

The Historical Conflict Dataset (HCD) v1.0*

Joseph Noonan, Stockholm University Jan Teorell, Stockholm University

15 August, 2023

Introduction

This document outlines the process to merge war data covering the period of 1816 to 1945, from the Correlate of War (CoW) inter-, intra-, extra-, and non-state war datasets, the Gleditsch (2004) inter- and intra- war dataset, and the Wimmer and Min (2009) war dataset while matching countries to the International System(s) Dataset (ISD).¹ This greatly expands the geographical scope of states in the 19th century in non-Western regions. Like Gleditsch (2004) and Wimmer and Min (2009), the Historical Conflict Dataset (HCD) follows the CoW threshold of 1,000 battle deaths for a conflict to be classified as a war. The most recent iteration of the CoW datasets defines battle deaths as deaths of armed personnel who died either during active combat or from injuries or diseases resulting from the war (Dixon and Sarkees 2016; Sarkees and Wayman 2010).² Battle related deaths are counted per year (12-month period beginning with the start of the conflict). Civilian fatalities are excluded. Both Gleditsch (2004) and Wimmer and Min (2009) use the same threshold of battle deaths for war.³

The primary motivations for this dataset are to update existing classifications of wars at the state level (interstate, intrastate, and extrasystemic) with a broader definition of statehood that reduces the Eurocentric bias in determining what constitutes a state, covering the period 1816 to 1945. Additionally, this dataset, in contrast to Gleditsch (2004), includes the extrasystemic classification of conflict, making it conceptually closer to what we consider the modern gold standard of conflict research, the UCDP/PRIO Armed Conflict Dataset. From a user perspective, this dataset will make it easier for researchers to work with war data before 1946 by offering a compiled war dataset, in contrast to CoW which separately stores each conflict type in a different dataset. The ambition of this dataset is not to make methodological changes to how conflict is defined or measured, but rather to reclassify war based on a more inclusive definition of statehood and to create a dataset that increases ease of access to war data, particularly for scholars working with country-level historical data between 1816 to 1945.⁴

*The authors thank Charles Butcher and Margareta Stoltenberg for their helpful feedback and suggestions on earlier drafts of this paper and dataset. We would also like to thank Steven V. Miller for sharing R code to help with data management process at the early stages of this project. The authors are also grateful to the Swedish Research Council and Stockholm University for their support of The Historical Data Archive (H-DATA) project.

¹The versions of the datasets used are as follows uses v4.1 of the inter-, extra- and non-state CoW datasets and v5.0 of the intrastate dataset. The v5 intrastate data was updated as part of the Dixon and Sarkees (2016). The Gleditsch (2004) data ‘Expanded War Data set’ uses v2.1 updated in November 2013, following CoW’s v4.1 releases. The Wimmer and Min (2009) dataset is from replication data which does not contain a version number. An unreleased and updated version of the International System(s) Dataset v2.0 is used, which has been provided by Charles Butcher. The International System(s) Dataset v2.0 is outlined in Butcher and Griffiths (2020).

²This standardization is recent, with previous iterations of the CoW datasets having different battle death requirements, for instance in Sarkees and Schafer (2000) data on battle deaths is cumulative for interstate wars while on an *annual basis* for extrasystemic wars. As such, there is some ambiguity of how battle deaths are calculated.

³The 1,000 battle deaths threshold is also used for the UCDP/PRIO Armed Conflict Dataset, which covers the post-1945 period, for their variable *intensity*. This indicator classifies conflicts as minor (25-999 battle deaths in a given year) and war (over 1,000 battle deaths in a given year).

⁴Lyall (2020) has developed a comprehensive dataset where there is a methodological shift in how to classify wars. In his Project Mars dataset, he uses a 500 battle death threshold for wars and focuses on “conventional” warfare while forgoing classifying conflicts as interstate, intrastate, or extrasystemic. As such, Lyall (2020) represents a departure from past datasets and is not used in HCD v1.0.

This dataset is available in two formats: country-war and country-year. These datasets cover all regions of the world, and run from 1816 to 1945. The primary variables of interest are the war classifications which consist of three different types of war:

1. **interstate**: A war between two states
2. **intrastate**: A war between a state and a non-state actor, inside of its territory.
3. **extrasystemic**: A war between a state and a non-state group outside its territory.

The country-year dataset contains the following variables:

- **year**: Year
- **inter_war**: Indicator for interstate war participation (1 = interstate war participation during the year)
- **count_inter_war**: The number of interstate wars
- **inter_war_name**: Name of the interstate war(s)
- **intra_war**: Indicator intrastate war participation (1 = intrastate war participation during the year)
- **count_intra_war**: The number of intrastate wars
- **intra_war_name**: Name of the intrastate war(s)
- **external_participant_intra_state**: Indicator for states which participated in *another* state's intrastate war
- **external_participant_intra_state_side**: Indicator which classifies **external_participant_intra_state** as either participating in a war as (1) against the government, (2) in support of the government.
- **extra_war**: Indicator for extrasystemic war (between as state and a non-state group outside its own territory) participation (1 = extrasystemic war participation during the year)
- **count_extra_war**: The number of extra systemic wars
- **extra_war_name**: Name of the extra systemic war(s)

The country-war dataset contains the following variables:

- **war_name**: Name of conflict
- **war_type**: Extrasystemic, interstate, or intrastate war classification.
- **min_year**: year when conflict began
- **max_year**: year when conflict ended
- **external_participant_intra_state**: Indicator for states which participated in another country's intrastate war.
- **external_participant_intra_state_side**: Indicator which classifies **external_participant_intra_state** as either participating in a conflict as (1) against government, (2) in support of government.
- **g_deaths**: Gleditsch (2004) death estimates.⁵
- **cow_state_deaths**: CoW estimate of state battle deaths.

⁵There are high levels of uncertainty and missingness with all the death estimates. In some cases, only a single estimate is given in CoW when the HCD determines that multiple states have participated. For instance, The British-Maratha War of 1817–1818, CoW provides causality estimates the total number for all members of the Maratha Confederacy (2,000). In the HCD each individual state of the Maratha Confederacy (Nagpur, Poona, Indore and Gwalior) is included in the dataset. As we do not know the specific spread of causalities between these four states, the estimates are evenly distributed between them. As such, these estimates should be treated with caution. Note that in Gleditsch (2004) and CoW, interstate wars are broken up into two cases if a state switches side during the war. HCD only has one record for each war participant for all interstate wars. In HCD the casualty estimates for cases where a participant switched sides are the combined total deaths of the two records in Gleditsch (2004) or CoW.

- `cow_non_state_deaths`: CoW estimate of non-state battle deaths (where available) as Gleditsch (2004) does not differentiate between state and non-state battle deaths in `g_deaths`. Only available for intrastate and extrasystemic wars.

Both datasets contain the following ID variables:

- `isd_code`: ISD numeric country code
- `isd_country`: ISD country name
- `isd_region`: ISD regional classification
- `gw_code`: Gleditsch and Ward (1999) numeric country code
- `gw_character`: Gleditsch and Ward (1999) character country code
- `cow_code`: Correlates of War numeric country code
- `cow_character`: Correlates of War character country code
- `cow_country`: Correlates of War country name
- `v_dem_country`: V-Dem country name
- `v_dem_character`: V-Dem character country code
- `v_dem_code`: V-Dem numeric country code

The remainder of this text goes into greater detail about the construction of this dataset.

State System

This section discusses the different country-state systems used for this merge. The three war datasets use different definitions of what a state is. The choice of state system is highly consequential for the classification of war, as it affects whether a war is interstate, intrastate, or extrasystemic. This section provides a brief overview of these extant state systems.

The Correlates of War (CoW) project differentiates between the pre-1920 period and the post-1920 period in determining whether an entity is a state using two main criteria: a population level and diplomatic recognition. In the pre-1920 period, recognition criteria include having diplomatic missions at or above the rank of charge d'affaires with Britain and France, as well as a population over 500,000. In the post-1920 period, the entity needs to be a member of the United Nations or the League of Nations, or have a population of 500,000 or more and be recognized by two or more major powers. This removes the 500,000 population limit for states that are UN or League of Nations members. In the 19th century, this classification, with its focus on European diplomatic recognition, creates validity issues. For instance, Argentina is classified as a state by CoW in 1841 despite having declared independence 25 years earlier in 1816. This state classification leads to anomalies in CoW war data where wars involving Argentina between 1816 and 1841 are classified as extrasystemic or non-state. For example, the Brazil-Argentine War of 1826 to 1828 is classified as a non-state war. This is purely a byproduct of the recognition criteria from CoW, as Brazil was only recognized by the United Kingdom and not France in 1828 while Argentina was recognized by neither.

To rectify some of these cases, Gleditsch and Ward (G-W) (1999) introduce a new state system that builds upon the CoW state system, and provides a different classification of independent states. The authors use three points in their criteria for determining if an entity is an independent polity:

- (1) it has a relatively autonomous administration over some territory,
- (2) is considered a distinct entity by local actors or the state it is dependent on and,
- (3) has a population greater than 250,000.⁶

This is the state system used by Gleditsch (2004) in his revised war list, which is an adaptation of CoW war data to the Gleditsch and Ward (1999) state system. While the G-W state system has a more inclusive definition of a state, which greatly increase the number of states in the pre-1900 period, it does not fully capture pre-colonial states in non-Western regions of the world.⁷

Wimmer and Min (2009) do not provide a systematic list of states, but rather specify that their data covers a variety of different units including colonial states, city-states, feudal kingdoms and centrally-organized tribes.⁸ This reflects a conceptual shift, moving from a focus on the state to a focus on territories, allowing for differentiation between war prone states and war prone territories, which they argue decreases the Western bias inherent in focusing on states recognized in the international system.

The Historical Conflict Dataset retains the state-based approach used by CoW, Gleditsch (2004), and UCDP but mitigates the issue of Eurocentric bias in state definitions by using the International System(s) Dataset (ISD) as the criteria for what units constitute states (see Butcher and Griffiths (2020) for a full description of this dataset). The ISD dataset is the result of a remarkable effort to catalog pre-colonial states in Africa, South Asia, Central Asia, and South-East Asia. A key concept for this dataset is the idea of multiple loosely connected international systems, which post-1920 consolidated into a single international system. This conceptualization of international system(s) emphasizes the importance of regional state systems.

The full criteria of statehood in the International System(s) Dataset (ISD) is as follows:

- (1) A population of at least 10,000

⁶Gleditsch and Ward also provide a list of formally independent microstates, which meet the criteria of (1) and (2) but have a population lower than 250,000. These countries are Andorra, Liechtenstein, Monaco, San Marino, Abkhazia, South Ossetia, and Vatican City (Kristian Skrede Gleditsch and Ward 2013).

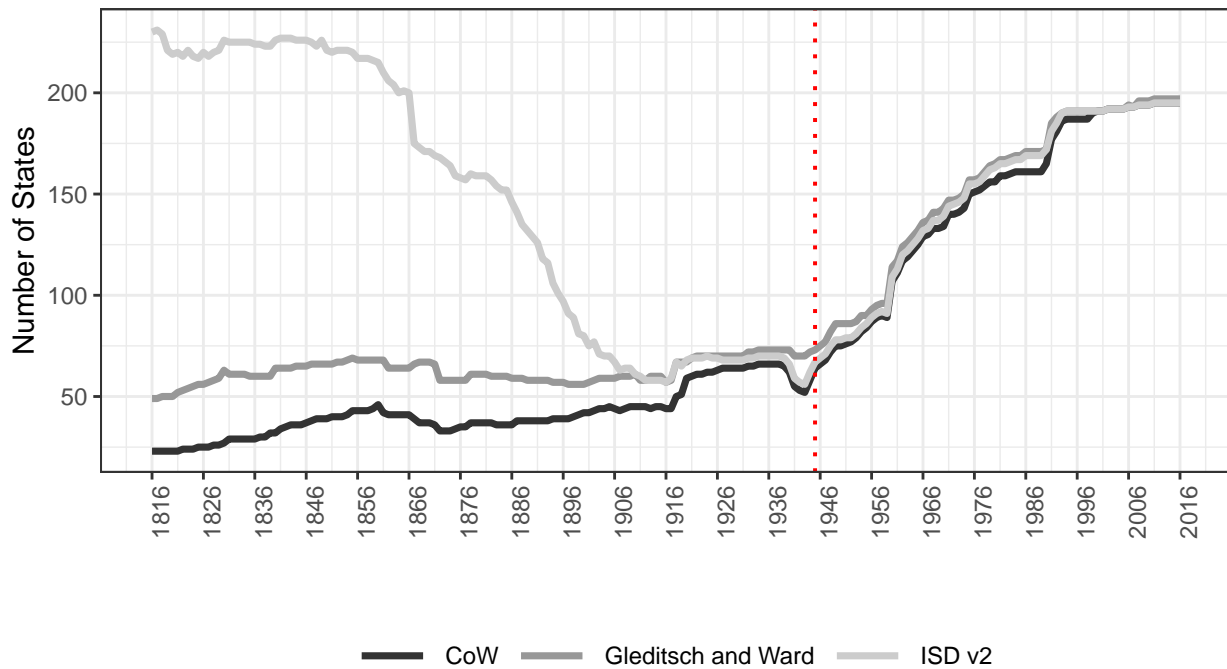
⁷The appendix provides further analysis of the descriptive differences between the G-W and the ISD state datasets which showcase these omissions.

