SOCIUM · SFB 1342
WorkingPapers No. 2

Bastian Becker

Introducing COLDAT: The Colonial Dates Dataset

socium
Forschungszentrum
Ungleichheit und Sozialpolitik

Globale
Entwicklungsdynamiken
von Sozialpolitik
SFB 1342

Bastian Becker
Introducing COLDAT: The Colonial Dates Dataset
SOCIUM SFB 1342 WorkingPapers, 2
Bremen: SOCIUM, SFB 1342, 2019

Universität Bremen
SOCIUM Forschungszentrum Ungleichheit und Sozialpolitik /
Research Center on Inaquality and Social Policy
SFB 1342 Globale Entwicklungsdynamiken von Sozialpolitik /
CRC 1342 Global Dynamics of Social Policy

Postadresse / Postaddress:
Postfach 33 04 40, D - 28334 Bremen

Websites:
https://www.socium.uni-bremen.de
https://www.socialpolicydynamics.de

[DOI https://doi.org/10.26092/elib/4395]

[ISSN (Print) 2629-5733]
[ISSN (Online) 2629-5741]

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG)
Projektnummer 374666841 – SFB 1342

DFG Deutsche Forschungsgemeinschaft

Bastian Becker

Introducing COLDAT: The Colonial Data Dataset

SOCIUM · SFB 1342
No. 2

Bastian Becker (bastian.becker@uni-bremen.de), University of Bremen, SOCIUM Research Center on Inequality and Social Policy, Mary-Somerville-Str. 3, 28359 Bremen, Germany.

Acknowledgements:
I thank Judith Ebeling, Johanna Kuhlmann, Carina Schmitt, Amanda Shriwise and two anonymous reviewers for helpful comments and suggestions, and Maximilian Pfeil for excellent research assistance. All remaining errors are my own. This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 755129). No conflicts of interest declared.

Abstract

Social scientists increasingly turn to historical research to understand long-term institutional and societal change. However, data availability and quality, including disagreements about basic historical facts, remain central challenges to this line of work. When it comes to research on European colonial empires and their legacies, social scientists draw on a number of secondary sources, which differ in scope, detail, and coding decisions. Thus, findings risk being driven by the choice of the data source rather than substantive differences. To address this shortcoming, I introduce the Colonial Dates Dataset (COLDAT), which aggregates information on the reach and duration of European colonial empires from renowned secondary sources. By aggregating secondary sources, rather than collecting from primary sources, the new dataset reflects the accumulated knowledge in the discipline and relieves researchers from making hard to justify choices between different historical datasets.

Zusammenfassung

Sozialwissenschaftler nutzen vermehrt historische Forschung um langfristigen institutionellen und gesellschaftlichen Wandel zu verstehen. Allerdings stellen die Verfügbarkeit und Qualität von Daten, auch in Bezug auf grundlegende historische Fakten, weiterhin eine zentrale Herausforderung für diese Arbeit dar. Forschungsarbeiten zu europäischen Kolonialimperien und ihren Vermächtnissen nutzen eine Reihe von Sekundärquellen, die sich allerdings in Umfang, Detail und Kodierungsentscheidungen unterscheiden. Daher ergibt sich die Gefahr, dass Forschungsergebnisse Resultat von Quellenwahl, und nicht inhaltlicher Natur, sind. Um dieser Gefahr entgegenzuwirken, stelle ich hier den Kolonialdaten Datensatz (COLDAT) vor, welcher aggregierte Information zur Reichweite und Zeitdauer europäischer Kolonialimperien, aufbauend auf renommierten Sekundärquellen, enthält. Durch die Aggregierung von Sekundärquellen, an Stelle einer Primärdatensammlung, reflektiert der neue Datensatz das akkumulierte Wissen der Disziplin und befreit Forscher davon Entscheidungen zwischen Sekundärdatensätzen treffen zu müssen, die schwer zu rechtfertigen sind.

Contents

1. Introduction

In recent years, historical research has rekindled the interest of many social scientists. This research hopes to overcome biases rooted in a too narrow focus on present times and to uncover processes that only unfold in the long-term. Historical research on colonialism, for example, has contributed greatly to our understanding of long-term institutional and societal change. However, historical research is no panacea. Data collection and quality, including disagreements about basic historical facts, remain central challenges to this line of work. This paper is committed to one such historical basic: the reach and duration of European colonial empires. Scholars draw on a variety of data sources, which differ in terms of samples and coding strategies and can thus bias findings decisively. In this paper, I present a new unified dataset, Colonial Dates Dataset (COLDAT), that combines the most prominent social science datasets on colonialism. COLDAT synthesizes the current knowledge about the basics of colonialism and relieves scholars from making hard to justify choices between different datasets.[1]

Social scientists have studied colonialism to shed light on a range of social science questions. Amongst others, scholars in the quantitative-comparative tradition, who are the first to benefit from a unified dataset on colonial dates, debate the effect of colonialism on economic growth (Acemoglu, Johnson, & Robinson, 2000; Glaeser, La Porta, Lopez-de Silanes, & Shleifer, 2004), social development (Lange, 2004; Mahoney, 2010), conflict (Mamdani, 1996; Wucherpfennig, Hunziker, & Cederman, 2016), democracy (Olsson, 2009; Woodberry, 2012), quality of government (La Porta, Lopez-de Silanes, Shleifer, & Vishny, 1999; Treisman, 2000), social policy (Frankema, 2012; Schmitt, 2015), fiscal capacity (Frankema & Waijenburg, 2014; Huillery, 2014), and international aid (Alesina & Dollar, 2000; Fuchs, Dreher, & Nunnenkamp, 2014). For comprehensive reviews of the literature, see De Juan and Pierskalla (2017) and Michalopoulos and Papaioannou (2018).

The explanatory variables on which social scientists interested in colonialism rely differ. Some simply want to know whether a country has been previously colonized and/or by whom. Others rely on specific dates or the duration of colonialism, overall or by colonizer. Some scholars also use information about colonialism to determine the samples and cases they want to study. As the different datasets available to scholars of colonialism differ in terms of samples and coding stategies, the choice of the dataset alone risks determining the inferences scholars make. For example, some datasets focus only on major empires or specific colonizers, e.g. the main or last colonizer. Even if they focus on the same colonizer, they often differ on the chosen start and end years. In this paper I discuss and resolve these differences. The resulting unified dataset, COLDAT, provides the broadest coverage in terms of sample and variables of any dataset on the reach and duration of European colonial empires to date.

2. Building a Unified Dataset of European Colonialism

The unified COLDAT dataset provides information on all contemporary nation states and how they have been affected by European colonial empires. To identify whether a country has previously been colonized and for what period, I lay out a procedure to ag-

[1] The complete dataset can be downloaded from Harvard's Dataverse (https://doi.org/10.7910/DVN/T9SDEW).

gregate information from different prominent secondary data sources. This procedure is largely automated, relying on predefined coding rules. The only exception is a small number of missing dates, which are added manually to complete the dataset, and erroneous dates in the data sources, which are excluded. Overall, COLDAT constitutes the most complete and accurate dataset on the historic and geographic reach of European colonial empires.[2]

Table 1 provides an overview of the secondary data sources on which COLDAT is based. In order to merge the different data sources, it is necessary to define key terms. A colony is a territory whose domestic and/or foreign affairs are dominated by a European nation state, and whose population is constructed as inferior to the colonizer. A European colony is a territory outside of Europe colonized by a European power. This definition encapsulates those of the secondary data sources used for the construction of COLDAT. Regarding the timing of colonialism, a territory is usually regarded as a colony once external control is established over a significant part of its territory. A formal declaration is not necessary.[3] Finally, a colony is considered disolved once control by a European power vanished.[4] According to this definition, mandated territories, which are administered by an external power, retain the status as a colony.

