

STATES WE PRETEND EXIST AND STATES WE IGNORE

Karen Ruth Adams

Associate Professor
Department of Political Science
University of Montana

Draft date December 2003

(Reformatted May 2009)

Abstract:

From 1816 to 1994, the Correlates of War (COW) data set includes 15 European and Middle Eastern non-state entities comprising a total of 91 entity years and lacks data for 99 empirically-sovereign European and Middle Eastern states comprising 3,153 state years -- 37% of the state years I have identified. In this article, I explore the selection biases in the COW data by comparing the states and state years in the COW data to those in the State Survival and Death (SSAD) data set, which includes all of the empirically-sovereign states in Europe and the Middle East from 1816 to 1994. Finding significant biases against small, young, dead, 19th century, and non-European states, I conclude that it is necessary to use the COW data with caution and work toward the development of data sets that include only -- and all of -- the empirically-sovereign states in particular times and places.

Acknowledgments:

For comments on earlier versions of this paper, I thank D. Scott Bennett, Doug Bond, Ken Menkhaus, John Mueller, and Christopher Muste. For statistical assistance, I thank Jake Bowers, and for research assistance, I thank Nikki Detraz, Katia Ivanova, Anita Leska, and Katie Weir.

Copyright 2003 by Karen Ruth Adams

In research on state survival and death in Europe and the Middle East from 1816 to 1994 (Adams, in progress), I have found that a number of non-state entities are included in the Correlates of War (COW), Polity, and other data sets used by international-relations and comparative scholars. I have also found that these data sets exclude a large number of the empirically-sovereign states that existed in Europe and Middle East between 1816 and 1994. COW, for example, includes data for 15 non-state entities comprising a total of 91 entity years and, more significantly, lacks data for 99 states comprising 3,153 state years -- 37% of all of the state years I have identified in Europe and the Middle East from 1816 to 1994. Although most of the missing states fall beneath the COW population cutoff of 500,000 people, a number of states with populations of less than 500,000 are included in the COW data. Moreover, COW excludes a number of Middle Eastern and Eastern European states with populations over that benchmark.

In this paper, I explore the selection biases in the COW data and assess their significance. I focus on the COW data for two reasons. First, because the COW project has one of the longest time horizons and the most extensive collection of data (on everything from inter-state and civil wars to alliances, militarized interstate disputes, and territorial change), it is the most widely used data collection in the field.¹ Second, researchers often use COW system membership as a filter when using Polity and other data, as well as when they collect new data.

I proceed as follows. First, I discuss the population and juridical sovereignty rules COW uses to define its universe of states, and I speculate about the biases these criteria are likely to introduce into the data. Then I propose alternative criteria for inclusion and discuss my alternative State Survival and Death Data (SSAD) set, which includes all of the empirically-sovereign states in Europe and the Middle East from 1816 to 1994. Next, I compare the states and state years in the COW data to those in SSAD. I do so in two steps. First, I discuss the non-state entities included in the COW data, then I enumerate and examine the states missing from that data set. In discussing both of these problems, I examine the biases in

¹For a description of the COW data, a list of the COW data sets, and a bibliography of works using COW

the COW data, consider their implications for the validity of past empirical findings, and suggest ways that scholars can test for and avoid the effects of these biases until the COW and other data projects address the problem so that scholars can continue to use their data with confidence.

To summarize, I find that the non-state entities included in the COW data are primarily states that were occupied by the Soviet Union after World War II, erstwhile great powers, wartime allies of Britain and France, or Middle Eastern states occupied by Britain and France. I also find that the state years missing from the data are disproportionately among small, young, dead, 19th century, and non-European states. Thus I conclude that it is necessary to use the COW and other data sets with caution and work towards the development of international and comparative data sets that include only -- and all of -- the empirically-sovereign states in particular times and places.

Defining the Universe of States

Although most political scientists define the state in Weberian terms -- as “a human community that (successfully) claims the monopoly of the legitimate use of physical force within a given territory” (1958, p. 78) -- COW and the other data sets most often used by international and comparative scholars do not employ Weberian indicators such as military control of at least some territory or foreign and domestic policy autonomy in coding the states that have existed over the past centuries. Instead of using empirical, Weberian criteria, they use juridical sovereignty -- the international legal condition of diplomatic recognition by other states -- as their indicator of statehood. Moreover, they typically contain data not on all states in the international system, but on states of a certain size. The result is a disconnect between the data scholars use, which is about states that exist on paper, and the questions they are trying to answer, which are about the policies that real states adopt and implement.

data, go to <http://www.umich.edu/%7Ecowproj/>.

