Routledge
Taylor & Francis Group

International Interactions

Empirical and Theoretical Research in International Relations

ISSN: 0305-0629 (Print) 1547-7444 (Online) Journal homepage: https://www.tandfonline.com/loi/gini20

States and their international relations since 1816: introducing version 2 of the International System(s) Dataset (ISD)

Charles R. Butcher & Ryan D. Griffiths

To cite this article: Charles R. Butcher & 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, DOI: 10.1080/03050629.2020.1707199

To link to this article: https://doi.org/10.1080/03050629.2020.1707199

INTERNATIONAL INTERACTIONS
2020, VOL. 46, NO. 2, 291–308
https://doi.org/10.1080/03050629.2020.1707199

Routledge
Taylor & Francis Group

SPECIAL DATA FEATURE

States and their international relations since 1816: introducing version 2 of the International System(s) Dataset (ISD)

Charles R. Butcher[a] and Ryan D. Griffiths[b]

[a]The Norwegian University of Science and Technology; [b]Syracuse University

ABSTRACT

We introduce version 2 of the International System(s) Dataset (ISD), a register of sovereign states across the 1816–2016 period that include numerous states that are missed in commonly used datasets like the Correlates of War (COW) Project. Whereas ISD version 1 identified 363 states between 1816 and 2011, version 2 identifies 482. This version also records valuable information on a range of corollary variables, including start dates, end dates, estimated population sizes, diplomatic relations with Europe, conflict episodes, the existence of borders, and the location of capital cities. This dataset makes an important contribution to the study of international relations. It provides a more accurate understanding of the development of the international system over the last two centuries, it moves beyond the Eurocentric bias that sits at the heart of existing quantitative IR scholarship, and it will enable scholars to pursue a range of research topics such as the historical importance of state borders and boundaries, the practices surrounding recognition, and the frequency and intensity of conflict across regions. In this article, we discuss the existing state system membership lists and show how the ISD addresses their shortcomings. We outline the key concept and operationalization of statehood that the ISD adopts. We detail the variables included in this version of the ISD, discuss the data collection process, and show temporal and spatial distributions that illustrate the uniqueness of the ISD. Finally, we demonstrate the utility of bringing the ISD into one of many potential research topics: the study of conflict.

KEYWORDS

Conflict; data; international system; states

In this article we introduce version 2 of the International System(s) Dataset (ISD).[1] The ISD is a register of sovereign states across the 1816–2016 period that include numerous states that are missed in commonly used datasets like the Correlates of War (COW) State System Membership List (Correlates of War Project 2017), particularly in the pre-1920 period. While version 1 of the ISD identified 363 states between 1816 and 2011, version 2 identifies 482. Most of these new states existed in the 19th century. For example, whereas

CONTACT Ryan D. Griffiths rgriff01@syr.edu Political Science, Syracuse University, Syracuse, NY

This article has been republished with minor changes. These changes do not impact the academic content of the article.

[1]For version 1, see Griffiths and Butcher (2013).

Supplemental data for this article can be accessed on the publisher's website.

version 1 of the ISD identified 134 states in 1816, version two identifies 211. The new version of the ISD also records valuable information on these states, including start dates, end dates, estimated population sizes, diplomatic relations with Europe, conflict episodes, and the existence of borders.

This dataset makes an important contribution to the study of international relations (IR). First, it provides a more accurate understanding of the development of the international system(s) over the last two centuries.[2] Second, it moves beyond the Eurocentric bias that sits at the heart of existing quantitative IR scholarship, best signified by COW, and enables a more systematic investigation into pre-colonial, non-western state systems that have largely been neglected (Acharya 2014; Butcher and Griffiths 2017). With these data, scholars can pursue a range of research topics such as the historical importance of state borders and boundaries, the practices surrounding recognition, and the frequency and intensity of conflict across regions.

The remainder of the article proceeds as follows. First, we discuss the existing state system membership lists and show how the ISD addresses their shortcomings. Second, we outline the key concept and operationalization of statehood that the ISD adopts. Third, we detail the variables included in this version of the ISD, discuss the data collection process, and show temporal and spatial distributions that illustrate the uniqueness of the ISD in relation to other state registers. In the final section we demonstrate the utility of bringing the ISD into one of many potential research topics: the study of conflict.

Existing Datasets

The original purpose for the ISD was to correct a bias in existing IR datasets that excluded states in Africa, South Asia, Central Asia, and South-East Asia in the pre-colonial period. For example, Figure 1 shows the number of states recorded in the COW state system membership list for the pre-League of Nations period (1816–1919) and their approximate locations in relation to contemporary states. One would get the impression from this figure that the African continent was stateless, as was South Asia, Southeast Asia and the Pacific. Yet, historical sources point to numerous state-like entities in these regions (Tambiah 1977; Wilks 1975; Herbst 2000; Oliver and Atmore 2001; Kang 2010; Ringmar 2012; Fenske 2013; Besley and Reynal-Querol 2014; Phillips and Sharman 2015; Wig 2016).

The regions appear empty because of the criteria used by COW to identify states in the pre-1920 era. For an entity to qualify as a state in the COW

[2]Following the example set in the ISD V1, we refer to international system(s) in a way that implies singular/plural ambiguity in the number of systems over time. Most readers will accept that in 2019 there is only one international system. Many will accept that in 1816 there were a number of loosely connected, even disconnected systems. That plurality gradually consolidated into a singularity over the last 200 years.

Figure 1. World Map of States in the COW Data, 1816–1919.

register during this period it needed to satisfy both of the following conditions: (1) Possess a population of at least 500,000 people; and (2) Receive diplomatic representation by France and Britain at the level of *charge d'affaires* or higher (Correlates of War Project 2017). The purpose behind these criteria was to identify viable states that were internationally recognized; population size and diplomatic relations with the British and the French were used as proxies. Notably, the criteria for inclusion in the COW list changed from 1920 onward, where states were now counted if they had membership in the League of Nations or the United Nations (UN), or a population of 500,000 or more and recognition by two or more major powers. As other scholars have noted, these criteria had two important effects (Bremer and Ghosn 2003; Butcher and Griffiths 2015; Butcher and Griffiths 2017; Fazal 2007; Gleditsch and Ward 1999; Griffiths and Butcher 2013). First, it generated a pre- and post-1920 inconsistency in population size given that small UN-member states like Tuvalu and Nauru with populations just over 10,000 make the COW list, and any state in the pre-1920 period with a population below 500,000 was excluded. Second, it tied international recognition to Britain and France because there were no globe-spanning associations of sovereign states like the League of Nations or the UN prior to 1920, and, as the originators of COW wrote, Britain and France were the two key "legitimizers." (Singer and Small 1966, 246).

