution*, 59(6), 1129-1149.