COW Criteria for System Membership

As mentioned, COW "system membership" is the criterion most often used for inclusion in international relations data sets.² To be considered a COW system member, an entity must have a population of at least 500,000 and "exercise a fair degree of sovereignty and independence," indicated from 1816 to 1920 by diplomatic recognition by Britain and France, and after 1920 by membership in the League of Nations or United Nations, or by diplomatic recognition by two major powers (Singer and Small 1972, pp. 20-21).

The population and juridical sovereignty rules are likely to introduce several biases into the COW data. To begin with, the population rule is obviously biased against small states. It also seems likely to be biased against historical states -- that is, states that are not alive today, especially states that lived and died in the 19th century -- for a state of 500,000 in 1816 would be larger relative to its peers than a state of the same population today.³ Thus a more reasonable population rule would be adjusted for population growth. Finally, if a significant number of small states or historical states are excluded by the population rule, dead states probably are too, because population is an important component of and a good proxy for national capability and because states that have been around longer have been exposed to more life-threatening situations.

The recognition rule also seems likely to create some biases in the data. First, because juridical recognition of states often lags behind their empirical creation and because juridical non-recognition of dead states is often delayed in the hopes that conquered or collapsed states will come back to life, this rule makes it likely that COW will lack early years of new states' lives (and thus that it will lack data on short-

²COW Interstate System data set (states.CSV). In this paper, I refer to the 1816 to 1997 data, available at <http://pss.la.psu.edu/intsys.html>. For updated data covering the period to 2002, see State System Membership v2002.1, available at <http://cow2.la.psu.edu/>.

³In 1816, the average state had a population of 3,195,000. In 1994, the average was 16,909,000.

lived states that were not around long enough to be recognized) and that it will contain data for non-state entities.

Second, because the post-1920 criteria for recognition are easier to meet than those for 1816 to 1919, the recognition rule is likely to be biased against 19th century and non-European states, many of which did not survive, either because of national consolidation in Europe or European colonialism in the Middle East. In this respect, Fazal's more relaxed criteria of pre-1920 diplomatic recognition by either Britain or France is an improvement (Fazal 2001, p. 85). Yet no matter how recognition is coded, there will probably be a bias against 19th century states because improvements in transportation and communication technologies mean that the transaction costs of bilateral recognition decreased from 1816 to the end of the 20th century. Moreover, the late emergence of international organizations with (more or less) universal membership meant there was no institutional architecture for (or expectation of) multilateral recognition before 1920. Furthermore, it is problematic to use France as one of just two sovereignty-granting great powers from 1816 to 1919, when for years during that time it was exhausted after the Napoleonic and Franco-Prussian wars and thus cannot be expected to have expended much energy recognizing small or distant states. A better indicator of juridical sovereignty would be recognition by any two great powers in a given era. But this would assume that the phenomenon of diplomatic recognition is not a Western phenomenon, which is also problematic given that between 1816 and 1919, Russia, Japan, and China entered into half as many treaties as the United Kingdom, France, and Italy.⁴ Yet this does not mean that the states Russia, Japan, and China did not formally recognize did not exist.

These problems with the recognition rule suggest the perils of using juridical sovereignty criteria. Denying the existence of states not recognized by great powers is like saying that trees do not fall in the forest unless great powers are there to see them do so. Moreover, it plays into the histories that great powers expound for their own purposes. If scholars are willing to study the (Western European)

⁴Derived from Correlates of War Alliance 3.0 data set, 1816-2000, available at

diplomatic system -- not the international-political system -- the COW data may be fine. But that is neither how the data are usually used nor how findings based on the data are usually interpreted.

Alternative Criteria for Statehood

To study the historical incidence and causes of state survival and death, I have developed a State Survival and Death (SSAD) data set of all of the European and Middle Eastern states that were empirically sovereign from 1816 to 1994. In doing so, I have employed the Weberian definition of statehood and indicators of sovereignty consistent with that definition -- specifically, military control of at least some territory and the final word on foreign and domestic policy in the land. Thus I code states as being born on the day that:

1. the last occupying troops leave the country if their purpose is to run the country and/or its foreign policy, not simply assist the state in implementing its own policies.
2. the occupying state recognizes the sovereignty of the national government in some concrete way (*e.g.*, by signing a treaty relinquishing control of foreign, defense, and domestic policy and beginning to withdraw its troops; or by signing a treaty relinquishing control of policy and stipulating that any remaining troops are there at the discretion of the other state and for the purpose of assisting that state in implementing its own policies), or
3. a sovereign government (one that has control of at least some territory) is declared, formed or returned to power (*e.g.*, after a national uprising or liberation by other states). To be considered alive, states born in this way must
 - a. seek independence from their erstwhile overlords (not just control of policy within the current state, which makes for a civil war, not the birth of a new state),

- b. have at least some independent administrative apparatus and military capability, and
- c. survive at least one month

By contrast, I code states as dead on the day that:

1. national troops are completely overrun by or surrender to occupying forces or, under the threat of force, the national government voluntarily surrenders sovereignty over foreign and domestic policy in all of its territory to another state (conquest)
2. the national government voluntarily surrenders sovereignty over foreign and domestic policy in all of its territory to another state (union)
3. national troops are completely overrun by or surrender to the troops of a domestic rival or rivals or the national government surrenders sovereignty over foreign and domestic policy in all of its territory to a new government (or governments) within the territory (revolution or disintegration)
4. the national government loses control of all of its territory without a new state arising to take its place (collapse).