The criteria used in dataset construction can have important downstream consequences. The architects of COW pinned international recognition in the pre-1920 era to Britain and France because those states maintained thorough records of diplomatic contact, and because they were arguably the core of the expanding European-based international system. But the international system(s) of the 19th century was not fully connected, and is

better envisaged as a set of regional systems that were gradually drawing together. Neither Britain nor France had sufficient presence in all regions of the world to recognize states at the required level. For some states that operated independently in a sovereign fashion, France and Britain simply were not the relevant regional powers. For the Luba Kingdom in Central Africa, the locally relevant states were Lunda and the Belgians, not Britain or France. In Bali, the relevant powers were the competing local kingdoms and the Dutch. At a basic level, Britain and France were simply more likely to recognize states that were geographically closer to them, which is perhaps one of the reasons for why China and Japan did not join the European-based system, according to COW, until 1860. Importantly, diplomatic recognition is a strategic act (Teorell 2017), and by conditioning statehood on the diplomatic practices of France and Britain, the biases and selection processes of French and British diplomats were encoded into the DNA of COW. As a thought experiment, consider that the COW list and the family of COW datasets keyed to it have formed the backbone of quantitative research in IR. Would that research have derived substantially different conclusions on a range of topics if the designers had pinned diplomatic relations not to Britain and France, but to China and the Ottoman Porte?

We are not the first to identify or attempt to redress these issues. Kristian Gleditsch and Ward (1999)(GW) offered a corrective that reduced the size threshold and relaxed the recognition criterion. For them, a state counted if it had a population of at least 250,000 and was regarded by local actors as a distinct entity. However, there is no requirement that the political unit be externally sovereign, and, as a result, entities were included even if their foreign policy was formally controlled by another, as Oman was from 1891 to 1971. In addition, GW missed a substantial number of states that were picked up in the ISD. In her work on state death, Fazal (2007) added 16 new states to the COW member list in the pre-1920 period, and modified the start date of another 22, by including states that had signed treaties of commerce, alliance, or navigation with either Britain or France, even if they did not receive diplomatic missions. In doing so, she demonstrated how quickly state membership expands by lowering the level of required diplomatic relations. Finally, the more recent Historical V-DEM dataset records the features of 80 states from as early as 1789 (Coppedge et al. 2019). In relation to the ISD, however, the Historical VDEM data identifies far fewer states due to the requirement that states must have a continuous existence into the 20th century in order to be included. In Africa, for example, most states disappeared with colonialism by the end of the 19th century and are therefore not captured in the Historical V-DEM. In sum, all of these datasets have made valuable contributions to the scholarship. The COW project helped kick start the quantitative work in IR. Nevertheless, they present an incomplete historical picture of the international system(s).

The ISD

The purpose behind the ISD was to provide a more accurate list of states and a more comprehensive understanding of the development of the international system(s) since the early 19th century. To qualify as a state in the ISD, a political entity needs to have:

(1) A population of at least 10,000.[3]
(2) Autonomy over a specific territory.
(3) Sovereignty that is either uncontested or acknowledged by the relevant international actors.

These criteria reflect the ISD's revised conception of statehood. Sovereignty has an internal and external dimension. Internally, states stand at the top of a hierarchy and are force-wielding organizations. In the ideal form, states possess a complete monopoly on the use of physical force within their borders (Weber 1946, 78), but some latitude is given since few if any states possess this in full. Externally, a sovereign state has recognized, formal control over its foreign relations. By using criteria that stresses the importance of internal and external control, we are taking the same general tack as COW, and are consistent with the IR scholarship on sovereignty and the study of the international system(s) (Branch 2014; Butcher and Griffiths 2017; Buzan and Little 2000; Donnelly 2012; Fazal 2007; Krasner 1999; Lake 2009; McConaughey, Musgrave, and Nexon 2018; Ruggie 1998; Spruyt 1994).

The first criterion regarding population size has a rather practical purpose. A population floor is a scope condition for the unit of analysis, and its inclusion helps with data collection by filtering small entities about which it can be hard to gather information, one of the stated reasons behind the 500,000-person minimum in COW (Bremer and Ghosn 2003, 24). Moreover, there is a sense in which smallish entities gradually appear less viable as states. Although any population threshold is, in a sense, arbitrary, we chose 10,000 because it is focal, because it was not too small to hinder our data-collection efforts, and because it provided consistency across the entire 1816–2016 period given that many UN Member states have populations that barely exceed(ed) 10,000. As we discuss below, moving the population threshold from 100,000 to 10,000 is one of the main differences between ISD version 1 and version 2.

The second criterion is about internal control; to count, a state must have autonomy over a specific territory. In part, this was meant to filter virtual states and governments in exile. But more generally, it implies that the

[3]In ISD Version 1 the population floor was 100,000.

government of the state is super-ordinate to all other entities and government structures within the stated territory. It need not have control over all areas of governance, and, in practice, there is significant variation within states in terms of how decision-making powers are distributed (Nexon 2009; McConaughey Musgrave, and Nexon 2018). For example, the Oyo Empire in 19th Century West Africa controlled little more than foreign policy, leadership succession, and judicial authority over capital crimes in relation to its constituent polities, while they retained wide-ranging autonomy over decision-making in relation to taxation methods, policing, and local law-making (Wilks 1975). These are issues of what Barry Buzan and Richard Little referred to as structural differentiation: the degree to which sovereigns delegate key prerogatives to sub-state actors and entities (Buzan and Little 2000, 87). A key element that defines a sovereign, and that which differentiates it from subordinate polities, is its ability formally manage its foreign relations (Butcher and Griffiths 2017). As such, the second criterion is connected theoretically to the third.