⁸Specifically, the authors state that they take a political anthropological approach to defining the state and defines a state as a centralized political organization with a three-level administrative-political hierarchy (Wimmer and Min 2009, 394)

- (2) Autonomy over a specific territory.
- (3) Sovereignty that is either uncontested or acknowledged by the relevant international actors.

Figure 1 shows the total number of states by each state system from 1816 to 2016. Overall, these three systems are in most disagreement during the 19th century, with CoW being by far the most restrictive with their classifications of states while ISD being the most permissive. However, by the beginning of the 20th century, these three state systems began converging. The dotted red line is in 1946, representing the end date of the Historical Conflict Dataset. Overall, the ISD dataset offers unparalleled pre-1900 country coverage, capturing many extinct and pre-colonial state units which are integral for understanding the dynamics of war in the pre-1900 period.

Figure 1. A comparison of state systems over time



Adapted from Butcher and Griffiths (2020)

Note: Gleditsch and Ward counts include their tentative list of microstates.

War Reclassification

By using ISD to establish statehood, the Historical Conflict Dataset has reclassified wars taken from CoW, Gleditsch (2004) and Wimmer and Min (2009) into interstate, intrastate, and extrasystemic wars. This section begins with an overview of the Historical Conflict Dataset and how it differs from existing measures. This is followed by specific discussions of how each war type has been recoded.

Overview

Statehood is integral to the classification of war. Whether a unit is considered a state or not, has dramatic effects on the interpretation of trends of war throughout history (Butcher and Griffiths 2020). To give an example, the First British-Afghan War (1839-1842), is coded as an extrasystemic war in CoW state system, as Afghanistan is only registered as a state after The Anglo-Afghan Treaty of 1919 in CoW, presumably because CoW's definition of a state prior to 1920 is contingent on diplomatic recognition from the United Kingdom and France. In contrast, both the ISD and G-W systems code Afghanistan as a state during two periods: first

from 1816-1879 and again from 1919 onward. As such, in both Gleditsch (2004) and the Historical Conflict Dataset, the First British-Afghan War (1839-1842) is recognized as an interstate war, not an extrasystemic war. This case illustrates the problems of the narrow coding of CoW countries in the pre-1920 which requires two colonizing powers (France and Britain) recognition for statehood.

To build this dataset, all the wars were matched between datasets, to identify new wars and areas of disagreement. Here, Gleditsch (2004) was instrumental as his dataset includes warcode numbers for the CoW wars, thus works as a key to merge CoW data.⁹ Wars from Wimmer and Min (2009) were then matched based on the name, location, time frame, and description of the conflict. After wars were matched to this key, the participants were then matched to the ISD state system. To match the ISD state system, the case descriptions in Sarkees and Wayman (2010) were used. In some cases, matching to participants was relatively simple, where the case descriptions use language to indicate states, with the names corresponding to the names used in ISD. For example, for the China-Kokand War, Sarkees and Wayman (2010) clearly outline the participants as “Kokand vs. China”, both of which exist in the ISD in 1830. In other cases, especially in Africa, identifying participants can be more difficult as the case descriptions refer to states in the ISD as ethnic groups or tribes. For example, in the First Tukolor War, the description of the war in Sarkees and Wayman (2010) refer to the “Bambara tribe”. In ISD this state is referred to primarily as “Kaarta” with “Bamana Empire” as a secondary name.

If participants were sovereign states at the start of the conflict according to the ISD state system, then wars were classified as interstate. In some cases, multiple states are involved in a new interstate war. Here, the case descriptions first in Sarkees and Wayman (2010) followed by Clodfelter (2017) and Phillips and Axelrod (2005), were used to determine whether participants mentioned in the case description fulfilled the minimum criteria of war involvement according to CoW¹⁰. If only one participant had statehood, then the war was classified as either an intrastate or extrasystemic war, depending on the location and nature of the war as determined by the case descriptions and Wimmer and Min (2009).

The HCD also uses the corrections in Gleditsch (2004) which fixes issues in CoW where, due to changes in statehood in the CoW country system, wars are artificially broken up into two parts. For example, in CoW “The Taiping Rebellion” is broken up into two wars (phase 1 and 2) due to China becoming a member of the CoW’s interstate system in 1860. As the G-W state system (as well as ISD) registers China as a state from 1816 onwards, Gleditsch (2004) treats this as a single war.

Wimmer and Min (2009) is used primarily to identify new conflicts. They argue that their data collection is more stringent than CoW, only including wars where there is clear evidence from two sources of battle deaths over 1,000. However, Wimmer and Min (2009) do not provide detailed information about these cases, with their dataset only including the type of conflict, time span, locations (based on 2002 country borders), and state participants. As such, this source is primarily used to identify new conflicts. The starting years and ending years for wars are taken first from Gleditsch (2004) if available, then from CoW (for new reclassified wars stemming from non-state and extrasystemic CoW wars), and finally from Wimmer and Min (2009). In cases where states exist in both CoW state system and the ISD state system, there are no changes made to the CoW coding of war. This means that ambiguous cases that exist in some of the CoW coding are carried over into the HCD.¹¹

The results of this recoding effort are summarized in Figure 2, which shows the percentage of CoW conflicts that have been recoded by decade, from 1816 to 1945. CoW is chosen as the dataset for comparison since, unlike Gleditsch (2004) and Wimmer and Min (2009), CoW includes the extrasystemic category of conflict.

⁹Gleditsch (2004) has not updated his dataset to v5 of the CoW intrastate war data.

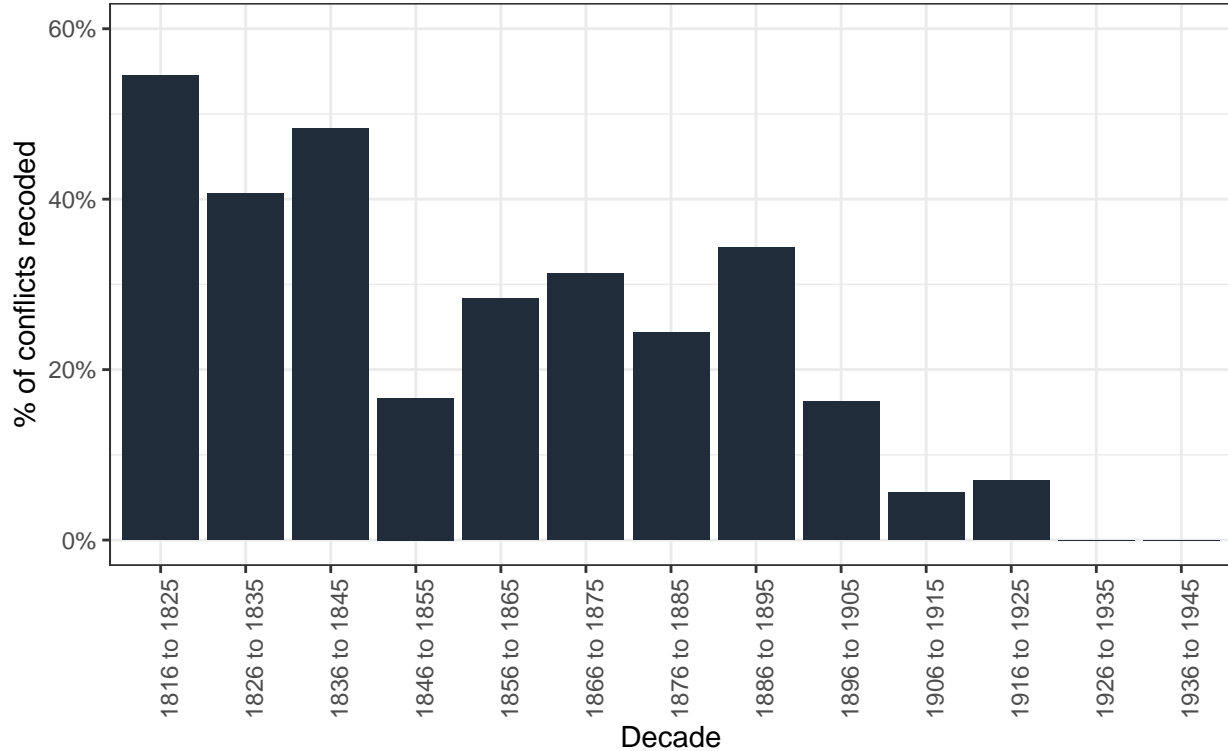
¹⁰CoW has two alternative criteria in determining war participation: (1) at least 100 battle related fatalities or (2) deploying 1,000 troops in active conflict (Sarkees and Wayman 2010).

¹¹Many of the ambiguous cases in the CoW data stem from the extrasystemic category of war. For instance, the Garibaldi Expedition of 1860 is coded as an extrasystemic war with Two Sicilies as the sole state participant. This war was a private expedition lead by a nonstate actor that invaded Two Sicilies. This does not follow the CoW coding of extrasystemic wars which is: “an extra-state war involves fighting by a state system member outside its borders against the armed forces of an entity that is not a member of the interstate system.” In this case, classifying this war as an extrasystemic war falsely gives the impression that Two Sicilies engaged in extraterritorial warfare in a territory not controlled by a member of the interstate system. Other examples of ambiguous cases include Austrian-Bosnian of 1878 and Serbian-Bulgarian 1885, both of which could be argued are interstate wars between Austria and the Ottoman Empire and Serbia and the Ottoman Empire.

Here, there is a clear drop off in the percentage of conflicts recoded after the turn of the century, with no conflicts being recoded in the post-1926 period. This follows of the dates of convergence between ISD, CoW and G-W state systems as shown in Figure 2.

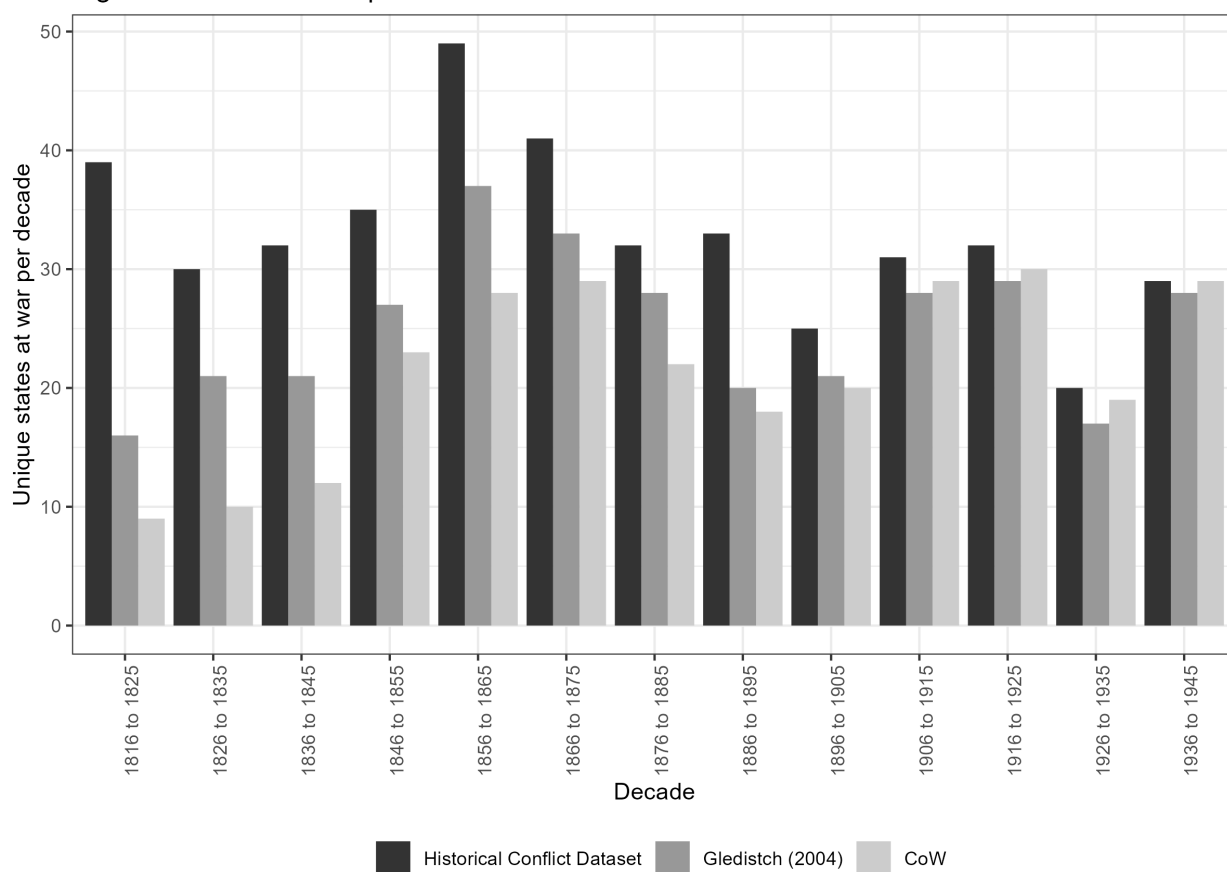
The decade with the highest percentage of recodes is 1816 to 1825. This includes the reclassification of colonial expansions by the British from extrasystemic to interstate, such as the First British-Ashanti, British-Bharatpuran, and First British-Burmese conflicts. Additionally, non-state conflicts classified as intrastate in Gleditsch (2004) were reclassified as interstate in HCD, such as the Siam-Kedah War, China-Kokand War, and Ottoman Conquest of Sudan.