2.1 Point of Departure: The "Correlates of War" Data Base

The core of COLDAT is based on the Correlates of War (CoW) data base. CoW consists of a multitude of datasets that provide annual snapshots of the nation state system, relationships between nation states, and other geopolitical units that are dependent on them (Sarkees & Wayman, 2010). Colonial histories of contemporary nation states register in two different ways in the CoW data base: (1) they are successors of geopolitical units that were previously part of colonial empires, where such dependencies can be direct or indirect (i.e. through dependencies on other geopolitical units), or (2) they absorbed geopolitical units that were previously dependent on one of the colonial powers. In the following I describe how information on European colonial empires can be extracted from CoW.

Contemporary Nation States

Contemporary nation states are identified as all members of the nation state system on December 31, 2016. CoW[5] defines nation states as follows, "the entity must be a member of the United Nations or League of Nations, or have population greater than 500,000 and receive diplomatic missions from two major powers [emphasis added]" (p.5). After excluding all European nation states,[6] the extracted list consists of 151 nation states.

2 COLDAT remains incomplete in so far as colonization of certain territories are not mentioned in any of the sources.

3 Note that this does not imply that the succeding nation state has to have been occupied itself for it to qualify as a former colony. It is sufficient if the nation state absorbed a territory that previously qualified as a colony. This further implies that it is possible for a contemporary nation state to have been exposed to different colonial powers at the same time.

4 Usually the dissolution of a European colony coincides with the achievement of independence. However, some countries fell under the control of other countries for an interim period, e.g. Namibia to South Africa or Bangladesh to Pakistan.

5 Correlates of War Project. 2016. State System Membership List, v2016. Online, http://correlatesofwar.org.

6 By my definition they cannot have been European colonies.

Table 1.
Data Sources and Definitions of Colonialism.

Source	Definition	Start points	End points	Empires covered[a]
CoW	"fairly durable status in which the entity exercised almost no control over its foreignaffairs, armed forces, immigration, or trade."	not specified, truncated at year 1816	not specified	All
Lange et al., 2006	"colonialism describes a broad range of institutions implanted by colonial occupiers, it is useful to disaggregate this concept into more specific institutions that operate as key mechanisms in shaping postcolonial development. We especially focus on those institutions that regulate (1) commerce and markets (e.g., the extent of free trade), (2) political authority (e.g., the degree to which a rule of law is present), and (3) race and ethnicity (e.g., the degree to which all groups have the same rights)." (p.1419)	Establishment of enduring control	Complete defeat or withdrawl of colonial authorities	Britain, Spain
Wimmer & Min, 2006	"[E]mpire is defined by the following institutional features: centralized bureaucratic forms of government, the domination of a core region over peripheries, an ethnically or culturally defined hierarchy between rulers and ruled, and claims to universal legitimacy–whether referring to a revolutionary ideology (e.g., the Soviet Union), a mission civilisatrice (e.g., colonial empires), or religious conversion (e.g., the Spanish empire)" (p.870)	Territory administered by occupying force or garrison with aim to expand military control or official status as protectorate or colony.	Creation of modern nation state[b]	all
Olsson, 2009	"A Western colony is a new and lasting political organization created outside Europe by Western countries (countries in Europe excluding Russia but including the Western offshoots United States, Australia, New Zealand, and Canada) from the 15th to the 20th centuries through either invasion and conquest, and/or settlement colonization. Its rulers are in sustained dependence on a geographically remote mother country or imperial center that claims exclusive rights of possession of the colony or in other ways strongly dominates politics in the country." (p.536)	Western colonizers became the major political power	Independence	Britain, France, Spain

Notes: a Only former colonial empires for which data is extracted listed here. Some data sources include information on other empires. *b* Wimmer and Min (2006, p.880) "coded as the year of nation-state creation the date when a territory began to be governed on the basis of a written constitution that identified a national group as the sovereign of the state, whether the nation was defined in multiethnic or monoethnic terms."

Colonial Legacies

CoW specifies different kinds of dependencies: colonies, mandates, occupations, protectorates, annexations, neutral/demilitarized zones, leasings, and claims.[7] The definition of colonies is based on Russet, Singer, and Small (1968) who define them as being "characterized by a fairly durable status in which the entity exercised almost no control over its foreign affairs, armed forces, immigration, or trade" (p.924). Before a later extension of the CoW data base, protectorates were also subsumed under this category (Sarkees & Wayman, 2010, pp.27-29). Following this initial approach, I count both dependencies, "colony" and "protectorate", as colonial dependencies. Furthermore, I also code mandates as colonial dependencies. CoW has maintained the original definition of mandates as "territories whose gradual transition to independence was the moral and legal responsibility of the metropolitan power assigned to it by the League of Nations or the United Nations" (Russet et al., 1968, p.924). As such, territories are still under the effective control of an external power, making them more like a colony than an independent nation state.

7 CoW makes this information available in pdf-form. The corresponding file ("entities.pdf") is a building block of many CoW datasets. Here the version released with the territorical change dataset (v2014) is used.

As mentioned above colonial dependencies can be either direct or indirect.[8] In the direct case, the geopolitical unit associated with a contemporary nation state had a colonial dependency with one of the European colonial powers. The corresponding information on the colonial power and years of the dependency are then extracted. In the indirect case, the associated geopolitical unit was dependent[9] on another geopolitical unit that had a colonial dependency with a European colonial power. In extracting the corresponding information time periods are limited to years in which the dependencies overlap.

Rather than having been directly or indirectly incorporated into colonial empires, contemporary nation states are also regarded to have a colonial legacy if they absorbed a geopolitical unit that previously had a colonial dependency. This information can be drawn from the CoW Territorial Change dataset (see Tir, Schafer, Diehl, & Goertz, 1998). This dataset contains information on exchanges of territories between nation states and other geopolitical units included in CoW from 1816–2014. Here, absorption of one geopolitical unit by another is defined as exchanges which (1) include a territory of significant size, (2) concern "homeland territory", and (3) are not later reverted.[10] If these conditions are satisfied, then the colonial legacy of the absorbed geopolitical unit before absorption constitutes a legacy of the absorbing geopolitical unit.

The CoW data base is neither complete nor is the information it contains undisputed. In addition, coding decisions about the extraction of information from the CoW data base drive when and what territories are considered colonies. This would change if one set different criteria, for example, for the treatment of mandates or absorbed territories. That being said, the decisions here were driven by the definition layed out upfront and are in line with common practice. It is also important to keep in mind that the CoW data is truncated at year 1816. As can be seen from the upper-left panel in Figure 1, CoW at best provides an incomplete picture of final two centuries of European colonial history. To complete the information on European colonial empires the data extracted from CoW needs to be complemented with information from other sources. The following section introduces the sources I use for this purpose.

2.2 Complementary Data Sources

Most researchers draw on a small number of datasets for information on colonial empires. In this section I introduce the most prominent alternatives to CoW. One reason why these alternatives are often preferred is that they provide data in a single table and do not require users to combine information from a multitude of tables.[11] Following this introduction, I explain how I aggregate information from these different sources to build the new COLDAT dataset. Table 1 provides an overview of all source datasets including informa-