The third criterion is meant to pick up external sovereignty that is either uncontested or acknowledged by the relevant international actors. Here, our purpose was to avoid the Eurocentric bias in COW and develop the concept of sufficient recognition (Griffiths and Butcher 2013, 757). For a period until the late 1800s, Japan existed in a relatively hermetic condition with few diplomatic linkages, but its sovereign independence was uncontested. Others, like Nicaragua, engaged in diplomatic relations and were regarded by their neighbors as sovereign, but just not with both Britain and France. By taking a regionally-sensitive approach, we can ascertain whether sufficient recognition was achieved by the polity in question. Accordingly, we define the relevant international actors as those states which, for all practical purposes, are the key legitimizers of the state in question. The general disconnectedness and gradual development of the international system(s) prior to 1920 renders it problematic to make any one state the key legitimizer. While the key actors can vary depending on context, we argue (and have found) that they are always identifiable.

Formal control over foreign relations is a key element for units in our dataset. It stresses the ability of states to interact with other states and be recognized by the relevant external actors as a state participating independently in an international system. We require that no other state has the formal right to determine the polity's foreign affairs. If we found evidence that neighboring states or major powers in the region recognized a state as independent, then we treated it as independent. Of course, this can be ambiguous in cases where other states gradually take control of another state's foreign policy. In such cases, we recorded a state exit (death) when the larger state and other relevant actors began to treat the state in question as non-sovereign (Fazal 2007). Overall, we conceive of states in the dataset as

"sovereign peaks" in political landscapes that look like contour maps. Provinces, federacies, and protectorates may also be peaks, but they are local maxima when they surrender control over their foreign affairs to a higher authority.

These criteria constitute the conceptual foundation of the ISD, which was originally elaborated in the 2013 article introducing the ISD (Griffiths and Butcher 2013). The purpose behind version 2 was to improve the data in three ways. First, external funding enabled us to form a large research team that could, as we detail below, more systematically and rigorously explore the data. Second, we lowered the population threshold from 100,000 to 10,000, because we felt that would create a more consistent dataset over time, and because we knew that with our added resources we could research the smaller states that were excluded in version 1. Third, we created corollary data on a host of variables like conflict, borders, and population size that are keyed to the ISD state list, a feat that raises the value and utility of the ISD.

Data Collection for ISD Version 2

Our data collection efforts were conducted over a three-year period from 2016 to 2019. The team consisted of two Principal Investigators (PIs) and 12 Research Assistants (RAs) who were doing doctoral work in one of the social sciences. We began by consulting a variety of secondary historical sources to identify states and code their features over the 1816–2016 period. Key sources for the first sweep of cases included the Cambridge History Series, the UNESCO History of Africa Series, the General History of Africa series, and the Correlates of War Colonial Entities List. Once a potential case was identified, RAs consulted secondary sources relating to those cases to establish whether the case met our criteria for statehood. Coding these first cases often led to the discovery of new states and the process snowballed until we concluded that we had identified all states meeting our criteria.

Once a state was added to our list, a number of additional variables were coded. Table 1 provides a partial list of the variables used (for the complete list see the online Codebook and Dataset).[4] Note that these additional variables are coded only for states in the 1816–1920 period and primarily for states that are not captured in the COW List. Coders were assigned cases based on regions, which led to the development of expertise and familiarity with sources related to the region. All of the states in the ISD were independently coded by at least two coders. Each state included in the dataset has a separate case narrative that provides a description of the state and outlines the sources used to reach coding decisions. The final dataset was produced

[4]Table 1 is truncated for presentational reasons. Some variables were excluded here because they may be less important for the reader at this stage (e.g. Was their ambiguity on the end date?).

Table 1. Sample variables from ISD 2.

Variable	Explanation
Start Date	Date at which the entity gains external sovereignty and enters the dataset
Early start date	If the state already exists in 1816, approximate earlier start date
Start Type	1. Territorial Consolidation = created from stateless space. 2. Secession = seceded from larger state. 3. Dissolution = fragment of a larger state that ceased to exist. 4. Unification = voluntary merging of states. 5. Decolonization = state born from decolonization 6. Other
Start Settlement	Was there a treaty or formal settlement attending the birth of the state?
Region	COW index code for region in which state is located.
Declare	Was there a declaration of independence?
End Date	Date at which the entity exited the dataset
End Type	1. Colonization = Non-contiguous annexation. 2. Conquest / annexation = State death via foreign take-over. 3. Unification = Voluntary unification. 4. Dissolution = State death via dissolution of state. 5. Partition = State death via partition by outside powers. 6. Other
End Settlement	Was there a treaty or formal settlement attending the death of the state?
Borders	Did the state have clearly defined borders separating it from other states or sources of authority, as opposed to undefined border regions characterized by shifting, uncertain, and overlapping authority claims?
Europe Diplomacy	Did the State have diplomatic relations with a European power?
Geolocation	Latitude and Longitude coordinates for capital city or center of the state
Violent Start	Was the state born violently?
Violent End	Did the state end violently?
Violent Other	All recorded conflicts for state

after we, the PIs, read each case narrative and coding, and reconciled this information with the coding criteria to ensure the consistency of the data in relation to the variable operationalization across coders, regions, and time periods. Each variable has an ambiguity score of 0 or 1 that reflects whether a coder thought that a data point was highly uncertain due to conflicting sources or scarce information.

Version 2 of the ISD contains information on the variables in Table 1 for 282 states over the 1816–1920 period. Importantly, the core variables are those that focus on state membership in the system(s). Did the polity qualify as a state given our terms? When and how did it begin? When and how did it end? Other variables regarding borders, declarations, conflict, etc., are more exploratory. As we discuss in the online Codebook, our team recorded all instances of the phenomena in relation to the states. Some of these data will require further systemization. At the moment, they constitute a wealth of information that be used for multiple research purposes, all keyed to a rigorously created set of sovereign states. As an example, Figure 2 shows approximate locations of the central regions/capital cities of these states.