Figure 2. Percentage of CoW conflicts recoded by decade



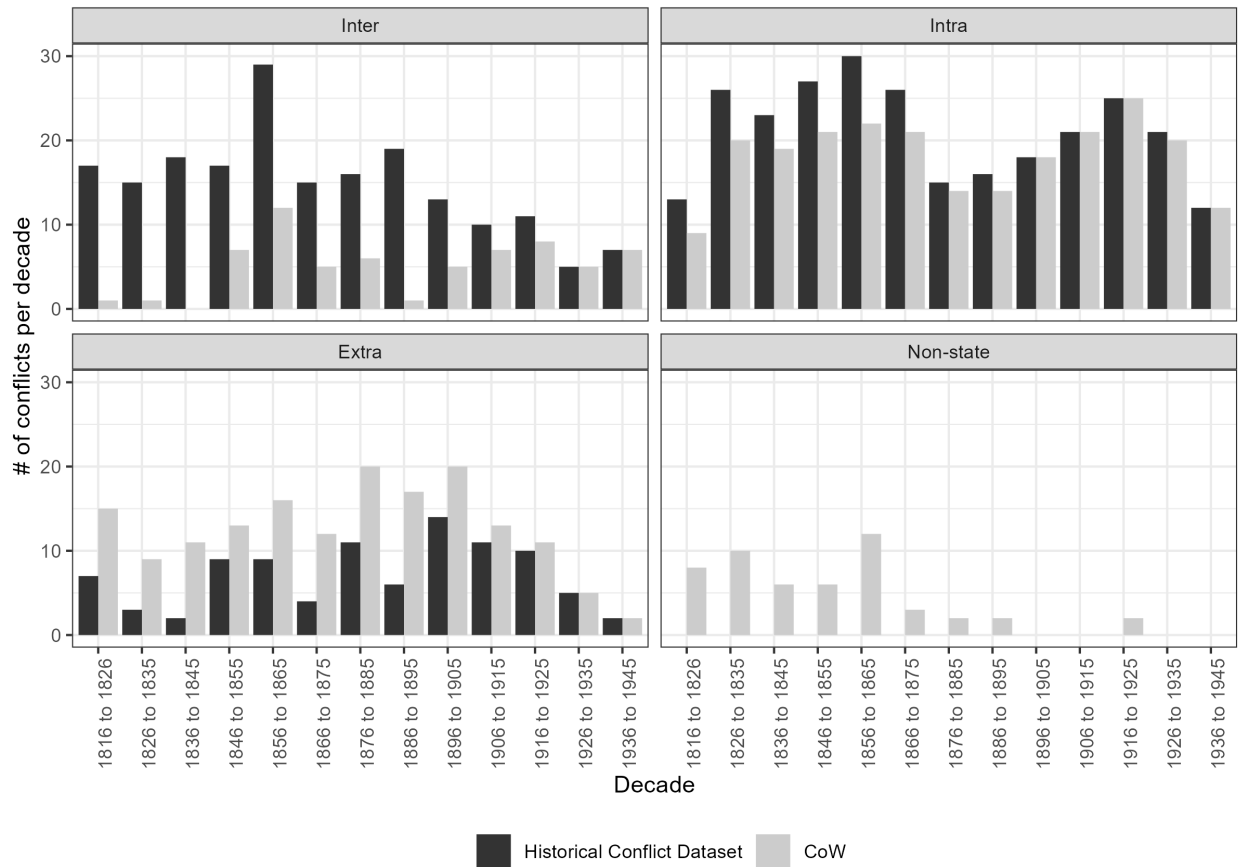
Using the ISD criteria for statehood dramatically increases the number of states involved in wars. Figure 3 graphs the number of states engaged in any type of war (inter-, intra-, extra-) by decade for CoW, Gleditsch (2004), and the Historical Conflict Dataset. Here, the HCD shows clear differences in the trends during the 19th century compared to the other two datasets, with the highest number of states in war occurring from 1816 to 1825. This is in stark contrast to CoW and Gleditsch (2004), which both record the lowest number of states involved in wars during this period. Immediately after this decade, in 1826-1835, the number of states involved in wars decreased in HCD, with a steady rise in each subsequent decade before peaking in 1856 to 1866. In addition to the differences in trends, there are large differences between the datasets in the absolute number of states engaged in war per decade. Here, the Historical Conflict Dataset registers the same number or more states engaged in conflict than Gleditsch (2004) or CoW for every decade that the HCD covers, although the difference between the datasets becomes narrower after the end of the 19th century.

Figure 3. States at war per decade



Lastly, Figure 4 shows the number of wars per decade by subtype for both CoW and the Historical Conflict Dataset. CoW is chosen as the dataset of reference here due to its inclusion of extrasystemic wars. Here, there are stark differences in the number of interstate wars. The CoW dataset records the period from 1816 to 1855 as remarkably peaceful with regard to interstate war, with the decade of 1836 to 1845 recording no wars in the CoW dataset. In contrast, the decade of 1836-1845 has the third highest number of interstate wars for any decade in the Historical Conflict Dataset. This is due to the reclassification of extrasystemic and non-state wars in CoW to interstate wars in the Historical Conflict Dataset.

Figure 4. Wars per decade, Historical Conflict Dataset and CoW



The remainder of this text goes into greater detail on the process of reclassification for each of the war types, starting with a discussion of war that take place outside the bounds of state territory (extrasystemic and non-state conflicts) followed by a discussion of wars that take within the bounds of state territory (interstate and intrastate).

Reclassification of Extrasystemic Wars

Of the datasets of interest, only CoW classify wars as extrasystemic, with both Gleditsch (2004) and Wimmer and Min (2009) only making distinctions between intra- and inter- state wars, along with a number of subcategories of these conflicts.

CoW’s “extra-state wars” are coded as follows:

“An extra-state war involves fighting by a state system member outside its borders against the armed forces of an entity **that is not a member of the interstate system**. Within the current CoW war typology, an extra-state war must meet same definitional requirements of all wars in that the war must involve sustained combat, involving organized armed forces, resulting in a minimum of 1,000 battle-related combatant fatalities within a twelve month period”.

Gleditsch (2004), forgoes classifying wars as extrasystemic arguing that the distinction between extrasystemic and inter- and intra- state wars is much more blurred than distinguishing between interstate and intrastate. However, Gleditsch (2004) does not include subtypes of wars (such as colonial conquests). This leads to unusual cases such as the British colonization of Maratha (The British-Maratha War of 1817–1818), is classified as an intrastate war in Gleditsch (2004), despite Maratha being outside of the control of the British colonial Empire when the conflict began. This coding it makes it impossible to determine if the war occurred with in the core territory of the country.

Although out of the timeframe of this dataset, the UCDP/PRIO Armed Conflict Dataset also codes extrasystemic conflict similarly to CoW:

“Between a state and a non-state group **outside** its own territory, where the government side is fighting to **retain** control of a territory **outside the state system**”

In the Historical Conflict Dataset, extrasystemic war is included as a war type as this captures colonial warfare while allowing distinctions between whether a combatant is a state or not. In order to match match extrasystemic wars to the ISD system the following steps are made. First, extrasystemic wars from CoW have been examined, and, if both participants are members of the ISD state system at the time the conflict occurred, then they are classified as interstate conflicts. For example, the CoW extrasystemic war First Franco-Dahomeyan (1890), is coded as an interstate war in this merge as ISD includes Dahomey as a state during the time of the war, while neither Gleditsch (2004) nor CoW recognizes Dahomey as a state.

In total, **66** extrasystemic wars from CoW are recoded as interstate. This can be contrasted with Gleditsch’s (2004) reclassification of extrasystemic wars to fit into the Gleditsch and Ward country system which resulted in only **31** additional interstate wars, with the other wars being classified as intrastate, many of which were colonial conquests. In cases in which there were only one ISD codes for a given war at a given time, these wars continued to be classified as extrasystemic. There were **4** cases from Wimmer and Min (2009) which had one side which was not recognized as a state in CoW or Gleditsch (2004) which were coded as extrasystemic conflicts. Because neither CoW nor Gleditsch (2004) had data regarding these conflicts, as part of this merge they were coded using Clodfelter (2017) and Phillips and Axelrod (2005) as the main source for information on these conflicts.

In total, 74 wars from CoW remain extrasystemic, with an additional 2 extrasystemic wars coming from reclassification of CoW non-state wars and an additional 4 extrasystemic wars coming from Wimmer and Min (2009).

Notably, there are four wars which move from interstate in Gleditsch (2004) to back to the CoW classification of extrasystemic in the HCD: Spanish-Santo Dominican (1865), Russo-Afghan (1885), Serbian-Bulgarian (1885), and the First Boer War (1881). This is due to the differences in coding start-years between the G-W and ISD state classifications (see the Appendix for more information). Table 1 lists all of the new extrasystemic wars from Wimmer and Min (2009), from CoW’s non-state wars and, the four wars which have moved back to extrasystemic from interstate in Gleditsch (2004).

Table 1: New Extrasystemic Wars

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Portuguese vs. Latin American patriots	Extra	-	-	Intra	Portugal	1816-1817
Bolivia’s war of independence	Extra	-	-	Intra	Spain	1824-1824
Russia vs. Circasians	Extra	-	-	Inter	Russia	1829-1840
Sayyid Said War	Extra	Intra	Non-state	-	Oman	1829-1830
Kashmir-Dards of Chilas War	Extra	-	Non-state	-	Hunza	1852-1852
Spanish-Santo Dominican	Extra	Inter	Extra	Intra	Spain	1863-1865
First Boer War	Extra	Inter	Extra	Intra	United Kingdom	1880-1881
Russo-Afghan	Extra	Inter	Extra	Inter	Russia	1885-1885
Serbian-Bulgarian	Extra	Inter	Extra	Inter	Serbia	1885-1885

British vs. Matabele and Shona	Extra	-	-	Intra	United Kingdom	1896-1897
--------------------------------	-------	---	---	-------	----------------	-----------

This expanded scope adds further granularity to these historical wars, while exemplifying issues with conflict datasets which use less inclusive state systems. As an example, the series of British-Ashanti wars during the 19th century can shed light on how the expanded ISD state system can inform conflict dynamics and state death. The Ashanti Empire was a regional powerhouse, whose population peaked at approximately 3,000,000 people, this state is not included in CoW as it lacked diplomatic representation from Britain and France. It is also not included in Gleditsch and Ward (1999). The series of wars lasted spanned 70 years during the 19th century, starting in 1824, and eventually resulting in the 1894 victory for the British wherein the Ashanti Empire became a vassal state. By using the ISD country system, the first three conflicts, until the Ashanti Empire loses sovereignty, are treated as interstate conflicts. Following the 1894 victory, the Ashanti Empire dies as a state in 1896 in ISD and therefore the final conflict in 1900, is treated as extrasystemic. In CoW all of these wars are treated as extrasystemic, while in Gleditsch (2004) they are classified as intrastate conflict. For CoW, each of these wars is an *imperial war* subtype which is against an independent political entity that does not meet CoW’s threshold for statehood while Gleditsch (2004) does not provide subtypes. In both cases, the changing status of the Ashanti Empire is not reflected in the data.

Reclassification of Non-state Wars

In CoW, non-state wars are wars that are either (1) between two non-state entities in a territory not controlled by a CoW State, or (2) conflicts between two non-state actors that take place across two different state territories. No conflicts pre-1945 have been registered as the latter type of non-state conflict.

Like with extrasystemic conflicts, Gleditsch (2004) reclassifies non-state wars according to the Gleditsch-Ward country system by classifying them as inter- or intra- state wars. In total, Gleditsch (2004) reclassifies 31 non-state wars, transforming 10 into inter- and 21 into intra- state wars. All of the CoW non-state wars have been re-examined, checking interstate wars against the ISD country system. If both participants are in the ISD country system then the conflict is classified as an interstate war. If only one side of the conflict is a state actor the conflict is either coded as an intra- or extra- state war. In these cases, coding was determined by using Sarkees and Wayman (2010) description of the conflict as a source, attempting to replicate the coding of CoW while using the ISD system.

For example, The Sayyid Said War of 1829–1830 has been classified as an extrasystemic conflict as the conflict was between Oman (which existed as a state in the ISD in 1829) and the Mazaria tribe in Mombasa (which is not recognized as a state in ISD). During this conflict Oman asserted control over Mombasa to gain access to the coastline. Mombasa before this war (having a garrison in Mombasa in 1828), however, the case description notes that full control was only established by 1830. Thus, this war was coded as extrasystemic.

In total, 20 additional non-state wars from are CoW recoded into different war types compared to Gleditsch (2004). The remaining non-state wars are not included in the final dataset.¹² Table 2 lists all of the recoded non-state wars.

Table 2: CoW Non-state Wars Reclassification

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Kashmir-Dards of Chilas War	Extra	-	Non-state	-	Hunza	1852-1852
Shaka Zulu-Bantu War	Inter	-	Non-state	Intra	Ndwanwe Kingdom, Zululand	1819-1820

¹²The appendix provides a list of all of these non-state conflicts before 1945 that are not included in the HCD.