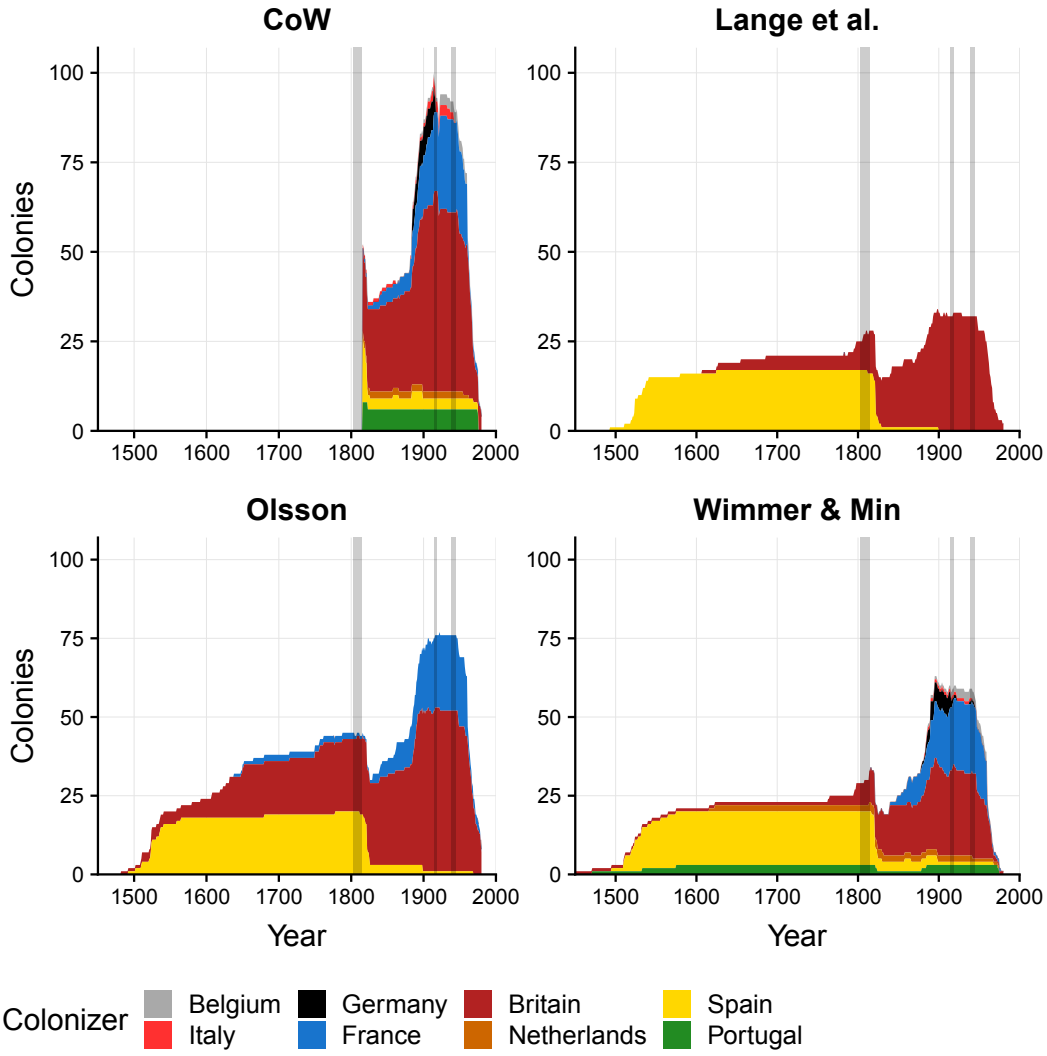
8 Direct and indirect here simply refers to the relational coding scheme in CoW. It should not to be confused with the common distinction of direct and indirect rule.

9 Either in terms of a colonial dependency or as a possesion, leasing, occupation, or simply by being part of the other geopolitical unit. Note that indirect dependencies can consist of even longer chains of dependencies.

10 An absorbed territory is considered of significant size if, after accounting for later losses, it exceeds 5% of the area of the contemporary nation state. Thus, the colonial legacy of small absorbed territories, such as Zanzibar (absorbed by Tanzania) or French possesions in India, is not carried forward.

11 The complexity of CoW is due to its goal to serve wide variety of research needs, whereas the other datasets introduced here are limited to specific purposes.

Figure 1.
Timeline of European Colonial Empires (Based on Source Datasets).



Note: Counts of colonies computed based on dates and identities of colonizers as indicated in each dataset. For Olsson, full duration of each colony is assigned to last colonizer as no information on other colonizers is provided. The grey-shaded areas indicate the Napoleonic Wars (1803-1815), World War I (1914-18), and World War II (1939-45).

tion on how colonialism is defined and which European colonial empires are covered.

Lange et al., 2006

A frequently used dataset on the geographical and temporal reach of the British and Spanish colonial empires is provided by Lange et al. The authors adopt an institutionalist perspective on colonialism that emphasizes the control a colonial power exerts over its dependencies. This perspective is best captured by the definition of the key concept in their article. "We define level of colonialism as the extent to which a colonizing power installs economic, political, and sociocultural institutions in a colonized territory [emphasis added]" (p.1414).[12]

12 The authors further state: "Because level of colonialism describes a broad range of institutions implanted by colonial occupiers, it is useful to disaggregate this concept into more specific institutions that operate as key mechanisms in shaping postcolonial development. We especially focus on those institutions that regulate (1) commerce and markets (e.g., the extent of free trade), (2) political authority (e.g., the degree to which a rule of law is present), and (3) race and ethnicity (e.g., the

With regards to the selected start and end points of the colonial period, Lange et al. state the following: "Identifying a single starting point of colonialism is problematic for many cases, given that the initiation of the process was often gradual and informal. For the Spanish-American cases, we date the onset of colonialism with the foundation of major settlements or expeditions that established enduring control over the indigenous population. The initiation of British colonialism is often especially difficult to date. [...] We note several cases in which multiple dates could be used to mark the beginning of British colonialism. The conclusion of colonialism corresponds with the more or less complete defeat and/or withdrawal of colonial authorities rather than simply the declaration of independence." (p.1418) While Lange et al. subsume protectorates under their definition, they exclude most mandated territories from their dataset.

Wimmer & Min, 2006

Together with their highly-cited article, Wimmer & Min provide a dataset that includes information on time periods of empires. Although they do not limit themselves to colonial empires, their definition encapsulates European colonial empires. "[E]mpire is defined by the following institutional features: centralized bureaucratic forms of government, the domination of a core region over peripheries, an ethnically or culturally defined hierarchy between rulers and ruled, and claims to universal legitimacy–whether referring to a revolutionary ideology (e.g., the Soviet Union), a mission civilisatrice (e.g., colonial empires), or religious conversion (e.g., the Spanish empire)" (p.870). Due to the similarity of definitions, the more detailed information on European empires contained in Wimmer & Min's dataset can be used as a complement. Wimmer & Min determine start and end dates of empires as follows. "In order to determine the year in which a territory was considered to be part of a larger political entity (usually an empire), we searched for evidence of one of the following and coded the year of incorporation to whichever came first:

- The territory is effectively administered by an occupying force.
- A garrison is established that aims at expanding military control over the territory.
- The territory becomes a protectorate or colony.

The establishment of military posts that serve only to provide military protection to foreign traders, however, was not treated as a case of imperial incorporation. Temporary military occupation that lasted three or fewer years and that was not intended to permanently 'absorb' the occupied territory into the state was not coded as imperial incorporation and were coded as periods of military occupation." (Codebook accompanying Wimmer and Min (2006), p.8)

Olsson, 2009

Another comprehensive dataset of European colonialism is provided by Olsson. Olsson specifies the following definition of colonialism to identify 143 contemporary nation states (reference year 2009) that were former colonies. "A Western colony is a new and lasting political organization created outside Europe by Western countries (countries in Europe excluding Russia but including the Western offshoots United States, Australia, New Zealand, and Canada) from the 15th to the 20th centuries through either invasion and conquest, and/or settlement colonization. Its rulers are in sustained dependence on a geographically remote mother country or imperial center that claims exclusive rights

degree to which all groups have the same rights)." (p.1419)

of possession of the colony or in other ways strongly dominates politics in the country." (p.536)

Based on this definition, Olsson determines start and end dates of colonial dependencies. End dates are chosen to correspond to years of independence. In recognizing the problematic nature of determining start dates of colonialism, the author states, "[w]e have tried to stay close to our definition of a colony when determining the date of colonization. In particular, we have tried to identify a date when Western colonizers became the major political power in the region. This date is sometimes hundreds of years before the formal declaration of colony status, which some previous authors have used" (p.536).

For the purposes of building a unified dataset, there are two shortcomings in Olsson's data. First, it only identifies the last colonizers (and only Britain, France, and Spain). Thus, prior changes of the colonizer, as for instance in the case of Cameroon, are not documented. Second, Olsson takes years of independence to indicate the end of European colonizations. There are cases were several years pass between the end of European colonizations and the achievement of independence, such as in the case of Bangladesh or Namibia. Below I detail a procedure to ascertain that only information aligned with the definition outlined above is used.