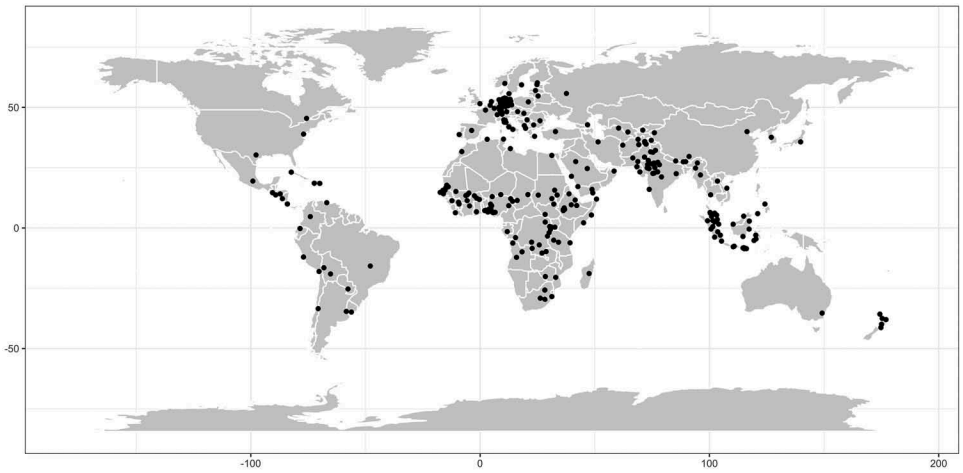

Figure 2. Approximate Locations of Capitals/Centers of ISD States, 1816–1920.

The final dataset lists 482 states over the 1816–2016 period. This represents a substantial increase over COW and other datasets, which identify roughly half that number. Most of the new cases were located in Africa, South Asia, and South East Asia during the 1816–1920 period. For example, COW identifies just six states on the African continent prior to decolonization: Ethiopia (1898–1936, from 1941), South Africa (from 1920), Morocco (1847–1912), Egypt (1855–1882, from 1937), Liberia (from 1920), and Tunisia (1825–1881). In contrast, Version 2 of the ISD identifies 95 states during this time, which is also a substantial increase in relation to projects that have specifically aimed to identify states in Africa.[5] Likewise, whereas COW identified only two states in South Asia prior to decolonization: Afghanistan (from 1919) and Nepal (from 1920), we identified 48. Figure 3 shows the number of ISD-identified states that existed within what are now the boundaries of a contemporary sovereign state. Standouts include Nigeria, Indonesia, and, of course, India. These states were once the locations of regional systems.

There are limitations to the data. Although we have comprehensively surveyed these regions, there may be other states that we have missed. Some cases were truly marginal. For example, although Swat was included in version 1 of the ISD, we removed it because we struggled to find a state-like entity that possessed clear control over the mountainous region. A different problem existed in more densely populated regions of complex, nested political control like South Asia and in pre-unification Germany where it was difficult to determine whether a polity had formal control over its foreign relations. Our research team came to refer to these problems as, respectively, the low-density and the high-density problems. In addition,

[5]For example, Besley and Reynal Querol (2014) identified 19.

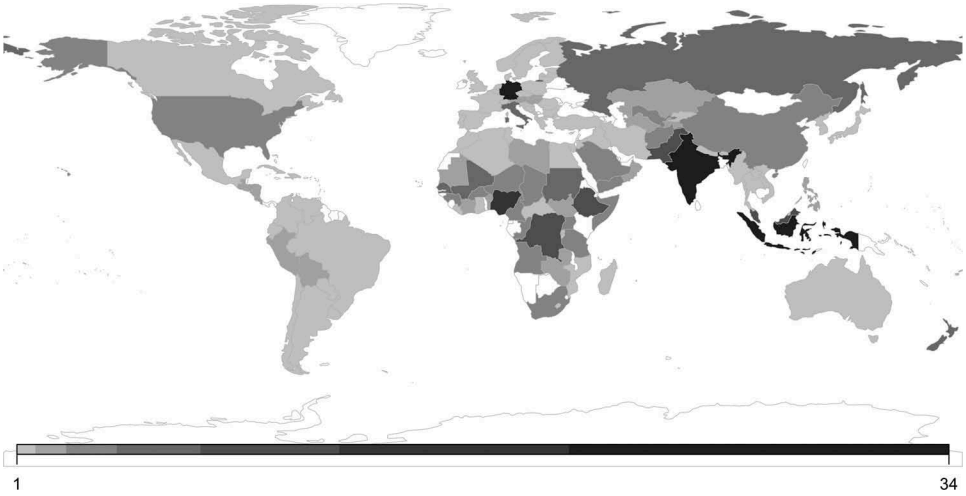

Figure 3. Number of ISD States, 1816–1920, within Contemporary States.

many of the state-level variables are constant across state episodes and cannot be used to examine changes over time. Information on many of these historical states is scarce, and while we have consulted a wide range of secondary historical and anthropological sources to collect the data, the sparsity of quality information led to caution on our part. As mentioned, each variable has a binary ambiguity indicator to signal whether the coders felt that the value for a particular variable was uncertain because of a lack of source material or conflicting accounts. Finally, some of the more exploratory variables (all noted in the Codebook) are ready for qualitative and small-N research, but will require more systemization before they can be utilized for advances quantitative work.