Table 2: CoW Non-state Wars Reclassification (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Second Maori Tribal War	Inter	-	Non-state	-	Ngapuhi, Waikato	1821-1823
First Tukolor War	Inter	-	Non-state	-	Tokolor, Kaarta	1852-1854
Second Tukolor War	Inter	-	Non-state	-	Tokolor, Massina, Segou	1860-1862
Rabih Zubayr-Bornu War	Inter	-	Non-state	Intra	Kanem-Bornu, Rabih	1893-1893
First Nejd-Hejaz War	Inter	-	Non-state	-	Nejd (Saudi Arabia), Hejaz	1919-1919
Second Nejd-Hejaz War	Inter	-	Non-state	-	Nejd (Saudi Arabia), Hejaz	1924-1924
Boer-Zulu War	Intra	-	Non-state	-	Zululand	1838-1838
Dissolution of the Central American Confederation	Intra	-	Non-state	-	United Provinces of Central America	1839-1840
Han-Nien War	Intra	-	Non-state	-	China	1855-1858
First Zulu Internecine War	Intra	-	Non-state	Intra	Zululand	1856-1856
Peru-Gran Colombia War	Inter	Inter	Non-state	-	Colombia, Peru	1828-1829
Siam-Cambodia-Vietnam War	Inter	Inter	Non-state	Inter	Thailand, Annam	1831-1834
Bolivia Conquest of Peru	Inter	Inter	Non-state	Inter	Peru, Bolivia	1835-1836
Dissolution of the Bolivia-Peru Confederation	Inter	Inter	Non-state	Inter	Peru-Bolivian Confederation, Chile, Argentina	1837-1839
Persian Siege of Herat	Inter	Inter	Non-state	-	Iran, Herat	1837-1838
First Haiti-Santo Domingo War	Inter	Inter	Non-state	Intra	Haiti, Dominican Republic	1844-1845
Second Haiti-Santo Domingo	Inter	Inter	Non-state	-	Haiti, Dominican Republic	1855-1856
Filibuster War	Inter	Inter	Non-state	Inter	Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica	1856-1857

Table 2: CoW Non-state Wars Reclassification (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Central American War	Inter	Inter	Non-state	-	Guatemala, Honduras, El Salvador, Nicaragua	1863-1863
Sayyid Said War	Extra	Intra	Non-state	-	Oman	1829-1830
Burma-Assam War	Inter	Intra	Non-state	-	Myanmar, Assam	1819-1821
China-Kokand War	Inter	Intra	Non-state	-	China, Kokand	1819-1822
Siam-Kedah War	Inter	Intra	Non-state	-	Thailand, Kedah	1821-1821
Viang Chan- Siamese War	Inter	Intra	Non-state	Intra	Thailand, Vientiane	1826-1827
First Boer-Basuto War	Inter	Intra	Non-state	-	Orange Free State, Lesotho	1858-1858
Second Boer-Basuto War	Inter	Intra	Non-state	-	Transvaal, Orange Free State, Lesotho, Zululand	1865-1866
Ethiopia-Mahdi War	Inter	Intra	Non-state	-	Ethiopia, Sudan	1885-1889
Buenos Aires War	Intra	Intra	Non-state	-	Argentina	1820-1820
China-Kashgaria War	Intra	Intra	Non-state	Intra	China, Kokand	1825-1828
Mexico-Yaqui Indian War	Intra	Intra	Non-state	-	Mexico	1825-1827
Central American Confederation War	Intra	Intra	Non-state	-	United Provinces of Central America	1826-1829
Argentine War for Unity	Intra	Intra	Non-state	Intra	Argentina	1829-1831
Argentina-Ranqueles Indian War	Intra	Intra	Non-state	-	Argentina	1833-1834
War of Seven Khojas	Intra	Intra	Non-state	-	China	1847-1848
Kucha and Khoja Uprising	Intra	Intra	Non-state	Intra	China	1857-1857
First Ethiopian War	Intra	Intra	Non-state	-	Ethiopia	1858-1861
Transvaal War	Intra	Intra	Non-state	-	Transvaal	1862-1864
Second Ethiopian War	Intra	Intra	Non-state	-	Ethiopia	1868-1872
Uruguay Colorados-Blancos War	Intra	Intra	Non-state	-	Uruguay	1870-1872
Oman-Ibadi War	Intra	Intra	Non-state	-	Oman	1883-1884

Intrastate Wars

Intrastate wars in CoW are classified as wars that take place in the boundaries of a state, with sustained conflict where there is ‘effective resistance’ on both sides. CoW further subdivides these conflicts into three subtypes (civil wars, regional internal, and intercommunal), in the first two a government is a combatant, while in the intercommunal classification both participants are non-state entities. Gleditsch (2004) does not include intercommunal conflicts in his reclassification of CoW wars, only civil wars and regional internal. In the CoW intrastate war dataset, participants in regional internal conflicts are identified by the sub-regional unit i.e., Yucatan instead of Mexico. The HCD follows Gleditsch (2004) in assigning the national polity’s country code.¹³ In this merge we have dropped intercommunal conflicts. This decision is in line with Gleditsch (2004) and UCDP/PRIO Armed Conflict Dataset which only classifies conflicts with a state participant as intrastate.¹⁴

The indicator `external_participant_intra_state` is used to identify whether the country is a primary state experiencing an intrastate conflict, that is, that the conflict is occurring within its territorial boundaries, or whether the state is a participant in an external intrastate conflict. For example, the Spanish Civil War of 1936-1939, which took place in Spain where CoW and Gleditsch (2004) code Spain, Germany, Italy, Portugal and the Soviet Union as participants in the conflict. For Spain, it is not recorded as an external participant, as the conflict takes place in the territorial bounds of the country. In contrast, the other participating countries have indicators for being external participants in the intrastate conflict (`external_participant_intra_state`). This variable was coded by using the CoW variable `int1` to identify wars that were internationalized, and then taking all participants with country codes to identify state actors.¹⁵ This coding was then checked by using the Wimmer and Min (2009) location variable where available. In order to distinguish whether an external participant in an intrastate conflict fought for or against the government the indicator `external_participant_intra_state_side` has been created. Using the Spanish Civil War example, Portugal, Germany, and Italy are coded as fighting with the government (`external_participant_intra_state_side = 2`) while the USSR is coded as fighting against the government (`external_participant_intra_state_side = 1`).

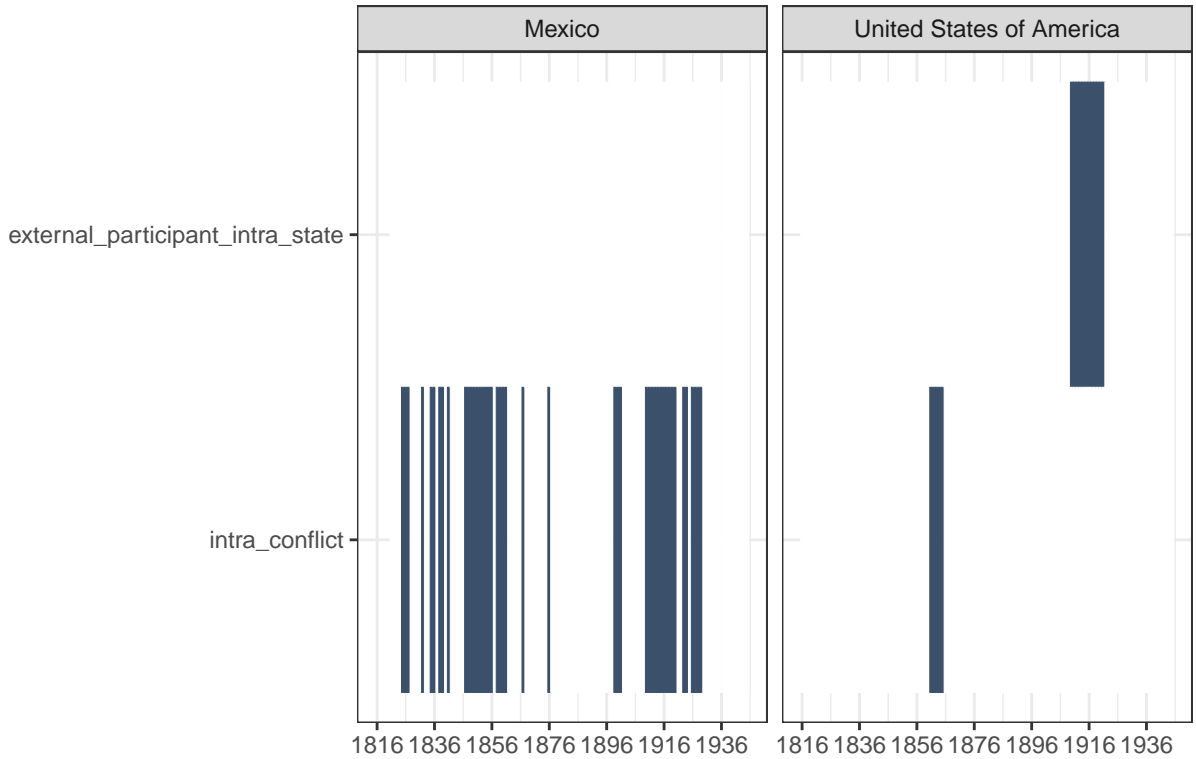
As an illustrative example of how the intrastate indicator versus the intrastate participant indicator works in practice, the graph below plots yearly occurrences of intrastate and internationalized intrastate wars in Mexico and the United States of America. Only the American Civil War from 1860 to 1865, and the Minnesota-Sioux War of 1862 are registered as an intrastate war in the USA. In 1915-1917 the United States, was a participant in the Fourth Mexican civil war and from 1918 to 1920 the US along with Finland, Japan, the UK, and France participated in the Russian Civil War. In the dataset the US is registered as being a participant in an intrastate conflict, but does not have a indicator for the intrastate conflict variable because no conflict occurred within the boundaries of the state. In contrast, Mexico has only intrastate conflict flags, including the Fourth Mexican of which, the United States was an international participant.

¹³Examples of these wars include “the Minnesota-Sioux War of 1862”, in which troops were deployed against the Santee Sioux at the direction of the Governor Ramsey of the sub-national unit of Minnesota. In the CoW dataset this conflict lacks a country code in the Historical Conflict Dataset this conflict has been assigned to the United States.

¹⁴Examples of inter-communal conflicts include “The Spartacist Rising of 1919” in Germany between socialists and German Freikorps and “The Italian Fascist War of 1920-1922” between the Fascist Blackshirts and leftist groups. UCDP classifies inter-communal conflicts as non-state, which despite bearing the same name, is conceptually distinct from the ‘non-state’ classification used by CoW.

¹⁵There is only one case, “China-Kashgaria War”, where a war that was previously classified as extrasystemic or non-state in CoW has been coded as an internationalized intrastate conflict. In this war, Kokand provided military support for an uprising in Kashgaria, a territory controlled by China. This war was coded using the case descriptions in Sarkees and Wayman (2010).

Figure 5. Intrastate Wars in Mexico and USA



In three cases, wars are classified as intrastate in all of the datasets, but due to the inclusion of both parties in the ISD dataset at the time of the war, these are classified as interstate wars. These wars are:

- First Sino-Tibetan (1912)
- Kokand Rebellion (1864-1865)
- Liberation of Peru (1824-1825)

In total, 10 intrastate wars are included that were not classified as such in either CoW or Gleditsch (2004), of which 6 are new intrastate wars from Wimmer and Min (2009). Table 3 lists these new intrastate cases. Due to the inclusion of version 5 of the CoW intrastate war dataset, the Historical Conflict Dataset includes an additional 63 intrastate wars compared to Gleditsch (2004), which uses an earlier version of the CoW Intrastate war dataset. These wars are not included in the table below.

Table 3: New Intrastate Wars

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Durrani succession wars	Intra	-	-	Intra	Afghanistan	1818-1826
Boer-Zulu War	Intra	-	Non-state	-	Zululand	1838-1838
Colorados vs. Blancos	Intra	-	-	Intra	Uruguay	1839-1842
Dissolution of the Central American Confederation	Intra	-	Non-state	-	United Provinces of Central America	1839-1840
France vs. Royalists	Intra	-	-	Intra	France	1851-1851

Table 3: New Intrastate Wars (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Han-Nien War	Intra	-	Non-state	-	China	1855-1858
First Zulu Internecine War	Intra	-	Non-state	Intra	Zululand	1856-1856
Buganda Protestant vs. Catholics	Intra	-	-	Intra	Buganda	1888-1892
Ottoman Empire vs. Cretans of 1888	Intra	-	-	Intra	Turkey	1888-1889
Austria vs. Socialists	Intra	-	-	Intra	Austria	1934-1934

Lastly, there are three intrastate wars from Wimmer and Min (2009) that are dropped to the state not being recognized by ISD, and are thus non-state wars which are not included in the dataset.¹⁶

Interstate Wars

Wars are reclassified into interstate wars if at least two sides in the conflict exist in the ISD dataset at the start of the conflict.¹⁷ The HCD reclassifies 90 wars from CoW, with 66 originating from extrasystemic wars and 24 from non-state wars. This is over twice the number of wars reclassified compared to Gleditsch (2004), which only reclassified 41 extrasystemic or non-state wars as interstate wars. Additionally, the Wimmer and Min (2009) dataset contributed an additional 14 wars. Table 5 lists all interstate wars that were previously classified as another type of war in CoW or Gleditsch (2004) along with unique new wars from Wimmer and Min (2009).