Timelines

Based on the different source datasets, it is possible to construct timelines of European colonial empires. As Figure 1 illustrates these timelines vary considerably. Unlike CoW, the three complementary datasets do not contain a sharp cut-off and instead portray the complete temporal reach of the empires. However, they suffer from other shortcomings. The data provided by Lange et al. (2006) and Olsson (2009) does not cover all eight European colonial empires. Wimmer and Min (2006) provide information on all empires but exclude many island nations and mandated territories. Furthermore, as Olsson's dataset only includes information on the last colonizer, the respective timeline could only be constructed by associating the start date with the last colonizer.[13] However, this approach is problematic as colonies changed hands in several cases. In constructing COLDAT I follow a more dedicated strategy (described below) to adress this challenge. It should finally be noted that timelines cannot show all discrepancies between the datasets. If one looks closer at specific cases, many further discrepancies, especially regarding start dates, can be revealed.

3. Aggregating Data Sources

Before aggregating the data sources a number of preparatory steps are necessary. First, all data sources are brought into a common format. The units of observation are contemporary nation states and each is characterized by sixteen variables: A start and an end date for each of the eight European colonial powers. In case a nation state was not part of a given colonial empire, these dates are simply coded as missing. Second, start dates in CoW equaling 1816 are deleted as this constitutes the first year the data source covers, i.e. start dates are truncated. Third, Olsson provides information only on the identity of the last colonizer. Unless one wants to assume colonizers never changed, which is empirically false, the "anonymous" start dates indicated in Olsson cannot be associated with a specific colonizer. However, in the main text of his article (but not in the accompany ing dataset), Olsson lays out short explanations for each chosen start date.

13 In the paper accompanying his dataset, Olsson follows the same approach.

Based on these explanations, I am able to identify the colonizers relating to all "anonymous" start dates.[14] An additional benefit of this coding strategy is that the start dates are now coded for more than the three colonizers in Olsson's original dataset. Fourth, all end dates in Olsson correspond to independence years. These often coincide with the end of European colonizations. However, I ascertain this for all entries individually. If they do not coincide, I exclude the respective value from the aggregation.[15]

As a final fifth step before the aggregation, the issue of missing dates needs to be addressed. Most missing data results from the incorporation of Olsson's data. In his dataset only the last colonizer is identified explicitly, such that the colonizer can be associated with the provided end date but not the start date. For the start dates in Olsson I manually coded the identity of the first colonizer. This, however, has the effect that there is no associated end date, if the first colonizer differs from the last colonizer. Furthermore, unless other data sources provided corresponding information, start dates are missing for colonies that were only in CoW and their beginning truncated at 1816. In total there are 25 missing dates.[16] They are filled manually, primarily with information from The World Factbook (CIA, 2016). Following these preparatory steps, the raw data can be aggregated.

Table 2.
Information on Colonial History of Canada in Source Datasets.

Source	Colonizer	Start date	End date
CoW	Britain	(1816)	1919
Lange et al.	Britain	1686	1867
Wimmer & Min	Britain	1763	1866
Olsson	Britain		1867
Olsson	France	1608	
Manual coding	France		1763

Note: CoW date in parantheses due to indicate trunction in 1816. End date of French colonization coded manually as no information contained in source datasets.

There is no single, or best, approach to aggregate data from secondary sources. One option is to regard discrepancies between datasets as an indication of measurement error. If this is the case and one chooses not to prioritize any source, unweighted means can suitably be used for aggregation. One disadvantage of this approach is that the resulting figure might not be related to a specific event (such as an invasion or declaration) but rather lie in between them. While it can be argued that the beginning and end of colonialism is a transitional process, others will insist that it is a discrete phenomenon and therefore has to be associated with specific events. In fact, all sources included here base their coding of start and end dates on event-based historical research. In line with this thinking, one might require that all sources agree that an event has occured that concludes the establishment/dissolution of

14 Some explanations specifically refer to the identity of the colonizers, in other cases additional research to identify the colonizer related to the given explanation was necessary. In total, colonizers for 121 start dates were identified.

15 For the following cases, end dates from Olsson (2009) are excluded: British colonization of Belize and Eritrea, French colonization of Cambodia, and Spanish colonization of Mexico and Nicaragua. Thanks to a comment by a reviewer, I also exclude the second end date of the British colonization of Bangladesh from Lange et al. (2006) for the same reason, and the erroneous CoW start date of the British colonization of South Africa.

16 British start dates: Dominica, Grenada, Senegal, Seychelles, St. Lucia; British end dates: Marshall Islands; Dutch end dates: Guyana, South Africa; French start dates: Vanuatu; French end dates: Dominica, Grenada, Seychelles, St.Lucia, Canada, Mauritius; German end dates: Nauru, Solomon Islands; Portuguese end dates: Equatorial Guine, Malaysia, Sri Lanka; Spanish end dates: Belize, Jamaica, Micronesia, Palau, Trinidad & Tobago.

a colony. Aggregation can then be implemented by taking the last date mentioned across all sources, i.e. all sources agree that a certain territory has become/ceased to be a European colony.[17] Instead of imposing a choice on researchers, I include mean aggregates, which aim at measurement accuracy, as well as last date aggregates, which aim at event consensuality, in COLDAT.[18]

Canada provides an instructive example of the aggregation process. As can be seen from Table 2, all four source datasets contain information on Britain's colonization of the country. France's undertakings find mention only in Olsson's explanation of the chosen start date. As no source dataset contains information on the end date of French colonization, it needs to be coded manually. The chosen year is 1763 as Britain had by then established control over most of New France. For the aggregation, the start date of the British colonization mentioned in Correlates of War is ignored as it corresponds to the year in which the data is truncated (1816). As such, mean aggregation implies that British colonization of Canada began in 1724 and ended in 1880, based on two respectively four sources.[19] Last date aggregation identifies 1763 and 1919 as start and end dates of the British colonization. As there is only one data point for each the beginning and the end of French colonization, both aggregation methods lead to the same result, 1608 as start year respectively 1763 as end year.

The resulting COLDAT dataset is made available in wide and long format. The wide format includes all contemporary nation states (according to Correlates of War) as observations and variables indicating the presence and duration of different colonization. This includes colonizer-specific dummies to indicate whether a colonial power was ever present in a given territory as well as the aggregated start as well as end dates. Appendix I: Variable Description includes an overview of all variables included in COLDAT. The long format constitutes dyadic data, with one line for each pairing of a contemporary nation state and European colonial power (irrespective of whether they were present in the corresponding territory). The table includes the same information as the wide table, indicated by a colonial dummy and the respective start and end dates.

4. Comparing COLDAT to the Data Sources

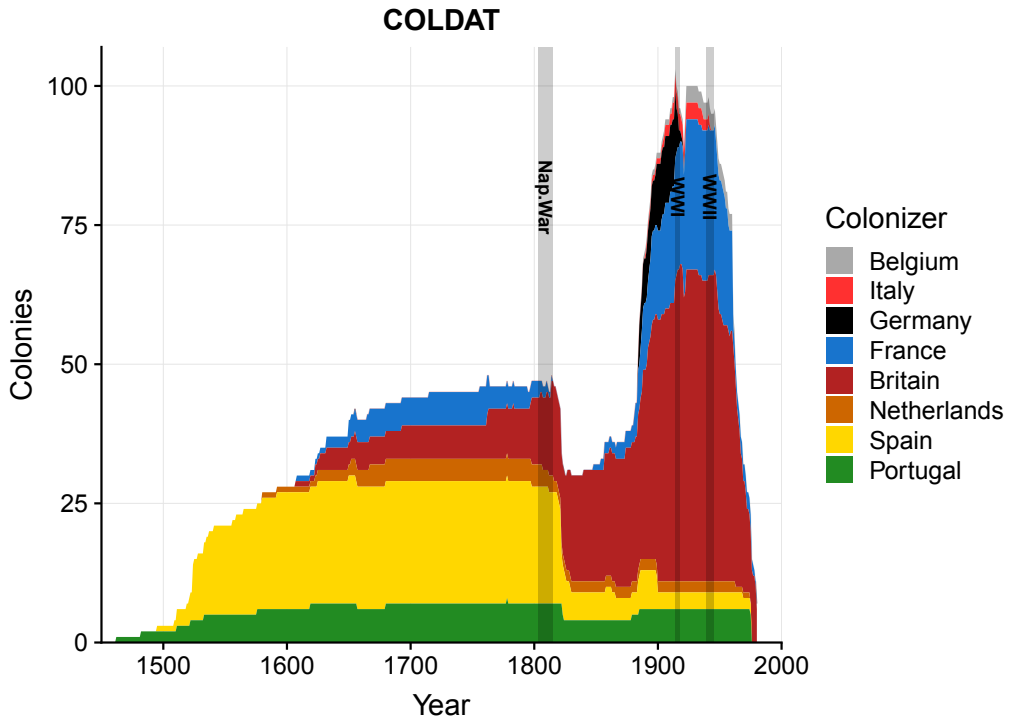
In this section, I shortly compare the contructed COLDAT dataset to the data contained in the source datasets on which it is built. For the comparisons, I focus on last date aggregates, which for each date equals the maximum value indicated in any source. As shown in Figure 2, COLDAT fully captures the rise and fall of the European colonial empires over the course of 500 years. For reference, the Napoleonic Wars and the two World Wars are also indicated. A detailed list of all contemporary nation states that were part of each colonial empires, and the respective start and end dates, is included in the Appendix II: Tables & Graphs (Table 4). Corresponding summaries using mean ag-