We argue that version 2 of the ISD provides the most comprehensive portrait of the development of the international system(s) since 1816. As Figure 4 illustrates, the ISD shows a very different picture of how the number of states around the world has changed over time, compared to the COW and GW lists. The three datasets are very similar from 1920 onwards and vary only slightly in relation to different rules around the inclusion of microstates and how periods of occupation are dealt with. However, the picture is altogether different in the pre-1920 era. Whereas COW identifies 23 states in 1816 and GW list 47, the ISD discovered 211. When taking the entire pre-1920 period as a whole, the ISD identifies 315 states, roughly 14 times the number of states in COW and seven times the number in GW. Most of these states existed in the arc of territory stretching from Africa up through Central and South Asia and down to Southeast Asia. This is a vastly different picture of the international system(s), one that shows a concave trend to the number of states over time, rather than one in which the number has been gradually

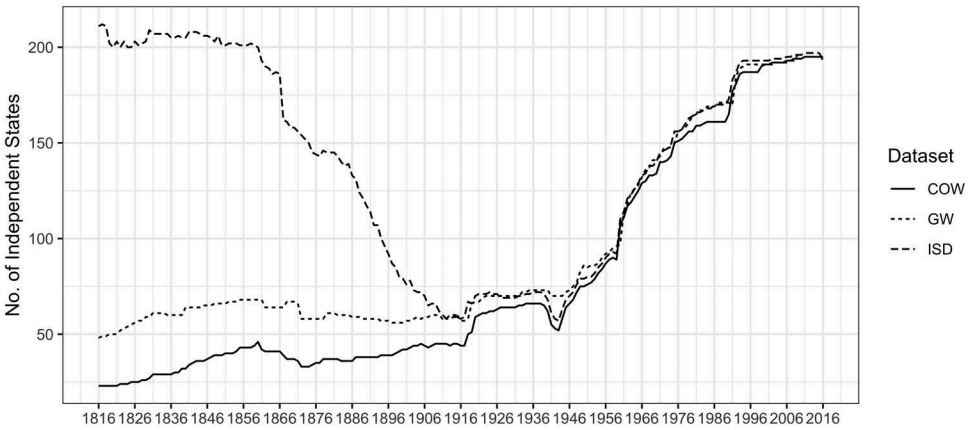

Figure 4. Independent States According to ISD, COW, and GW, 1816–2016.

increasing. To some extent, the distance between the two trends in the early 1800s, particularly where COW is concerned, is a legacy of the Eurocentric bias in IR research.

The ISD reveals a myriad of interesting dynamics. For example, Figure 5 shows patterns of state birth and death. Fazal (2007) observed that state death became uncommon after 1945, a pattern confirmed by the ISD. The ISD shows that state deaths were very common in the 19th century, especially with the acceleration of colonialism through Africa and South Asia. However, state birth was also fairly common at the beginning of the 19th century. As Figure 6 shows, state birth in Africa and the Americas outstripped state death in the early 19th century. Africa was especially dynamic in this regard with the rise of numerous states such like Sokoto, Tokolor, the Wassulu Empire, the Gaza Empire, and the Zulu Kingdom at the same time that the British and French were expanding across the continent. This

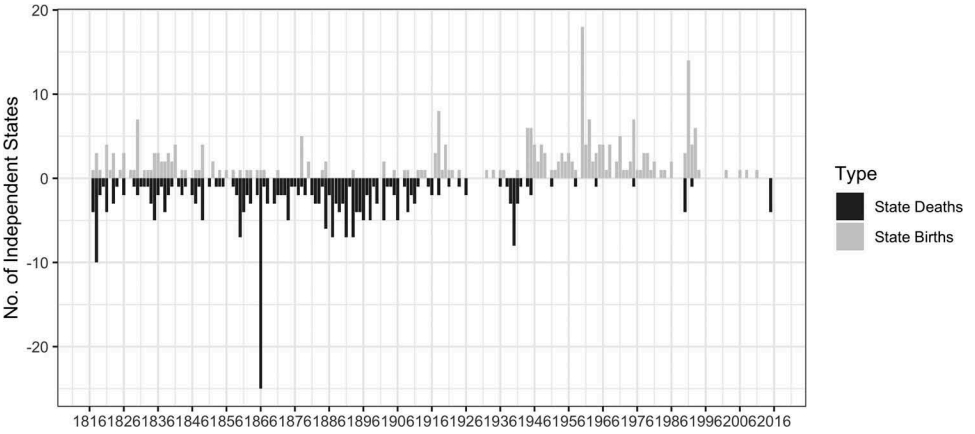

Figure 5. State Births and Deaths in the ISD Data, 1817–2016.

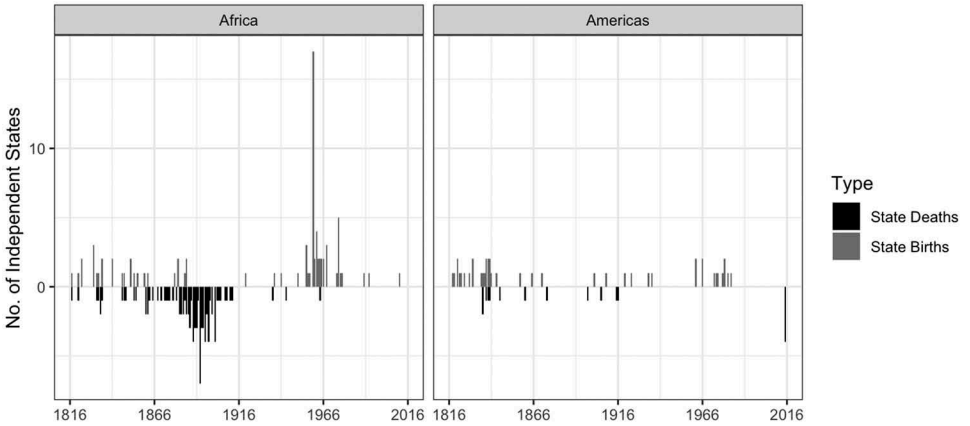

Figure 6. State Births and Deaths in Africa and Americas, 1817–2016.

empirical pattern points to interesting dynamics of state creation and destruction prior to the enclosure of the international system(s) through colonial acquisition. Overall, secession was the most common start type, followed by territorial consolidation, dissolution, unification and decolonization. Colonization was the most common form of state death, followed by conquest (by a local territorial power), unification, and dissolution.