Table 4: New Interstate Wars

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Ottoman-Wahhabi	Inter	Intra	Extra	Inter	Turkey, Nejd (Saudi Arabia)	1816-1818
British-Maratha	Inter	Intra	Extra	Inter	United Kingdom, Indore, Poona, Gwalior, Nagpur	1817-1818
Dahomey's expansion	Inter	-	-	Inter	Dahomey, Oyo Empire	1818-1833
Burma-Assam War	Inter	Intra	Non-state	-	Myanmar, Assam	1819-1821
China-Kokand War	Inter	Intra	Non-state	-	China, Kokand	1819-1822

¹⁶See the Appendix for more information on these cases.

¹⁷In 10 cases one state participant is conquered leading to state death. Here, the war ends when only one participant exists in the ISD. For example, the First Dutch-Achinese War runs from 1873 to 1878 in CoW. However, in ISD, Aceh stops being a state in the end of 1874. In the HCD the the conflict only runs from 1873 to 1874.

Table 4: New Interstate Wars (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Shaka Zulu-Bantu War	Inter	-	Non-state	Intra	Ndwanwe Kingdom, Zululand	1819-1820
Ottoman Conquest of Sudan	Inter	Intra	Extra	Inter	Turkey, Funj Sultanate, Shilluk Kingdom	1820-1821
Second Bolivar Expedition	Inter	Intra	Extra	-	Colombia, Peru, Argentina, Spain	1821-1822
Second Maori Tribal War	Inter	-	Non-state	-	Ngapuhi, Waikato	1821-1823
Siam-Kedah War	Inter	Intra	Non-state	-	Thailand, Kedah	1821-1821
First British-Ashanti	Inter	Intra	Extra	Inter	United Kingdom, Ashanti	1824-1826
Liberation of Peru	Inter	Intra	Extra	Intra	Colombia, Peru, Chile, Argentina, Spain	1824-1825
British-Bharatpuran	Inter	Intra	Extra	Inter	United Kingdom, Bharatpur	1825-1826
Viang Chan- Siamese War	Inter	Intra	Non-state	Intra	Thailand, Vientiane	1826-1827
Bolivia vs Peru	Inter	-	-	Inter	Peru, Bolivia	1827-1829
Ottoman-Bilmez-Asiri War	Inter	Intra	Extra	-	Turkey, Egypt, Asir	1832-1837
Afghanistan vs. Sikh	Inter	-	-	Inter	Afghanistan, Punjab	1836-1836
First British-Zulu	Inter	Intra	Extra	Inter	United Kingdom, Zululand	1838-1838
Iran vs. Afghanistan	Inter	-	-	Inter	Iran, Afghanistan	1838-1838
Russo-Khivan	Inter	-	-	Inter	Russia, Khiva	1839-1839
Thailand vs. Vietnam over Cambodia	Inter	-	-	Inter	Thailand, Annam	1841-1845
British-Sind	Inter	Intra	Extra	Inter	United Kingdom, Sind	1843-1843

Table 4: New Interstate Wars (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
First British-Sikh	Inter	Intra	Extra	Inter	United Kingdom, Punjab	1845-1846
First Tukolor War	Inter	-	Non-state	-	Tokolor, Kaarta	1852-1854
Siege of Montevideo	Inter	-	-	Inter	Brazil, Argentina, Uruguay, United Kingdom, France	1852-1852
Tukolor vs. Bambara	Inter	-	-	Inter	Tokolor, Segou	1852-1862
French-Tukolor War	Inter	Intra	Extra	Inter	France, Tokolor, Fouta Toro	1854-1857
Hodeida Seige	Inter	Intra	Extra	-	Turkey, Asir	1856-1856
First Boer-Basuto War	Inter	Intra	Non-state	-	Orange Free State, Lesotho	1858-1858
Netherlands-Bone	Inter	Intra	Extra	-	Netherlands, Bone	1859-1860
Italo-Roman	Inter	-	-	Inter	Italy, Papal States	1860-1860
Second Tukolor War	Inter	-	Non-state	-	Tokolor, Massina, Segou	1860-1862
Russian-Kokand	Inter	Intra	Extra	Inter	Russia, Kokand	1864-1865
British-Bhutanese	Inter	Intra	Extra	Inter	United Kingdom, Bhutan	1865-1865
Second Boer-Basuto War	Inter	Intra	Non-state	-	Transvaal, Orange Free State, Lesotho, Zululand	1865-1866
Spanish-Chilean	Inter	-	-	Inter	Peru, Chile, Spain	1865-1866
Russian-Bukharan	Inter	Intra	Extra	-	Russia, Bukhara	1866-1866
Ottoman Conquest of Arabia	Inter	Intra	Extra	-	Turkey, Asir	1870-1872
First Dutch-Achinese	Inter	Intra	Extra	Inter	Netherlands, Aceh	1873-1874

Table 4: New Interstate Wars (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Second British-Ashanti	Inter	Intra	Extra	Inter	United Kingdom, Ashanti	1873-1874
Egyptian-Ethiopian	Inter	Intra	Extra	Inter	Ethiopia, Egypt	1875-1876
Kokand Rebellion	Inter	Intra	Extra	Intra	Russia, Kokand	1875-1876
Second British-Zulu	Inter	Intra	Extra	Inter	United Kingdom, Zululand	1879-1879
Franco-Tunisian	Inter	Intra	Extra	Inter	France, Tunisia	1881-1882
Ethiopia-Mahdi War	Inter	Intra	Non-state	-	Ethiopia, Sudan	1885-1889
French-Mandinka	Inter	Intra	Extra	Inter	France, Mandinka Empire	1885-1886
Sudan vs. Ethiopia	Inter	-	-	Inter	Ethiopia, Sudan	1889-1889
First Franco-Dahomeyan	Inter	Intra	Extra	Inter	France, Dahomey	1890-1890
Franco-Jolof	Inter	Intra	Extra	Inter	France, Tokolor, Fouta Toro	1890-1891
Franco-Senegalese of 1890	Inter	-	-	Inter	France, Fouta Toro	1890-1891
Second Franco-Dahomeyan	Inter	Intra	Extra	-	France, Dahomey	1892-1893
British vs. Matabele	Inter	-	-	Inter	United Kingdom, Ndebele Kingdom	1893-1893
Mahdist-Italian	Inter	Intra	Extra	-	Italy, Sudan	1893-1894
Rabih Zubayr-Bornu War	Inter	-	Non-state	Intra	Kanem-Bornu, Rabih	1893-1893
Third British-Ashanti	Inter	Intra	Extra	Inter	United Kingdom, Ashanti	1893-1894
Second Dutch-Bali	Inter	Intra	Extra	Inter	Netherlands, Mataram Lombok	1894-1894
Portuguese-Gaza Empire	Inter	Intra	Extra	-	Portugal, Gaza Empire	1895-1895

Table 4: New Interstate Wars (*continued*)

HCD War Name	HCD Conflict Type	Gleditsch (2004)	CoW	Wimmer and Min (2009)	Countries	Years
Second British-Mahdi	Inter	Intra	Extra	Inter	United Kingdom, Sudan	1896-1898
British-South Nigerian	Inter	Intra	Extra	Inter	United Kingdom, Sokoto Caliphate	1897-1897
French Conquest of Chad	Inter	Intra	Extra	-	France, Rabih	1899-1900
Bailundu Revolt	Inter	Intra	Extra	-	Portugal, Ovimbundu	1902-1902
British Conquest of Kano & Sokoto	Inter	Intra	Extra	Inter	United Kingdom, Sokoto Caliphate	1903-1903
French Conquest of Wadai	Inter	Intra	Extra	Intra	France, Wadai	1909-1909
First Sino-Tibetan	Inter	Intra	Extra	Intra	China, Tibet	1912-1913
First Nejd-Hejaz War	Inter	-	Non-state	-	Nejd (Saudi Arabia), Hejaz	1919-1919
Second Nejd-Hejaz War	Inter	-	Non-state	-	Nejd (Saudi Arabia), Hejaz	1924-1924

Summary

This section provides a brief overview of how the Historical Conflict Dataset has taken wars from each of the extant datasets and reclassifies them. Figure 6 below shows the reclassifications from the CoW datasets to the Historical Conflict Dataset. On the left side of these graphs are the classifications from wars in an original dataset with the classifications on the right side of the graph. The bands between them show the movement between the different classifications. Note the following visualizations only include wars that are in both the original dataset and the Historical Conflict Dataset. Here, the largest changes are the reclassification of extrasystemic wars to interstate and the reclassification of non-state wars. Intrastate wars remain stable. This is due to the stringent state classification that CoW uses, meaning there are no cases in which there is an intrastate conflict in a state that is recognized by CoW but not by the ISD.

Figure 6. CoW Reclassifications

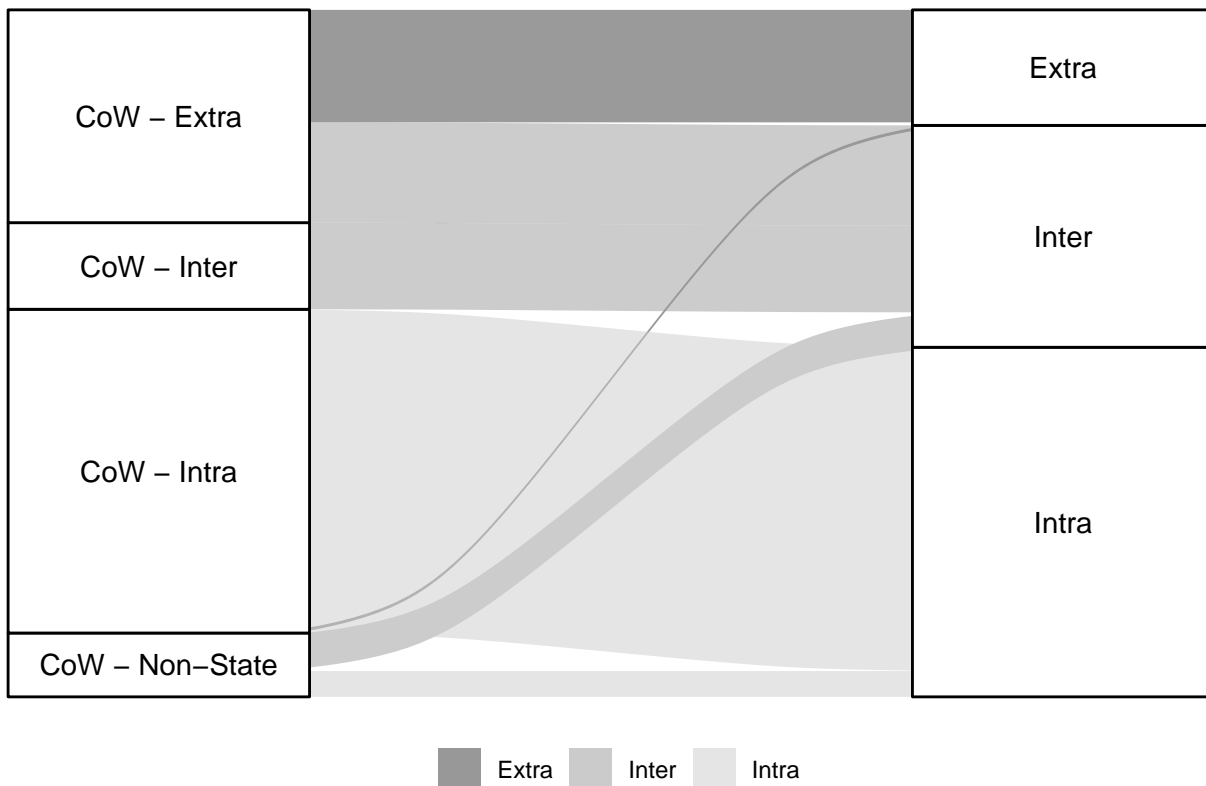


Figure 7 repeats the same style of visualization but for Gleditsch (2004). Here, the re-introduction of the extrasystemic category is the biggest change. Additionally, a number of intrastate wars in Gleditsch (2004) are recoded as interstate – these are wars that were originally coded as extrasystemic in CoW but reclassified as intrastate wars by Gleditsch (2004). Lastly, there are 4 wars that were coded as interstate in Gleditsch (2004) but have been reverted to extasystemic status (Spanish-Santo Dominican, First Boer War, Russo-Afghan, Serbian-Bulgarian). This is due to differences between ISD and G-W statehood requirements, likely due to the formers heightened importance of external sovereignty.