[17] A particularity of the Lange et al. (2006) dataset is that for some dates it provides two years. In case of mean aggregation, these two years are averaged before averaging across all sources. In this way, equal weighting is preserved. In case of last date aggregation, the later date is used as input for the aggregation process. Note that I exclude the second end date the authors indicate for the British colonization of Bangladesh as it conincides with independence. However, independence was achieved from Pakistan, not Britain.

[18] The calculated means are rounded to the closest integer.

[19] Note that results are rounded to the closest integer.

Figure 2.
Timeline of European Colonial Empires (Last Date Aggregation).



Note: Based on COLDAT, last date aggregation. Territories are multiply counted in the case of simultaneous colonizations (e.g Cameroon from 1922-1960, colonized by Britain and France, is counted twice). Nap. War = Napoleonic Wars.

gregates can also be found there (see Figure 4 and Table 5).

From a visual comparison of the COLDAT timeline and the source timelines (see Figure 1), it becomes apparent that COLDAT surpasses all sources in terms of temporal and geographic coverage. In total, COLDAT contains information on 165 colonizations and respective start and end dates. Table 3 shows what number of these dates are included in each source dataset. Table 3 shows that most start and end dates, 135, can also be found in CoW. The dataset by Lange et al. provides the smallest coverage, which is unsurprising as it only includes information on the British and Spanish empires. Of course, information on many dates is contained in multiple source datasets. However, the fact that no source dataset comes close to full coverage shows the complementary gains of merging them.

Table 3.
COLDAT Start and End Dates, Coverage by Source.

Source	N Start	%COLDAT	N End	%COLDAT
CoW	135*	81.8*	135*	81.8*
Lange et al.	55	33.3	55	33.3
Olsson	121*	73.3*	102	61.8
Wimmer & Min	95	57.6	95	57.6

Note: Starred dates are derived from/manually coded based on the respective sources and coding rules described in sections Point of Departure: The "Correlates of War" Data Base and Complementary Data Sources.

In addition to providing complementary information, merging different data sources allows for the aggregation of existing knowledge, especially when the sources disagree on a specific date. As described above, this is done by computing means respectively determining last dates across the different sources (see Aggregating Data Sources). To

illustrate the last date aggregation process, Figure 3 plots the source dates (vertical axis) against the resulting COLDAT dates (horizontal axis).[20] Points that are on the diagonal indicate that the respective source dataset and COLDAT agree on a given date. That most points cluster closely around the diagonal indicates little disagreement between the source datasets themselves and with COLDAT. This impression is supported by the high correlation coefficients–the smallest being .972–of all pairwise comparisons with COLDAT.[21] Furthermore, the correlation coefficients are very close to each others, which is evidence that all source datasets are considered equally in the construction of COLDAT.

Notwithstanding the strong correlation between the source dates and COLDAT, in some cases discrepancies span several decades, or even centuries in the case of start dates (points far removed from the diagonal).[22] It is important to keep in mind that the information on historical events is usually less easy to attain the further the event lies in the past. Furthermore, colonization is a transitional process and scholars might therefore come to very different conclusions about when exactly a territory has been effectively colonized.[23] Despite some larger discrepancies, the mean differences between all source datasets and the COLDAT dates are all within single digits.[24]

5. Studying European Colonial History

When studying European colonial empires and their legacies, researchers rely on a variety of datasets. As these datasets do not have the same coverage and frequently disagree on specific dates, researchers risk that their findings are driven by the choice of dataset rather than substantive differences. Instead of relying on primary data collection, COLDAT joins the information contained in different datasets in a way that best represents current knowledge about colonial empires. By joining information from different datasets, COLDAT not only constitutes the most complete dataset on European colonial history to date, it also provides more detailed information than any of the source datasets. Overall, COLDAT offers a comprehensive source of information on the reach and duration of European colonial empires and relieves researchers from making hard to justify choices between different historical datasets. Nevertheless, researchers should keep a number of considerations in mind when drawing on COLDAT data. Most importantly, the two different aggregation schemes reflect different methodological predispositions: mean aggregation focuses on measurement accuracy, last date aggregation on event consensuality. While the former is usually better suited for statistical estimation, the latter is preferable for sample selection or qualitative comparative analysis. Furthermore, colonialism is not a binary phenomenon but comes

20 The corresponding figure using mean aggregates can be found in the Appendix II: Tables & Graphs (Figure 5).

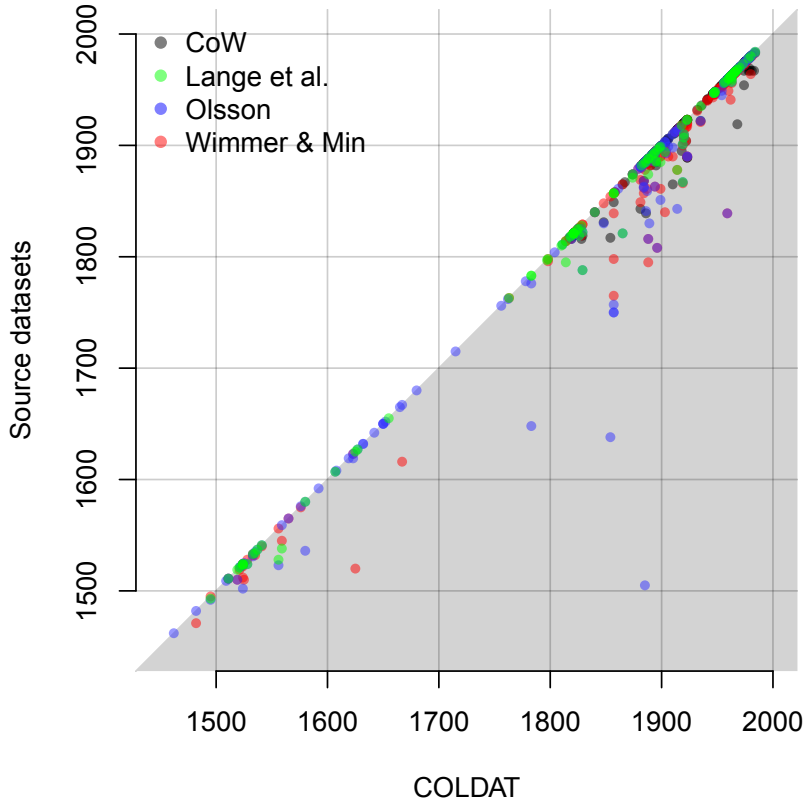
21 The correlation coefficients of each source datasets and the dates in COLDAT are as follows: CoW, .985; Lange et al., .998; Olsson, 972; Wimmer & Min, .989.