There are stark differences in the processes of state birth and state death across regions during the 1816–1920 period. In Africa, state creation was nearly evenly divided between territorial consolidation and secession. Thus, most states were created by either absorbing territory that contained smaller political entities such as ethnic groups or tribes (or in some cases establishing new entities on territory that was sparsely populated) or by breaking off from larger states. Put differently, state creation in Africa exhibited a relatively even balance between consolidation and secession. In terms of state death in Africa, most states were extinguished by colonization (65%) but a sizable portion died through conquest by local states (31%). There are no cases of voluntary unification in Africa across the 1816–1920 period. Secession was a significantly more common start type across all regions and territorial consolidation was comparatively rare. Although dissolution and unification are rare generally, they are more common forms of state birth in East and South-East Asia and the Americas. Colonization was by far the leading type of state death in South Asia (80%) and South East Asia (79%). Meanwhile, unification was the most common form of state death in Europe (largely as a result of German and Italian unification).

Utility in Conflict Research

In this section, we provide a more focused example of the utility of the ISD in a specific research area: conflict studies. We contend that the ISD is useful for

a range of purposes, such as: (1) Opening possibilities for further, more comprehensive data collection on historical conflict; (2) Contributing to the studies that probe links between pre- and post-colonial institutions and expanding them beyond Africa; (3) Expanding the analysis of historical statehood and contemporary conflict beyond the assumed link between ethnic groups and historical statehood; (4) Adding to the debate on the decline of war.

First, the ISD can contribute to conflict research simply by assisting with data collection efforts. Projects likes the Uppsala Conflict Data Program (UCDP) typically begin with a set of states, and then investigate the incidence of conflict in which those states are a party. Since the list of states guides the data collection efforts, many conflicts will be missed or else, as we see with the COW Project, given a different classification like "non-state", "intra-state", or "extra-state". The more expansive register of states in the ISD can assist scholars in developing more accurate and comprehensive data on historical conflict.

Second, there is a growing literature linking pre-colonial patterns of statehood with post- colonial armed conflict, especially civil war (Besley and Reynal-Querol 2014; Englebert, Tarango, and Carter 2002; Paine 2019; Wig 2016). While some studies find that state-like structures from the pre-colonial past have facilitated bargaining between post-colonial states and ethnic groups (Wig 2016), thereby reducing the chance of war, others find that pre-colonial state structures can increase the probability of armed conflict (Besley and Reynal-Querol 2014; Paine 2019). However, these studies are conducted with limited data and focused almost exclusively on Africa. To date, no one has been able to mount a comparative and comprehensive analysis between regions. Expanding the empirical scope of quantitative studies of historical statehood and post-colonial conflict is important for clarifying whether existing findings, trained on Africa, also apply in other regions. The ISD makes this possible.

Third, the ISD would contribute to studies looking at the relationship between historical statehood, ethnicity, and conflict. Studies of pre-colonial statehood and post-colonial conflict often ask whether ethnic groups that survive into the modern period were previously organized into states. For example, Jack Paine coded pre-colonial statehood for ethnic groups only and his primary independent variable is a measure of ethnic groups that were ruled by a centralized, ethnic state. Tore Wig also coded levels of statehood for ethnic groups. The assumption here is that states in Africa were primarily or exclusively ethnic states. While ethnicity was certainly a common foundation upon which the core structures of states were built, this assumption is more problematic when we note that many multi-ethnic empires emerged in the 19th century in Africa. States such as the Sokoto Empire, Tokolor Empire, Wassulu Kingdom, Macina Empire, Oyo Empire, Ashanti

Kingdom, and the Yeke Kingdom, among others, were all states that ruled multiple ethnic groups and embarked upon ambitious state-building projects designed to incorporate conquered peoples into a centralized administration. It may be the case that post-colonial ethnic groups are often the product of successful pre-colonial state-building enterprises. This assumption becomes even more problematic if we look outside of Africa. In India, for example, statehood may have been less tied to ethnicity than caste, as both the Rajputs and the Marathas demonstrate. The ISD data can facilitate a closer connection between the study of pre-colonial statehood and post-colonial conflict while holding separate the interrelated but distinct phenomena of statehood and ethnicity.

Fourth, the ISD can contribute to the debate on the decline of war. Whether the incidence of war has in fact declined over the past two centuries is a contested topic. While some claim that war and the severity of war has declined (Pinker 2011),[6] others suggest that, especially in relation to interstate war, the low levels of war witnessed since the end of World War II would not be unusual in a world where the baseline probability of war has not changed (Braumoeller 2019; Clauset 2018). Recent work has recognized that understanding whether the probability of war has declined also requires accounting for the number of units with the potential to initiate or participate in conflict. Observing more wars in the past could be a consequence of a higher probability of war, but it also could be explained by more states in the past, each with an unchanged likelihood of conflict. In the same way that it is inaccurate to judge the likelihood of contracting a disease among two differently sized communities by looking only at the number of people who get the disease without adjusting for how many people there are in each community, it is also inaccurate to infer changes in the baseline probability of war by looking only at the number of wars.We probed the potential contribution that the ISD Version 2 can make to this debate. Figures 7 and 8 show trendlines in conflict per year depending on the state register being used (COW, GW, and the ISD). The main difference between the two figures is the source of conflict data. Figure 7 utilizes the Brecke Conflict Catalog (Brecke 1999),[7] which includes wars involving states in the ISD that are not in COW or GW such as the Ethiopia-Shoa conflict of 1865, the Luza-Kazembe conflict of 1830, and the Sokoto-Bornu war of 1824. Brecke's criteria for identifying state actors emphasized effective sovereignty, and is notable for the use of military histories and non-English language sources (Brecke 1999, 17). Figure 8 utilizes COW war data, including Extra-State wars and Non-State wars (Sarkees 2010).[8] Extra-state wars include conflicts between states in the COW schema and entities that do not meet the COW criteria for

[6]For an interesting critique, see Fazal (2014).

[7]We include the number of active Inter-state, Intra-state, Non-state and Extra state wars in COW (Sarkees and Wayman 2010).

[8]This may bias the data toward conflicts involving states that meet the COW criteria.