Figure 7. Gleditsch (2004) Reclassifications

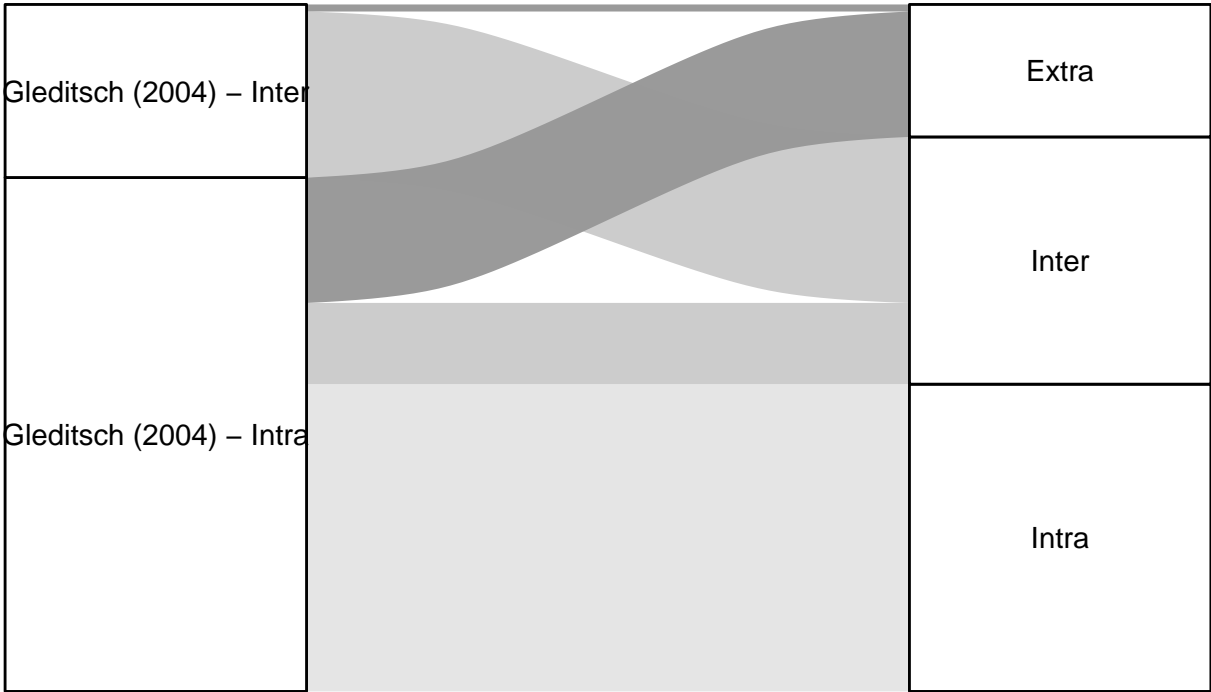


Figure 8 and Figure 9 plot the same chart for Wimmer and Min (2009). Here, Figure 8 shows how wars matched with CoW and Gleditsch (2004) have been recoded. Generally, intra- and interstate wars have remained in the same categories, with a smaller number being classified as extrasystemic. Figure 9 shows how new wars, that is, wars not matched to CoW or Gleditsch (2004), have been reclassified.

Figure 8. Wimmer and Min (2009) Reclassifications

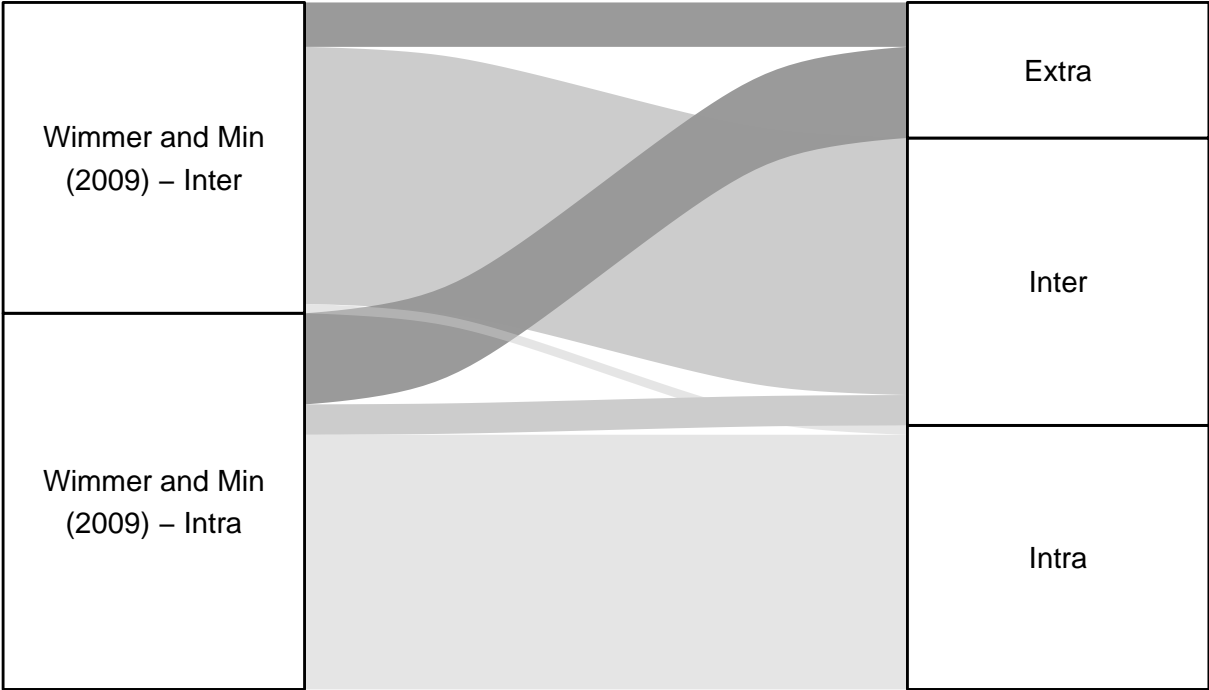
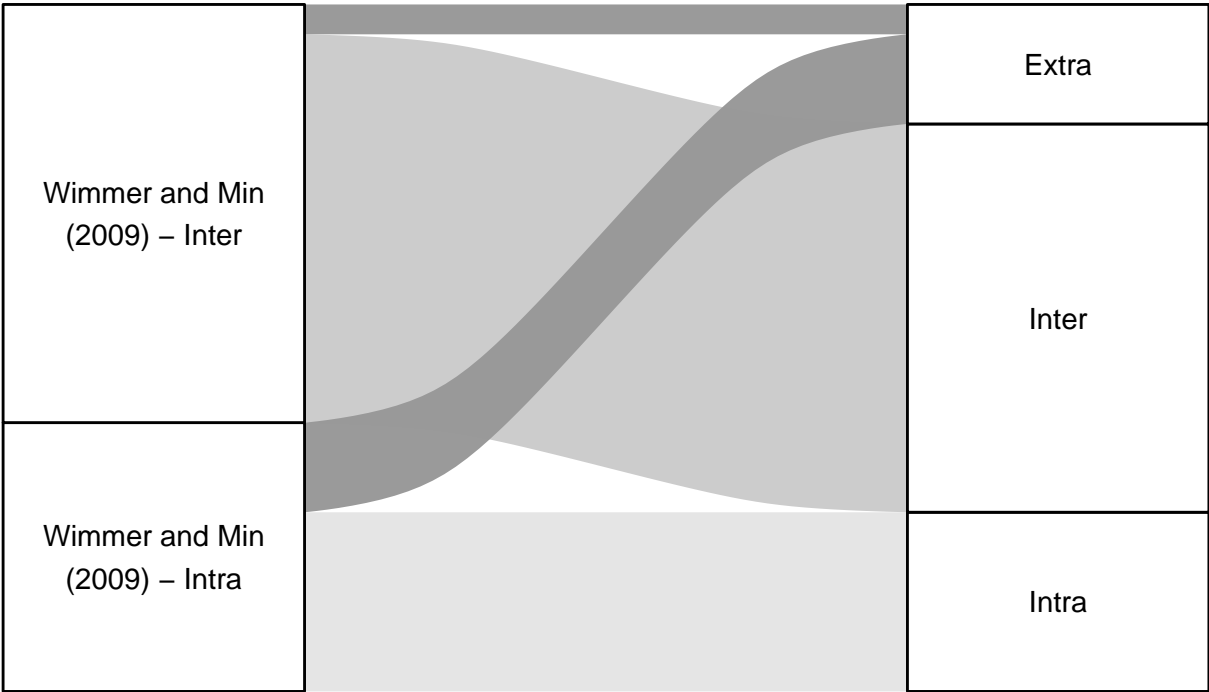


Figure 9. Wimmer and Min (2009) New War Coding



■ Extra ■ Inter ■ Intra

Limitations

While the Historical Conflict Dataset succeeds in its inclusion of additional non-Western states, reducing the euro-centric bias of CoW while having a broader classification of war compared to Gleditsch (2004), it should be noted that this is a first step in expanding our conception of conflict before 1945. This is in large part due to the heavy reliance of CoW data, of which, coverage for non-European polities may be more limited especially regarding intrastate wars (See Kristian Skrede Gleditsch, Metternich, and Ruggeri (2014)). As an attempt to mitigate this, additional wars have been brought in from Wimmer and Min (2009) whose explicit focus is to expand conflict coverage to one more global in nature. This has resulted in 23 new wars in total, of which 12 of which are in Asia, Africa or the Middle East. However, for certain indicators such as `external_participant_intra_state`, and the participants flagged will still be biased towards CoW's Eurocentric view of statehood, as non-state and extrastate war case descriptions in Sarkees and Wayman (2010) might miss important information about state participants and the nature of the wars, as data was collected assuming one side was a non-state participant. Further research should be conducted to find additional participants who meet the threshold of statehood in ISD in these conflicts. These future additions to this dataset could include using the more expansive data collected by Brecke (1999), Lyall (2020) or, alternatively, data on battles from Miller and Bakar (2022), all of which provide additional information and granular data about historical non-Western conflicts. Lastly, establishing accurate casualty counts to determine if a conflict reaches the 1,000 battle death threshold in the historical era becomes increasingly difficult going further back in time in the historical period, especially with regards to non-Western states.

Appendix

Compatibility with UCDP

The UCDP/PRIO Armed Conflict Dataset provides high quality, granular data on armed conflicts for the post-1945 period. As such, researchers may be interested in how compatible the Historical Conflict Dataset is with the UCDP/PRIO Armed Conflict Dataset, particularly if they are interested in attempting to merge or extend UCDP data into the past. This section of the appendix provides a brief overview of the compatibility of the Historical Conflict Dataset with UCDP.

State system

UCDP, classify a state in two ways:

- (1) that a state is internationally recognized, sovereign, and controlling a specific territory, or
- (2) the entity is unrecognized, controlling a territory which is not disputed by an internationally recognized government which has *previously* controlled the same territory.

Notably, in contrast to Gleditsch and Ward (1999), UCPD does not set a population threshold. Furthermore, recognition and contestation of sovereignty are made more explicit in the UCDP definition of a state in contrast with Gleditsch and Ward (1999). In practice, this definition aligns closely with Gleditsch and Ward (1999), with only small changes occurring in the post-1945 period (see Pettersson (2022) for a full description of the differences). However, the post-1945 period in general has higher levels agreement between different state-system datasets as this period represents a crystallization of the international system.

Using the ISD, the Historical Conflict Dataset captures many more states in the pre-1900 period than Gleditsch and Ward (1999), including “microstates” with populations under 250,000. This is inline with UCDP’s definition that does not put a minimum population threshold for states and results in a much more expansive coverage of states. Secondly, the ISD focus on acknowledgment or lack of contestation of sovereignty fulfills a critical dimension that is missing in the Gleditsch and Ward (1999) definition compared to the UCDP definition, namely that if the territory is unrecognized, then it should not be disputed by a sovereign government previously controlling the same territory. Criteria (3) from ISD and criteria (2) from UCDP are not identical, as UCDP is concerned with the contestation by states who **previously controlled** the territory in question. Thus, ISD has a broader scope with its conception of sovereignty by emphasizing the external dimension, which UCDP’s definition does not.

War

Regarding how conflict is conceptualized, there are clear difference between the Historical Conflict Dataset and the UCDP/PRIO Armed Conflict Dataset. Firstly, UCDP/PRIO Armed Conflict Dataset includes conflicts if the battle deaths are over 25 rather than the 1,000 battle death limit set by the CoW, Gleditsch (2004), Wimmer and Min (2009) and the Historical Conflict Dataset. Secondly, the UCDP/PRIO Armed Conflict Dataset includes a fourth conflict type *internationalized intrastate*, which is a conflict type between a government, one or more rebel groups, in which foreign states have deployed troops to either, or both sides. The Historical Conflict Dataset does include a `participant_intra_state` variable, which denotes if a given country was a participant in another states intrastate conflict, however, this inclusion is based of off CoW and Gleditsch (2004) in which states are only included as intrastate participants if they commit over 1,000 troops, or suffer 100 battle deaths. This is in contrast to the UCDP/PRIO definition which does not specify a troop limit, writing only that “there is involvement of foreign governments with troops”. Looking at conflicts where there is coverage between the two datasets (post-1945) there was agreement approximately ~45% of the time, with UCDP being more likely to code a conflict as being internationalized¹⁸.

To further examine the comparability of HCD with this dataset the period between 1946 to 1956 was coded.¹⁹

¹⁸Only conflicts with over 1,000 battle deaths are matched as this is the threshold for CoW’s war classification.

¹⁹For the extension, the analysis was done using an early draft of the HCD which used v4.1 CoW Intrastate dataset. Note that UCDP does not provide names to conflicts providing only a conflict code. To increase readability, conflicts are given unique names using the Side A and Side B variables in the UCDP dataset.