22 For example, start dates of the Portuguese colonization of Mozambique range from 1505 to 1885.

23 The same is true, although to a more limited extent, for the end of colonizations. For example, Britain established a number of dominions, such as Australia, Canada, or India, that delayed full independence and secured some of its influence. There is at least one data source (CoW) that considers governance models like dominions to qualify as colonies.

24 The mean absolute differences are as follows; CoW, 2.55; Lange et al., 3.49; Olsson, 9.96; Wimmer & Min, 8.55.

Figure 3.
Scatterplot of COLDAT (Last Date Aggregation) and Source Dates.



Note: The indicated COLDAT dates refer to last date aggregates. Only sources dates used as input to COLDAT are displayed.

in different coats and degrees. Researchers might thus want to focus on specific colonial powers, time periods, or add other qualifying information to the data COLDAT provides. Finally, country-level data necessarily brushes over sub-national variation. Such variation is particularly pronounced with regards to colonialism, where activities often focused on coastal areas and inlands were, if at all, affected much later. With these caveats in mind, and in anticipation of further, more detailed historical data, COLDAT hopefully proves to be a useful resource to researchers interested in colonialism and its legacies.

6. References

Acemoglu, D., Johnson, S., & Robinson, J. (2000, June). The Colonial Origins of Comparative Development: An Empirical Investigation (Tech. Rep. No. w7771). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w7771

Alesina, A., & Dollar, D. (2000, March). Who Gives Foreign Aid to Whom and Why? Journal of Economic Growth, 5(1), 33–63. doi:10.1023/A:1009874203400

CIA. (2016). The World Factbook. Central Intelligence Agency.

De Juan, A., & Pierskalla, J. H. (2017, June). The Comparative Politics of Colonialism and Its Legacies: An Introduction. Politics & Society, 45(2), 159–172. doi:10.1177/0032329217704434

Frankema, E. (2012, November). The origins of formal education in sub-Saharan Africa: was British rule more benign? European Review of Economic History, 16(4), 335–355. doi:10.1093/ereh/hes009

Frankema, E., & Waijenburg, M. v. (2014, November). Metropolitan Blueprints of Colonial Taxation? Lessons from Fiscal Capacity Building in British and French Africa, c. 1880–1940. The Journal of African History, 55(3), 371–400. doi:10.1017/S002185371400036X

Fuchs, A., Dreher, A., & Nunnenkamp, P. (2014, April). Determinants of Donor Gen erosity: A Survey of

the Aid Budget Literature. World Development, 56, 172–199. doi:10.1016/j.worlddev.2013.09.004

Glaeser, E. L., La Porta, R., Lopez-de Silanes, F., & Shleifer, A. (2004, September). Do Institutions Cause Growth? Journal of Economic Growth, 9(3), 271–303. doi:10.1023/B:JOEG.0000038933.16398.ed

Huillery, E. (2014, March). The Black Man's Burden: The Cost of Colonization of French West Africa. The Journal of Economic History, 74(01), 1–38. doi:10.1017/S0022050714000011

Lange, M. (2004, June). British Colonial Legacies and Political Development. World Development, 32(6), 905–922. doi:10.1016/j.worlddev.2003.12.001

Lange, M., Mahoney, J., & vom Hau, M. (2006, March). Colonialism and Development: A Comparative Analysis of Spanish and British Colonies. American Journal of Sociology, 111(5), 1412–1462. doi:10.1086/499510

La Porta, R., Lopez-de Silanes, F., Shleifer, A., & Vishny, R. (1999). The Quality of Government. Journal of Law, Economics, & Organization, 15(1), 222–279. https://www.jstor.org/stable/3554950

Mahoney, J. (2010). Colonialism and Postcolonial Development: Spanish America in Comparative Perspective. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511750328

Mamdani, M. (1996). Citizen and subject: contemporary Africa and the legacy of late colonialism. Princeton, NJ: Princeton Univ. Press. (OCLC: 318396100)

Michalopoulos, S., & Papaioannou, E. (2018). Historical Legacies and African Development. NBER. https://www.nber.org/papers/w25278

Olsson, O. (2009, December). On the democratic legacy of colonialism. Journal of Comparative Economics, 37(4), 534–551. doi:10.1016/j.jce.2009.08.004

Russet, B. M., Singer, J. D., & Small, M. (1968, September). National Political Units in the Twentieth Century: A Standardized List. American Political Science Review, 62(03), 932–951. doi:10.2307/1953441

Sarkees, M., & Wayman, F. (2010). Resort to War, 1816-2007. Washington, DC: SAGE Publications. doi:10.4135/9781608718276

Schmitt, C. (2015, June). Social Security Development and the Colonial Legacy. World Development, 70, 332–342. doi:10.1016/j.worlddev. 2015.02.006

Tir, J., Schafer, P., Diehl, P. F., & Goertz, G. (1998, February). Territorial Changes, 1816–1996: Procedures and Data. Conflict Management and Peace Science, 16(1), 89–97. doi:10.1177/073889429801600105

Treisman, D. (2000, June). The causes of corruption: a cross-national study. Journal of Public Economics, 76(3), 399–457. doi:10.1016/S00472727(99)00092-4

Wimmer, A., & Min, B. (2006). From Empire to Nation-State: Explaining Wars in the Modern World, 1816-2001. American Sociological Review, 71(6), 867–897. http://www.jstor.org/ stable/25472435

Woodberry, R. D. (2012, May). The Missionary Roots of Liberal Democracy. American Political Science Review, 106(2), 244–274. doi:10.1017/S0003055412000093

Wucherpfennig, J., Hunziker, P., & Cederman, L.-E. (2016). Who Inherits the State? Colonial Rule and Postcolonial Conflict. American Journal of Political Science, 60(4), 882–898. doi:10.11 11/ajps.12236

Appendix

Appendix I: Variable Description

Wide format (Country-level)

country Name of contemporary nation state.

col.* Dummy indicating whether *country* was ever colonized by (*) Belgium, Britain, France, Italy, Germany, Netherlands, Portugal, or Spain.

colstart.*_max First year during which *country* was colonized by (*) Belgium, Britain, France, Italy, Germany, Netherlands, Portugal, or Spain. Based on last date aggregation.

colend.*_max Last year during which *country* was colonized by (*) Belgium, Britain, France, Italy, Germany, Netherlands, Portugal, or Spain. Based on last date aggregation.

colstart.*_mean First year during which *country* was colonized by (*) Belgium, Britain, France, Italy, Germany, Netherlands, Portugal, or Spain. Based on mean aggregation.

colend.*_mean Last year during which *country* was colonized by (*) Belgium, Britain, France, Italy, Germany, Netherlands, Portugal, or Spain. Based on mean aggregation.

Long format (Country-colonizer dyads)

country Name of contemporary nation state. colonizer Name of the colonial power.

col Dummy indicating whether *country* was ever colonized by *colonizer*.

colstart_max First year during which *country* was colonized by *colonizer*. Based on last date aggregation.

colend_max Last year during which *country* was colonized by *colonizer*. Based on last date aggregation.

colstart_mean First year during which *country* was colonized by *colonizer*. Based on mean aggregation.

colend_mean Last year during which *country* was colonized by *colonizer*. Based on mean aggregation.

Table 4.

Former European Colonies by Colonial Power (Based on COLDAT, Last Date Aggregation).