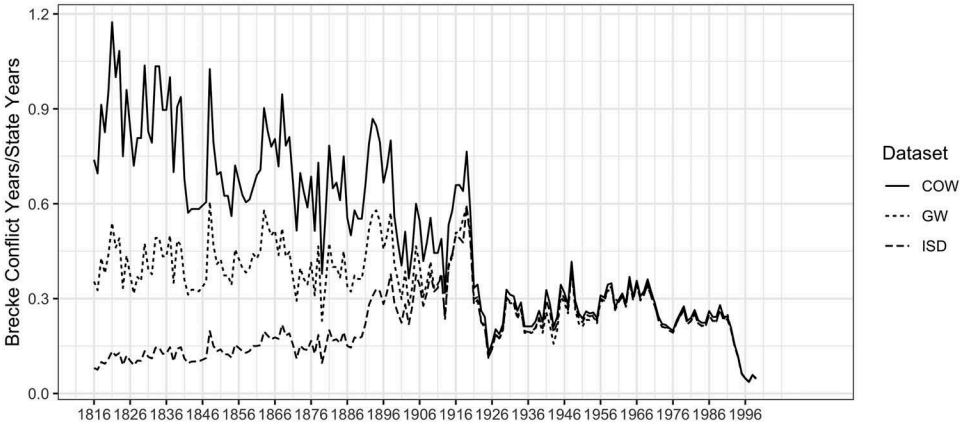

Figure 7. Brecke Armed Conflicts per State-Year, ISD, COW, and GW, 1816–1999.

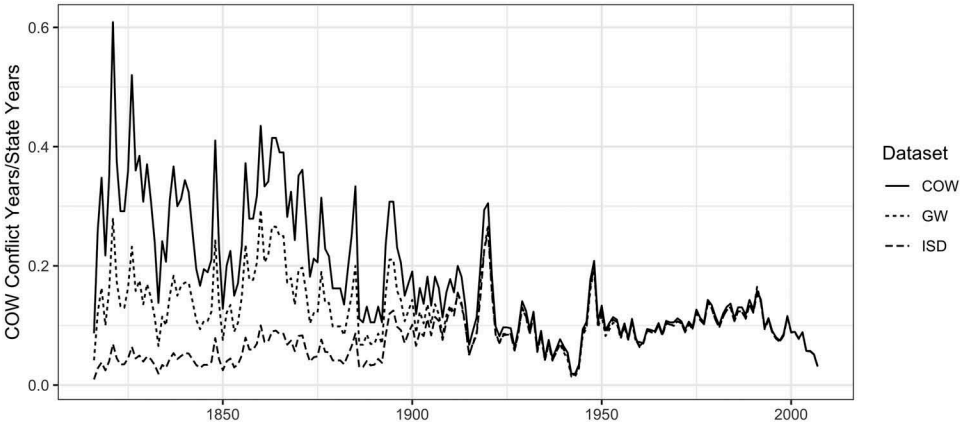

Figure 8. COW Armed Conflicts per State-Year, ISD, COW, and GW, 1816–1999.

statehood but might meet the ISD criteria. Non-state wars occur between entities that are not COW states but often meet the ISD criteria. For example, the French-Tukolor (1854–1857) conflict is classified as an extra-state war, and the Burma-Assam (1819–1822) conflict is coded as a non-state war. However, as we have argued elsewhere (Butcher and Griffiths 2015), under the ISD definition of statehood these are inter-state wars. Overall, we use these two conflict datasets to analyze how the number of active conflicts per state changes depending on state list that one uses.

Trends in the frequency of war vary dramatically depending on which state list is used to adjust for the number of entities at risk of conflict in a given year. As both Figures 7 and 8 indicate, using the COW conflict data would lead us to conclude that the frequency of war (wars per state-year) has dropped sharply since the 19th century. Adjustments using the GW list suggest that armed conflicts were more common before World

War II, after which the incidence fell. Adjusting for ISD states shows a more static picture in which the per-state war incidence was actually higher in the 1990s than it was in 1850. These patterns are broadly similar using the Brecke and the COW war data. To be clear, we are merely probing the data here, and we do not intend this to be the final word on whether war is declining. Comparisons across centuries will be hampered by a lack of reliable sources when compared to the vast amounts of newswire data that have been available to researchers in the past two decades (Weidmann 2016). Rather, our point is simply that the answers to this debate depend not only on the accuracy of conflict data, but also on the conceptualization and identification of states.

Conclusion

Datasets form the bedrock for quantitative research in IR. The ISD project began as an effort to contribute to that research by providing a more comprehensive and less Eurocentric set of states since 1816. In this second version of the dataset, the number of states is increased by lowering the population threshold from 100,000 to 10,000, and data is collected on a host of variables such as conflict, the existence of borders, and treaties. With these data, scholars can now go beyond a simple list of the states that have previously been excluded from quantitative work in IR and begin to study their international relations.

Acknowledgement

We thank the Australian Research Council (ARC) for funding this project with a Discovery Grant. We are grateful to Ben Goldsmith, Graeme Gill, and Seva Gunitsky for their advice and contributions. The data collection could not have been completed without the excellent work of the research team: Marigold Black, Keshab Girl, Jiye Kim, Aden Knaap, Haneol Lee, Christian Novak, Jeremy Simpson, Nik Skon, and Chris Watterson. We also thank the editors and reviewers at International Interactions for their feedback.

References

Acharya, Amitav. 2014. "Global International Relations (IR) and Regional Worlds: A New Agenda for International Studies." *International Studies Quarterly* 58 (4):647–59. doi: 10.1111/isqu.2014.58.issue-4

Besley, Timothy, and Marta Reynal-Querol. 2014. "The Legacy of Historical Conflict: Evidence from Africa." *American Political Science Review* 108 (2):319–36. doi: 10.1017/S0003055414000161

Branch, Jordan. 2014. *The Cartographic State: Maps, Territory, and the Origins of Sovereignty*. Cambridge: Cambridge University Press.

Braumoeller, Bear F. 2019. *Only the Dead: The Persistence of War in the Modern Age*. Oxford: Oxford University Press.

Brecke, Peter. 1999. "Violent Conflicts 1400 AD to the Present in Different Regions of the World, in 1999 Meeting of the Peace Science Society (International), October 8-10." Ann Arbor, Michigan, Unpublished Manuscript.