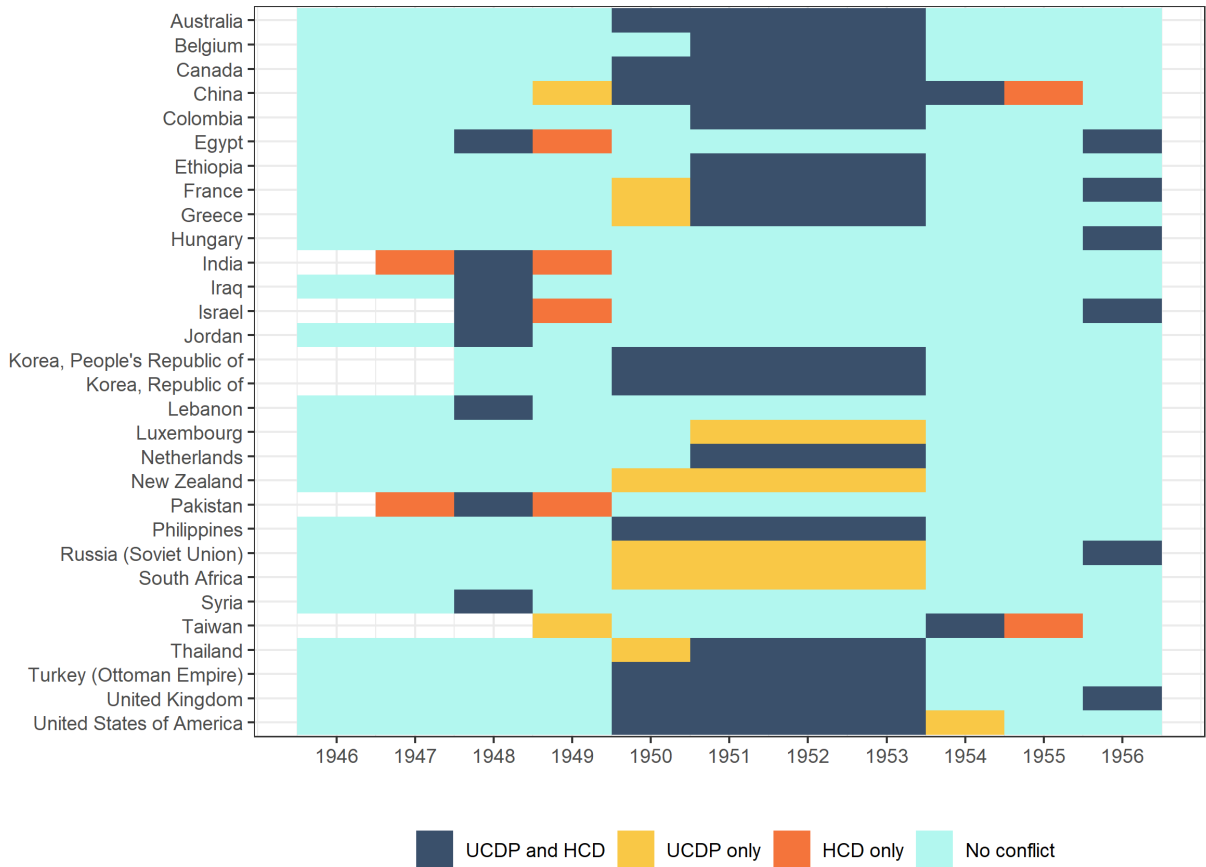
During this period UCDP identifies 69 conflicts, 30 with a maximum intensity of 1 indicating a conflict with 25 to 999 battle deaths in a conflict year, and 39 with a maximum intensity of 2 indicating over 1,000 combat deaths i.e. wars. 28 out of 39 UCDP conflicts with a maximum intensity of 2 indicating 1,000+ combat deaths have been matched. Additionally, 8 conflicts that CoW codes as having over 1,000 battle deaths, are matched which are classified as having a maximum intensity of 1 in the UCDP data.

Of all of the conflicts matched, only one conflict – Government of Hyderabad - Government of India, has a mismatch in classification, with UCDP recognizing it as an interstate conflict, while this merge, with the lack of Hyderabad in the ISD at the time, classifies this as an extrasystemic conflict. Hyderabad is only recognized by UCDP and not by Gleditsch and Ward (2004). In 7 of the conflicts, multiple UCDP conflicts have been matched to a single CoW conflict. There is also broad disagreement with conflicts starting and ending dates. For instance, the “First Burmese” conflict from CoW covers the period 1948 to 1951, which in UCDP covers two conflicts (Government of Myanmar (Burma) - KNUP and Government of Myanmar (Burma) - CPB, PVO - “White Band” faction), which together cover a period of 1948 to 1955.

In total, 18 of 31 unique matched UCDP conflicts, have a full match for years (58%) while all conflicts that have been matched have some partial year matches. To visualize the difference in year coverage, Figure A1-A3. graphs every polity during this period that has been involved in a conflict for each respective conflict type. For UCDP conflicts, only those years which have an intensity level of 2 that is, over 1,000 battle deaths, are included. Note, that these figures only show whether a country has been involved in a given conflict type in a given year, thus, differences in the amount of conflicts per year are not reflected in these visualizations.

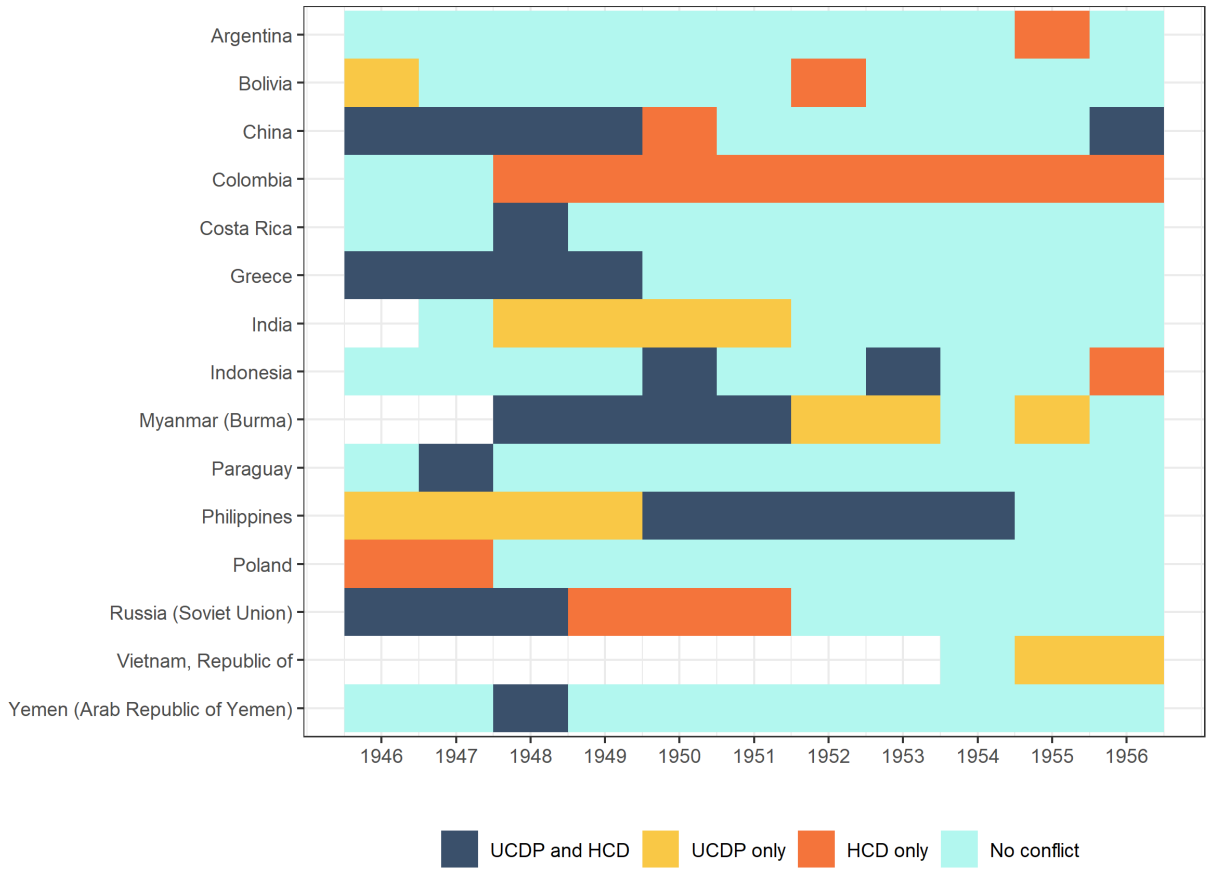
For interstate conflicts, there is generally high levels of agreement between the Historical Conflict Dataset and UCDP, with some disagreement on the start and end date of conflicts. Notably, UCDP includes New Zealand, Russia, South Africa and Luxembourg as participants in the Korean War, whereas the Historical Conflict Dataset does not. This is due to the lower threshold in UCDP to be recorded as a participant, as troop deployment is enough compared to CoW where a participant must either deploy over 1,000 troops to the conflict or suffer 100 battle deaths.

Figure A1. Interstate Conflict Comparison



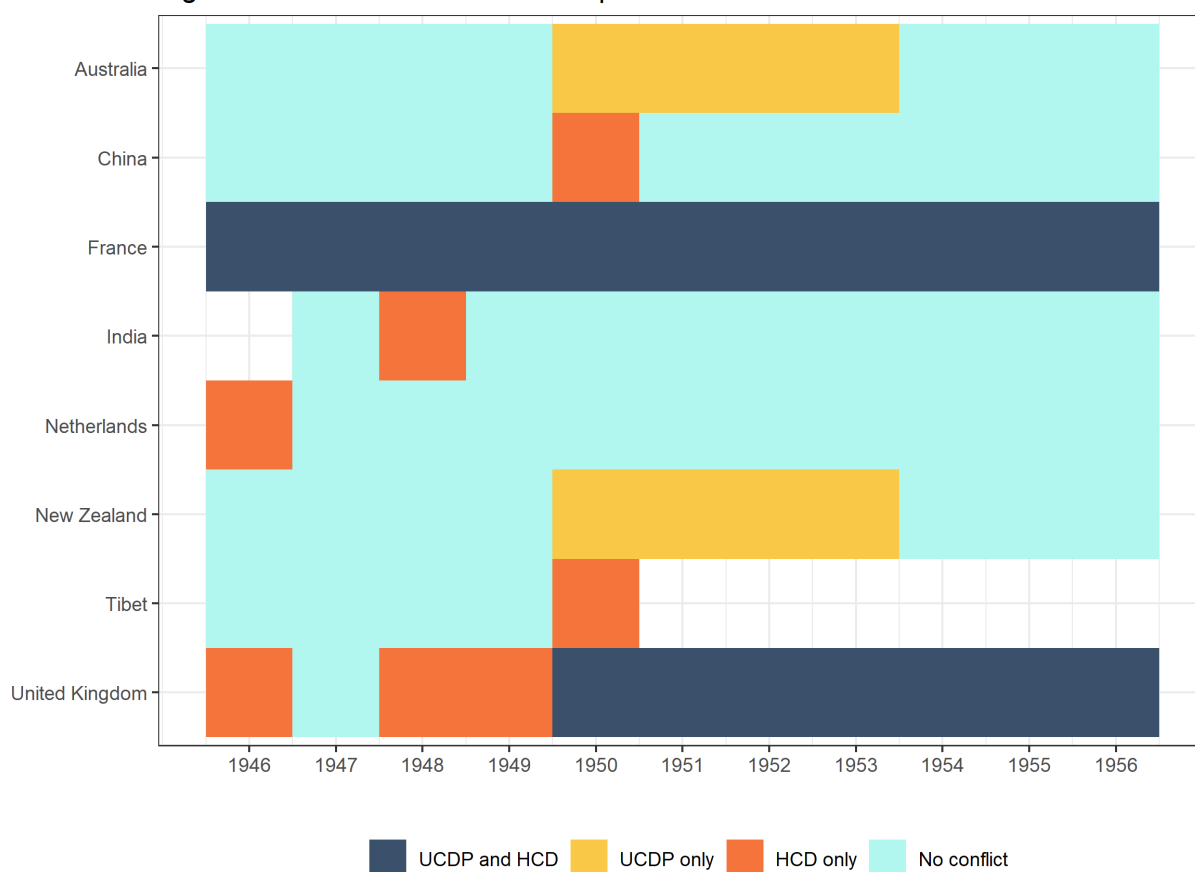
There is more disagreement between the two datasets regarding intrastate wars. While both the war merge and UCDP capture notable civil conflicts such as the Chinese and Greek Civil Wars, there are a number of country years which are only covered by one or the other datasets. Here, conflicts during this period in Colombia (Seventh Colombian “La Violencia” in CoW and Gleditsch (2004), Colombia vs. Liberals of 1949 and Colombia vs. Conservatives in Wimmer and Min (2009)) are not in the UCDP dataset. UCDP captures conflicts such in Bolivia (Government of Bolivia - Popular Revolutionary Movement), has different starting dates for the Huk conflict in the Philippines, and classifies the starting of the Vietnam conflict (Government of South Vietnam - FNL) as starting in 1956 as opposed to 1960 in CoW and Gleditsch (2004), similarly the starting dates for conflicts in Myanmar differ between the datasets.

Figure A2. Intrastate Conflict Comparison



Lastly, there are similar differences between the two datasets when looking at extrasystemic conflicts. Notably, UCDP captures Australia and New Zealand’s involvement in the Malayan Rebellion, an extrasystemic conflict fought primarily by the United Kingdom against Malay insurgents. The 1946 conflict in Indonesia with Netherlands and the United Kingdom participating, is missing in UCDP as that dataset does not register over 1,000 battle deaths for this conflict, while CoW does.