Empire	Colonies
Belgium (3)	Burundi (1923-1962), Congo - Kinshasa (1885-1960), Rwanda (1922-1962)
Britain (70)	Antigua & Barbuda (1632-1981), Australia (1829-1919), Bahamas (1783-1973), Bahrain (1861-1971), Bangladesh (1857-1947), Barbados (1627-1966), Belize (1798-1981), Bhutan (1910-1949), Botswana (1895-1966), Brunei (1888-1984), Cameroon (1922-1961), Canada (1763-1919), Cyprus (1914-1960), Dominica (1805-1978), Egypt (1882-1935), Equatorial Guinea (1827-1857), Eritrea (1941-1951), Fiji (1874-1970), Gambia (1888-1965), Ghana (1874-1957), Grenada (1762-1974), Guyana (1814-1966), India (1857-1947), Iraq (1920-1932), Israel (1923-1948), Jamaica (1655-1962), Jordan (1923-1946), Kenya (1895-1963), Kiribati (1892-1979), Kuwait (1914-1961), Lesotho (1884-1966), Libya (1946-1952), Malawi (1891-1964), Malaysia (1888-1963), Maldives (1887-1965), Marshall Islands (1886-1885), Mauritius (1810-1968), Myanmar (Burma) (1886-1948), Namibia (1915-1920), Nauru (1914-1968), New Zealand (1840-1920), Nigeria (1899-1960), Pakistan (1857-1947), Papua New Guinea (1884-1920), Qatar (1916-1971), Samoa (1914-1920), Senegal (1693-1817), Seychelles (1814-1976), Sierra Leone (1896-1961), Singapore (1826-1963), Solomon Islands (1893-1978), Somalia (1888-1960), South Africa (1814-1920), Sri Lanka (1798-1948), St. Kitts & Nevis (1623-1983), St. Lucia (1814-1979), St. Vincent & Grenadines (1762-1979), Sudan (1898-1956), Swaziland (1903-1968), Tanzania (1918-1963), Tonga (1900-1970), Trinidad & Tobago (1797-1962), Tuvalu (1892-1979), Uganda (1894-1962), United Arab Emirates (1892-1971), United States (1607-1783), Vanuatu (1906-1980), Yemen (1959-1967), Zambia (1923-1964), Zimbabwe (1923-1980)
France (34)	Algeria (1848-1962), Benin (1894-1960), Burkina Faso (1895-1960), Cambodia (1884-1953), Cameroon (1922- 1960), Canada (1608-1763), Central African Republic (1906-1960), Chad (1910-1960), Comoros (1914-1975), Congo - Brazzaville (1882-1960), Côte d'Ivoire (1889-1960), Djibouti (1884-1977), Dominica (1632-1763), Gabon (1886-1960), Grenada (1650-1762), Guinea (1881-1958), Haiti (1665-1804), Laos (1893-1954), Lebanon (1923-1946), Madagascar (1895-1960), Mali (1904-1960), Mauritania (1903-1960), Mauritius (1715-1810), Morocco (1912-1956), Niger (1922-1960), Senegal (1854-1960), Seychelles (1756-1794), St. Lucia (1650-1814), Syria (1923-1946), Thailand (1867-1941), Togo (1922-1960), Tunisia (1881-1956), Vanuatu (1887-1980), Vietnam (1887-1954)
Germany (12)	Burundi (1899-1922), Cameroon (1884-1916), Ghana (1884-1916), Marshall Islands (1885-1914), Namibia (1885-1915), Nauru (1888-1914), Palau (1885-1914), Rwanda (1899-1915), Samoa (1900-1914), Solomon Islands (1885-1918), Tanzania (1891-1917), Togo (1885-1916)
Italy (3)	Eritrea (1890-1941), Libya (1912-1942), Somalia (1905-1936)
Netherlands (4)	Guyana (1580-1815), Indonesia (1623-1962), South Africa (1652-1806), Suriname (1667-1974)
Portugal (12)	Angola (1576-1975), Brazil (1533-1822), Cape Verde (1462-1975), Equatorial Guinea (1778-1778), Ghana (1482-1823), Guinea-Bissau (1879-1974), Malaysia (1511-1641), Mozambique (1885-1975), São Tomé & Príncipe (1522-1975), Sri Lanka (1619-1656), Timor-Leste (1642-1975), Uruguay (1680-1822)
Spain (27)	Argentina (1580-1819), Belize (1524-1862), Bolivia (1559-1825), Chile (1541-1818), Colombia (1525-1820), Costa Rica (1524-1821), Cuba (1511-1899), Dominican Republic (1495-1865), Ecuador (1535-1829), El Salvador (1528-1821), Equatorial Guinea (1858-1968), Guatemala (1524-1821), Honduras (1524-1821), Jamaica (1509-1655), Mauritania (1884-1975), Mexico (1521-1821), Micronesia (Federated States of) (1650-1899), Morocco (1884-1975), Nicaragua (1524-1821), Palau (1886-1899), Panama (1519-1821), Paraguay (1537-1811), Peru (1533-1824), Philippines (1565-1898), Trinidad & Tobago (1592-1797), Uruguay (1625-1828), Venezuela (1556-1821)

Countries with multiple colonizers: Belize (2), Burundi (2), Cameroon (3), Canada (2), Dominica (2), Equatorial Guinea (3), Eritrea (2), Ghana (3), Grenada (2), Guyana (2), Jamaica (2), Libya (2), Malaysia (2), Marshall Islands (2), Mauritania (2), Mauritius (2), Morocco (2), Namibia (2), Nauru (2), Palau (2), Rwanda (2), Samoa (2), Senegal (2), Seychelles (2), Solomon Islands (2), Somalia (2), South Africa (2), Sri Lanka (2), St. Lucia (2), Tanzania (2), Togo (2), Trinidad & Tobago (2), Uruguay (2), Vanuatu (2).

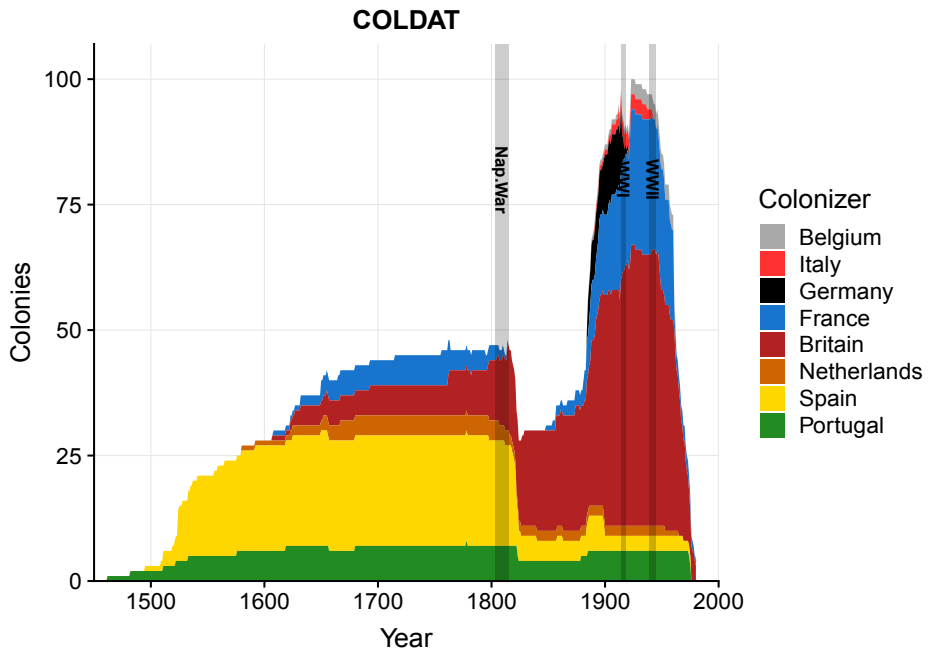
Table 5.

Former European Colonies by Colonial Power (Based on COLDAT, Mean Aggregation).