Bremer, Stuart A., and Faten Ghosn. 2003. "Defining States: Reconsiderations and Recommendations." *Conflict Management and Peace Science* 20 (1):21–41. doi: 10.1177/073889420302000102

Butcher, Charles R., and Ryan D. Griffiths. 2015. "Alternative International Systems? System Structure and Violent Conflict in 19th Century West Africa, Southeast Asia, and South Asia." *Review of International Studies* 41 (4):715–37. doi: 10.1017/S0260210514000436

Butcher, Charles R., and Ryan D. Griffiths. 2017. "Between Eurocentrism and Babel: A Framework for the Analysis of States, State Systems, and International Orders." *International Studies Quarterly* 61 (2):328–36. doi: 10.1093/isq/sqw057

Buzan, Barry, and Richard Little. 2000. *International Systems in World History: Remaking the Study of International Relations*. Oxford: Oxford University Press.

Clauset, Aaron. 2018. "Trends and Fluctuations in the Severity of Interstate Wars." *Science Advances* 4 (2):eaao3580. doi: 10.1126/sciadv.aao3580

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, et al. 2019. "V-Dem [Country-year/country-date] Dataset V9." *Varieties of Democracy (V-Dem) Project*. doi: 10.23696/vdemcy19

Correlates of War Project. (2017) "State System Membership List, V2016." Online http://correlatesofwar.org.

David, Singer, J., and Melvin Small. 1966. "The Composition and Status Ordering of the International System: 1815–1940." *World Politics* 18 (2):236–82. doi: 10.2307/2009697

Donnelly, Jack. 2012. "The Elements of the Structures of International Systems." *International Organization* 66 (4):609–43. doi: 10.1017/S0020818312000240

Englebert, Pierre, Stacy Tarango, and Matther Carter. 2002. "Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries." *Comparative Political Studies* 35 (10):1093–118. doi: 10.1177/001041402237944

Fazal, Tanisha M. 2007. *State Death: The Politics and Geography of Conquest, Occupation, and Annexation*. Princeton: Princeton University Press.

Fazal, Tanisha M. 2014. "Dead Wrong? Battle Deaths, Military Medicine, and Exaggerated Reports of War's Demise." *International Security* 39 (1):95–125. doi: 10.1162/ISEC_a_00166

Fenske, James. 2013. "Does Land Abundance Explain African Institutions?" *The Economic Journal* 123 (573):1363–90. doi: 10.1111/ecoj.12034

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. doi: 10.1080/03050629908434958

Griffiths, Ryan D., and Charles R. Butcher. 2013. "Introducing the International System(s) Dataset (ISD), 1816–2011." *International Interactions* 39 (5):748–68. doi: 10.1080/03050629.2013.834259

Herbst, Jeffrey. 2000. *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.

Kang, David C. 2010. *East Asia before the West: Five Centuries of Trade and Tribute*. New York: Columbia University Press.

Krasner, Stephen D. 1999. *Sovereignty: Organized Hypocrisy*. Princeton: Princeton University Press.

Lake, David. 2009. *Hierarchy in International Relations*. Ithaca, NY: Cornell University Press.

McConaughey, Mehgan, Paul Musgrave, and Daniel H. Nexon. 2018. "Beyond Anarchy: Logics of Political Organization, Hierarchy, and International Structure." *International Theory* 10 (2):181–218. doi: 10.1017/S1752971918000040

Nexon, Daniel H. 2009. *The Struggle for Power in Early Modern Europe*. Princeton: Princeton University Press. doi:10.1515/9781400830800.

Oliver, Roland, and Anthony Atmore. 2001. *Medieval Africa, 1250–1800*. Cambridge: Cambridge University Press.

Paine, Jack. 2019. "Ethnic Violence in Africa: Destructive Legacies of Pre-Colonial States." *International Organization* 73 (3):645–83. doi: 10.1017/S0020818319000134

Phillips, Andrew, and Jason C. Sharman. 2015. *International Order in Diversity: War, Trade and Rule in the Indian Ocean*. Cambridge: Cambridge University Press.

Pinker, Steven. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Penguin Books.

Ringmar, Erik. 2012. "Performing International Systems: Two East-Asian Alternatives to the Westphalian Order." *International Organization* 66 (1):1–26. doi: 10.1017/S0020818312000033

Ruggie, John. 1998. *Constructing the World Polity*. New York: Taylor and Francis Group.

Sarkees, M. R. 2010. "Extra-state Wars (Version 4.0): Definitions and Variables by Meredith Reid Sarkees." Accessed November 14, 2019. COW War Data, 1816–2007 (v4.0) Accessed November 14, 2019. http://www.correlatesofwar.org/data-sets/COW-war/extra-state-wars-codebook

Sarkees, M. R., and F. Wayman. 2010. *Resort to War: 1816–2007. Correlates of War*. Washington, DC: CQ Press.

Spruyt, Hendrick. 1994. *The Sovereign State and Its Competitors*. Princeton: Princeton University Press.

Tambiah, Stanley J. 1977. "The Galactic Polity: The Structures of Traditional Kingdoms in Southeast Asia." *Annals of the New York Academy of Sciences* 293:69–97. doi: 10.1111/j.1749-6632.1977.tb41806.x.

Teorell, Jan. 2017. "Rules of Recognition? Explaining Diplomatic Representation in the Long Nineteenth Century." STANCE Working Paper Series, Lund, Sweden: Lund University 3, 001041401984343.

Weber, Max. 1946. "Politics as Vocation." In *From Max Weber: Essays in Sociology*, edited by H.H. Gerth and Wright Mills. New York: Oxford University Press. 77-128.

Weidmann, Nils B. 2016. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* 60 (1):206–18. doi: 10.1111/ajps.12196

Wig, Tore. 2016. "Peace from the Past: Pre-colonial Political Institutions and Civil Wars in Africa." *Journal of Peace Research* 53 (4):509–24. doi: 10.1177/0022343316640595

Wilks, Ivor. 1975. *Asante in the Nineteenth Century: The Structure and Evolution of a Political Order*. Cambridge: Cambridge University Press.