Figure A3. Extrastate Conflict Comparison



Country Systems

This annex provides some descriptive comparisons between the ISD and the G-W state system.²⁰

Table 5 shows all of the cases where the start and end date differ between G-W and ISD. There are a number of notable differences between ISD and G-W. In several cases, ISD has a much earlier start date than G-W (Ethiopia -39 years, Saudi Arabia -25 years, South Africa -25) while in other cases, ISD is much stricter than G-W and closer in agreement to CoW (Canada +52, Bulgaria +30, Luxembourg +23, Bhutan +22). The examples of commonwealth countries such as Canada, Australia and New Zealand are illustrative of the effect of the criteria of external sovereignty, each with a much later starting date in ISD (1920) than G-W. In the post-1950 period, five countries are exceptionally late in being qualified as states in ISD compared to G-W: Oman (1971), Bhutan (1971), and the microstates of Monaco (1993), San Marino (1992), and Andorra (1993) all of which gained statehood after joining the United Nations.

Lastly, the G-W state system includes the contested territories of Abkhazia and South Ossetia, as formally independent microstates, meaning they fulfill all criteria of statehood with exception of the population threshold, whereas ISD does not recognize these states. This difference explains the -2 net difference between G-W and ISD from 2008 forward.²¹

²⁰The comparisons here include the list of microstates provided by Kristian Skrede Gleditsch and Ward (2013), which includes a number of states with a population of less than 250,000.

²¹These states are not recognized in the UCDP datasets.

Table 5: ISD G-W comparison

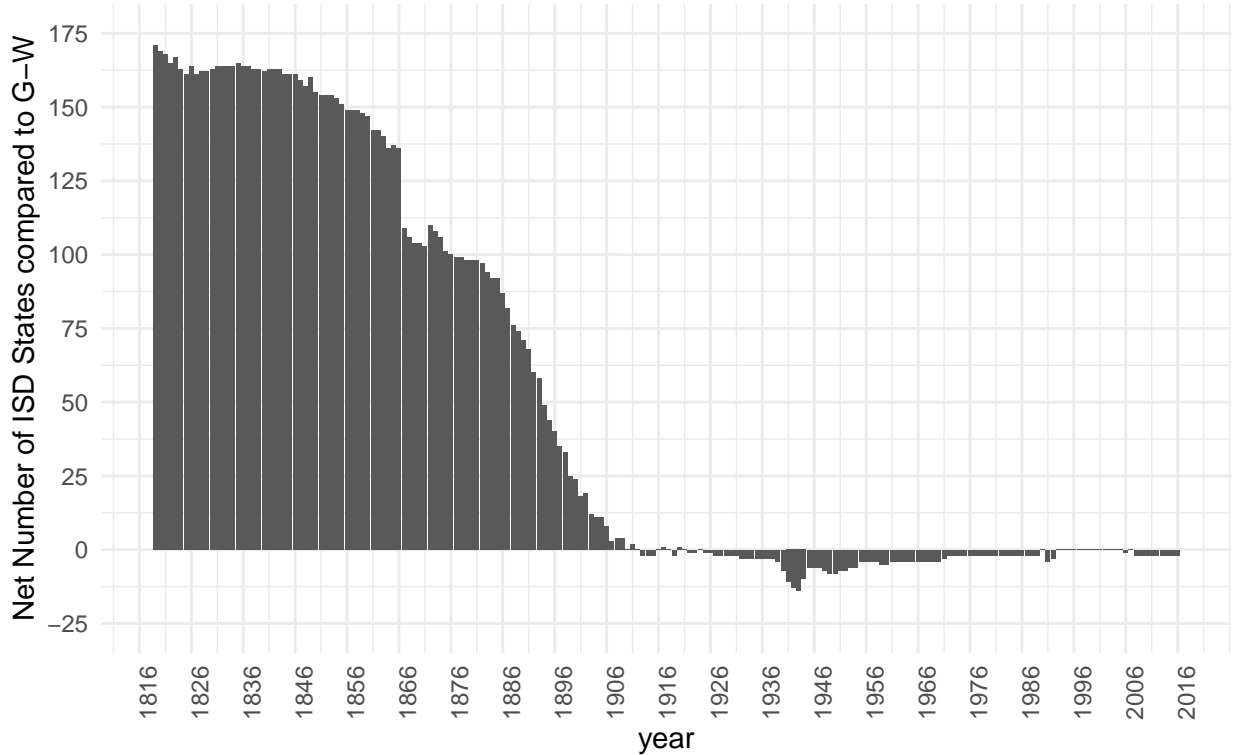
Country	ISD Start (min)	GW Start (min)	Start Difference	ISD End (max)	GW End (max)	End Difference
Afghanistan	1816	1816	0	1879	1888	-9
Algeria	1816	1816	0	1847	1830	17
Annam	1816	1816	0	1885	1893	-8
Ethiopia	1816	1855	-39	2016	2016	0
Hanover	1816	1816	0	1866	1871	-5
Hesse	1816	1816	0	1866	1871	-5
Electoral						
Hesse Grand	1816	1816	0	1867	1871	-4
Ducal						
Korea	1816	1816	0	1905	1910	-5
Libya	1816	1816	0	1835	1834	1
Madagascar	1816	1816	0	1895	1896	-1
Mecklenburg	1816	1816	0	1866	1871	-5
Schwerin						
Modena	1816	1816	0	1860	1861	-1
Morocco	1816	1816	0	1912	1904	8
Parma	1816	1816	0	1860	1861	-1
Saxony	1816	1816	0	1866	1871	-5
Tuscany	1816	1816	0	1860	1861	-1
Two Sicilies	1816	1816	0	1860	1861	-1
Wuerttemberg	1816	1816	0	1870	1871	-1
Colombia	1819	1830	-11	2016	2016	0
Peru	1821	1824	-3	2016	2016	0
United	1823	1823	0	1840	1839	1
Provinces of						
Central						
America						
Uruguay	1828	1830	-2	2016	2016	0
Venezuela	1830	1829	1	2016	2016	0
Egypt	1833	1827	6	2016	2016	0
Transvaal	1852	1852	0	1877	1910	-33
Orange Free	1854	1854	0	1902	1910	-8
State						
Montenegro	1878	1868	10	1918	1915	3
South Africa	1885	1910	-25	2016	2016	0
Luxembourg	1890	1867	23	2016	2016	0
Nejd (Saudi	1902	1932	-30	2016	2016	0
Arabia)						
Bulgaria	1908	1878	30	2016	2016	0
Tibet	1912	1913	-1	1951	1950	1
Australia	1920	1901	19	2016	2016	0
Canada	1920	1867	53	2016	2016	0
New	1920	1907	13	2016	2016	0
Zealand						
Ireland	1922	1921	1	2016	2016	0
Lebanon	1946	1944	2	2016	2016	0

Table 5: ISD G-W comparison (*continued*)

Country	ISD Start (min)	GW Start (min)	Start Difference	ISD End (max)	GW End (max)	End Difference
Indonesia	1949	1945	4	2016	2016	0
South Korea	1949	1948	1	2016	2016	0
Laos	1953	1954	-1	2016	2016	0
German Democratic Republic	1954	1949	5	1990	1990	0
Nauru	1968	1979	-11	2016	2016	0
Bhutan	1971	1949	22	2016	2016	0
Bangladesh	1972	1971	1	2016	2016	0
San Marino	1992	1816	176	2016	2016	0
Andorra	1993	1816	177	2016	2016	0
Macedonia	1993	1991	2	2016	2016	0
Monaco	1993	1816	177	2016	2016	0

Figure A4 shows the net difference in the total amount of states between ISD and G-W, a positive number indicates that there are more ISD states while a negative number indicates that there are more G-W states in a given year. Here, a few key differences are evident. Firstly, the ISD includes many more states in the pre 1900 period compared to G-W. However, in the post-1900 period, the ISD becomes slightly more restrictive than G-W, with fewer relative states registered per year albeit at a much lower net difference than in the 19th century.

Figure A4. ISD and G-W Net Difference (1816 to 2016)



These differences between these datasets largely stem from two factors. First, ISD places a strong emphasis on undisputed or acknowledged sovereignty, with sovereignty encompassing both internal sovereignty (a monopoly of violence in a specified territory) and external sovereignty (control over foreign policy) (Butcher and Griffiths 2020). This recognition or undisputed status applies to “relevant international actors,” which takes a regional approach. There is a temporal dimension to this criterion as “relevant international actors” become increasingly globalized over time. Second, G-W implements a “ten year rule” in which states are considered to retain their status if their sovereignty is restored within ten years e.g. France during World War II. This explains why there is a drop in the number of states during WWII in ISD but not in G-W.

In summary, using the ISD state system greatly increases the coverage of countries in the 19th century allowing greater coverage for specifically non-Western state units. However, ISD’s stricter definition of sovereignty, with an emphasis on external recognition rather than internal control, leads to a slight reduction in the number of country-year units in the post-1900 international system compared to G-W. In short, while ISD is generally more permissive and expansive during the 19th century, with a high focus on the regionality of international systems, the dataset becomes more conservative in the 20th century. For the purposes of extending and reclassifying conflicts 1816 to 1945, many of which occur during the long 19th century, the pros of using the ISD as a country unit basis outweigh the cons.

Non-state Wars Excluded from HCD

Table 6: Intrastate Wars in Wimmer and Min (2009) Excluded in HCD

W-M War Name	Wimmer and Min (2009)	Countries	Years
Boers vs. Matabele	Intra	Boer state	1836-1837
Gilgit, India	Intra	Principality of Jammu	1852-1852
Xhosa Civil War	Intra	Xhosa	1818-1818

Table 7: Non-state Wars in CoW Excluded in HCD

CoW War Name	Side A	Side B	Start	End
Boer-Matabele War	Boer Trekkers	Matabele	1836	1837
First Maori Tribal War	Te Rauparaha's Ngati Toa	Taranaki, Ngai Tahu, Waikato, Ngati Ira, Rangitikei	1818	1824
Dogra-Tibet War	Jammu	Tibet	1841	1842
First Australian Aboriginal War	Aborigines	White Settlers	1864	1865
Second Zulu Internecine War	Cetewayo faction	Zibelu faction	1883	1884
Second Australian Aboriginal War	Aborigines	White Settlers	1884	1894
German East Africa Company War	German East Africa Company	Arabs, Swahili	1888	1889

References

- Brecke, Peter. 1999. "Violent Conflicts 1400 A.D. To the Present in Different Regions of the World." In Ann Arbor, Michigan. <https://cpb-us-w2.wpmucdn.com/sites.gatech.edu/dist/1/19/files/2018/09/Brecke-PSS-1999-paper-Violent-Conflicts-1400-AD-to-the-Present.pdf>.
- Butcher, Charles R., and Ryan D. Griffiths. 2020. "States and Their International Relations Since 1816: Introducing Version 2 of the International System(s) Dataset (ISD)." *International Interactions* 46 (2): 291–308. <https://doi.org/10.1080/03050629.2020.1707199>.
- Clodfelter, Micheal. 2017. *Warfare and Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures, 1492-2015*. Fourth edition. Jefferson, North Carolina: McFarland & Company, Inc., Publishers.
- Dixon, Jeffrey S., and Meredith Reid Sarkees. 2016. *A Guide to Intra-State Wars: An Examination of Civil, Regional, and Intercommunal Wars, 1816-2014*. 2455 Teller Road, Thousand Oaks California 91320: CQ Press, SAGE Publications, Inc. <https://doi.org/10.4135/9781452234205>.
- Gleditsch, Kristian. 2004. "A Revised List of Wars Between and Within Independent States, 1816-2002." *International Interactions* 30 (3): 231–62. <https://doi.org/10.1080/03050620490492150>.
- Gleditsch, Kristian Skrede, Nils W Metternich, and Andrea Ruggeri. 2014. "Data and Progress in Peace and Conflict Research." *Journal of Peace Research* 51 (2): 301–14. <https://doi.org/10.1177/0022343313496803>.
- Gleditsch, Kristian Skrede, and Michael D Ward. 2013. "System Membership Case Description List," May. http://ksgleditsch.com/data/iisyst_casedesc.pdf.
- Gleditsch, Kristian S., and Michael D. Ward. 1999. "A Revised List of Independent States Since the Congress of Vienna." *International Interactions* 25 (4): 393–413. <https://doi.org/10.1080/03050629908434958>.
- Lyall, Jason. 2020. *Divided Armies: Inequality and Battlefield Performance in Modern War*. Princeton Studies in International History and Politics. Princeton: Princeton University Press.
- Miller, Charles, and K. Shuvo Bakar. 2022. "Conflict Events Worldwide Since 1468BC: Introducing the Historical Conflict Event Dataset." *Journal of Conflict Resolution*, August, 002200272211190. <https://doi.org/10.1177/00220027221119085>.
- Pettersson, Therese. 2022. "UCDP/PRIO Armed Conflict Dataset Codebook v 22.1." <https://ucdp.uu.se/downloads/>.
- Phillips, Charles, and Alan Axelrod. 2005. *Encyclopedia of Wars*. Facts on File Library of World History. New York: Facts on File, Inc.
- Sarkees, Meredith Reid, and Phil Schafer. 2000. "The Correlates of War Data on War: An Update to 1997." *Conflict Management and Peace Science* 18 (1): 123–44. <http://www.jstor.org/stable/26273568>.
- Sarkees, Meredith Reid, and Frank Whelon Wayman. 2010. *Resort to War: A Data Guide to Inter-State, Extra-State, Intra-State, and Non-State Wars, 1816-2007*. Correlates of War Series. Washington, D.C: CQ Press.
- Wimmer, Andreas, and Brian Min. 2009. "The Location and Purpose of Wars Around the World: A New Global Dataset, 1816-2001." *International Interactions* 35 (4): 390–417. <https://doi.org/10.1080/03050620903328837>.