Empire	Colonies
Belgium (3)	Burundi (1922-1962), Congo - Kinshasa (1885-1960), Rwanda (1919-1962)
Britain (70)	Antigua & Barbuda (1632-1974), Australia (1798-1904), Bahamas (1738-1973), Bahrain (1861-1971), Bangladesh (1796-1947), Barbados (1627-1966), Belize (1798-1980), Bhutan (1888-1949), Botswana (1887-1966), Brunei (1888-1984), Cameroon (1922-1961), Canada (1737-1877), Cyprus (1887-1960), Dominica (1805-1972), Egypt (1882-1926), Equatorial Guinea (1827-1857), Eritrea (1941-1951), Fiji (1872-1970), Gambia (1852-1965), Ghana (1874-1957), Grenada (1762-1970), Guyana (1814-1966), India (1803-1947), Iraq (1917-1932), Israel (1920-1948), Jamaica (1655-1962), Jordan (1923-1946), Kenya (1891-1963), Kiribati (1892-1979), Kuwait (1914-1961), Lesotho (1874-1966), Libya (1946-1952), Malawi (1891-1964), Malaysia (1847-1958), Maldives (1887-1965), Marshall Islands (1886-1885), Mauritius (1810-1968), Myanmar (Burma) (1879-1948), Namibia (1915-1920), Nauru (1914-1944), New Zealand (1840-1909), Nigeria (1874-1960), Pakistan (1830-1947), Papua New Guinea (1884-1912), Qatar (1914-1970), Samoa (1914-1920), Senegal (1693-1817), Seychelles (1814-1976), Sierra Leone (1838-1961), Singapore (1821-1961), Solomon Islands (1893-1978), Somalia (1885-1956), South Africa (1801-1912), Sri Lanka (1797-1948), St. Kitts & Nevis (1623-1975), St. Lucia (1814-1973), St. Vincent & Grenadines (1762-1974), Sudan (1898-1956), Swaziland (1897-1968), Tanzania (1912-1961), Tonga (1900-1970), Trinidad & Tobago (1797-1962), Tuvalu (1892-1978), Uganda (1892-1962), United Arab Emirates (1892-1971), United States (1607-1781), Vanuatu (1906-1980), Yemen (1879-1967), Zambia (1900-1964), Zimbabwe (1900-1975)
France (34)	Algeria (1836-1962), Benin (1873-1960), Burkina Faso (1895-1960), Cambodia (1868-1952), Cameroon (1922-1960), Canada (1608-1763), Central African Republic (1898-1960), Chad (1899-1960), Comoros (1878-1975), Congo - Brazzaville (1881-1960), Côte d'Ivoire (1869-1960), Djibouti (1869-1977), Dominica (1632-1763), Gabon (1855-1960), Grenada (1650-1762), Guinea (1858-1958), Haiti (1665-1804), Laos (1892-1952), Lebanon (1921-1946), Madagascar (1890-1960), Mali (1897-1960), Mauritania (1882-1960), Mauritius (1715-1810), Morocco (1912-1956), Niger (1910-1960), Senegal (1770-1960), Seychelles (1756-1794), St. Lucia (1650-1814), Syria (1922-1944), Thailand (1867-1941), Togo (1922-1960), Tunisia (1877-1956), Vanuatu (1887-1980), Vietnam (1869-1951)
Germany (12)	Burundi (1894-1922), Cameroon (1884-1915), Ghana (1884-1916), Marshall Islands (1885-1914), Namibia (1882-1914), Nauru (1888-1914), Palau (1885-1914), Rwanda (1894-1915), Samoa (1900-1914), Solomon Islands (1885-1918), Tanzania (1888-1916), Togo (1884-1914)
Italy (3)	Eritrea (1885-1940), Libya (1912-1942), Somalia (1905-1936)
Netherlands (4)	Guyana (1580-1815), Indonesia (1621-1952), South Africa (1652-1806), Suriname (1642-1964)
Portugal (12)	Angola (1576-1974), Brazil (1533-1822), Cape Verde (1462-1975), Equatorial Guinea (1778-1778), Ghana (1476-1823), Guinea-Bissau (1879-1973), Malaysia (1511-1641), Mozambique (1695-1974), São Tomé & Príncipe (1522-1975), Sri Lanka (1619-1656), Timor-Leste (1642-1975), Uruguay (1680-1822)
Spain (27)	Argentina (1565-1818), Belize (1524-1862), Bolivia (1545-1825), Chile (1541-1817), Colombia (1518-1819), Costa Rica (1516-1821), Cuba (1511-1898), Dominican Republic (1493-1839), Ecuador (1534-1823), El Salvador (1525-1821), Equatorial Guinea (1858-1968), Guatemala (1524-1821), Honduras (1524-1821), Jamaica (1509-1655), Mauritania (1884-1975), Mexico (1520-1821), Micronesia (Federated States of) (1650-1899), Morocco (1884-1975), Nicaragua (1523-1821), Palau (1886-1899), Panama (1514-1820), Paraguay (1537-1811), Peru (1532-1823), Philippines (1565-1898), Trinidad & Tobago (1592-1797), Uruguay (1590-1823), Venezuela (1534-1820)

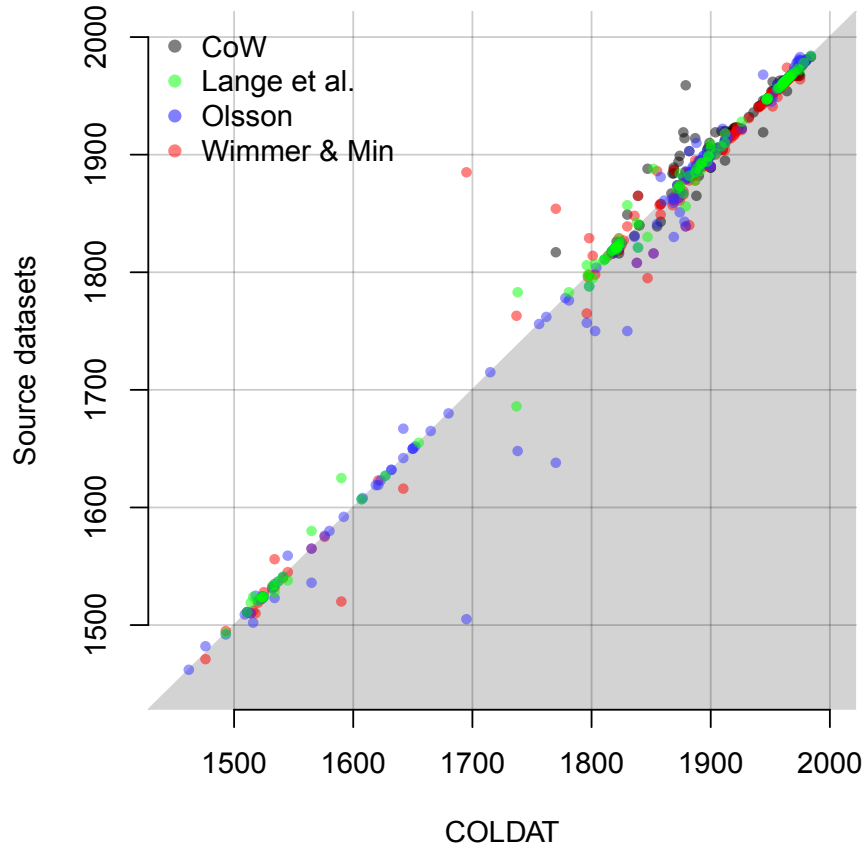
Countries with multiple colonizers: Belize (2), Burundi (2), Cameroon (3), Canada (2), Dominica (2), Equatorial Guinea (3), Eritrea (2), Ghana (3), Grenada (2), Guyana (2), Jamaica (2), Libya (2), Malaysia (2), Marshall Islands (2), Mauritania (2), Mauritius (2), Morocco (2), Namibia (2), Nauru (2), Palau (2), Rwanda (2), Samoa (2), Senegal (2), Seychelles (2), Solomon Islands (2), Somalia (2), South Africa (2), Sri Lanka (2), St. Lucia (2), Tanzania (2), Togo (2), Trinidad & Tobago (2), Uruguay (2), Vanuatu (2).

Figure 4.
Timeline of European Colonial Empires (Mean Aggregation).



Note: Based on COLDAT, mean aggregation. Territories are counted multiply in the case of simultaneous colonizations (e.g Cameroon from 1922-1960, colonized by Britain and France, is counted twice). Nap. War = Napoleonic Wars.

Figure 5.
Scatterplot of COLDAT (Mean Aggregation) and Source Dates.



Note: The indicated COLDAT dates refer to mean aggregates. Only sources dates used as input to COLDAT are displayed. The correlation coefficient and mean absolute error with respect to the COLDAT dates are as follows: CoW (.979, 3.86), Lange (.998, 3.94) Olsson