Of these means of state death -- conquest, union, revolution, disintegration, and collapse -- all but revolution alter the number of empirically-sovereign states in the international system. Thus the SSAD data provide a window on the non-state entities included in and the sovereign states excluded from the COW data.

States We Pretend Exist

By comparing the states and state years in the SSAD data (listed in the Appendix) to those in the COW data, I have identified a number of European and Middle Eastern non-state entities that are included

in COW from 1816-1994. As shown in Table 1, there are 15 such entities comprising a total of 91 entity years. As expected, most of the entities (11 or 75%) and entity years (53 or 58%) are European. Also as expected, most of the years (83 or 91%) are from the 20th century.

[Table 1 about here]

These entities were all far from sovereign. As noted in the Appendix, two (Lebanon from 1976 to 1994 and Albania from 1915 to 1919) had yet to recover from collapsing, while the other 14 were occupied by conquerors.

The particular non-state entities included in the COW data are interesting to consider. Of those in Europe that are included for more than one year, four (for a total of 37 or 41% of the extraneous years) were occupied by the Soviet Union after World War II. Soviet troops were not withdrawn from Bulgaria until 1947, and there were no treaties or other agreements limiting the role of Soviet troops in Hungary, Romania, or Poland until 1956 or 1957. Thus these entities did not have a monopoly of force over territory.⁵ Poland was nevertheless a member of the United Nations after 1945 (probably because the Polish resistance assisted the Allies, who set up the UN). But none of the other three entities became a UN member until 1955 or 1956. Thus, according to COW's coding rules, they should not have been included in the data. The COW project does not explain why it regards these entities as states. If it is because the Soviet Union, one of just two great powers at the time, claimed they were independent even as it occupied them, their inclusion illustrates the political nature and limitations of the juridical sovereignty rule.

⁵For example, "the chief Soviet representative of the Warsaw Treaty Organization in Hungary exercised day-to-day control of both the Soviet army and the Hungarian People's Army" (Burant 1989). Thus it was the Soviet Union, not Hungary, that monopolized force in the territory, as illustrated by the events of 1956, when the Soviets used military force to reverse the Hungarian government's decision to withdraw from the Warsaw Pact and pursue political liberalization.

The other three European non-state entities included in the COW data for more than one year are Portugal from 1816 to 1819, France from 1816 to 1817, and Serbia from 1916 to 1917. During these periods, Portugal was occupied by the British, and France was occupied by the Allies. Their inclusion indicates that erstwhile great powers receive special consideration in the COW data, especially France, whose recognition of other states is a key to their inclusion in the COW system.

Russett, Singer, and Small, who initiated the COW project, explain that they code Serbia as an independent state throughout World War I because it continued to mount “an effective fighting force of 100,000 or more troops” (1968, p. 934). Yet Serbia was able to do so only because the French occupied Corfu in January 1916 and, without the consent of the Greek government, landed Serbian troops there (Langer 1980, p. 955). Because it was not Serbia but France that controlled the territory occupied by the Serbs, Serbia was not empirically sovereign. Here again, being an ally of Britain and France has its privileges.

So, in a peculiar way, does being their target. Other than Lebanon, all three of the Middle Eastern non-state entities included in the data were occupied by Britain and France. Perhaps to maintain an aura of decency, these great powers found it useful to “recognize” the governments they installed. But this no more means these governments were empirically sovereign than does contemporary American recognition of the Karzai regime in Afghanistan, which controls no territory outside of Kabul and survives in that city thanks only to an international peacekeeping force dedicated to the day-to-day security of its leaders.

The inclusion of these entities in the data should be of concern to scholars interested in studying the actions and interactions of empirically-sovereign states, for two reasons. First, although the number of state years is not very large, if these entities were included in the SSAD data (which contains 8,507 state years), they would overstate the number of state years in Europe and the Middle East by 11%. Thus, even if the exclusion of these cases reflects no selection bias, their inclusion could affect the significance of variables in statistical models.

Second, because these non-state entities were not sovereign, they were unable to make decisions about whether to democratize, fight wars, and undertake the other actions that scholars use this data to study. Thus scholars who wish to study phenomena unrelated to the niceties and whims of diplomatic recognition should eliminate these entities from the universe of states they consider.

States We Ignore

Table 2 summarizes the state years I have found to be missing from the COW and other prominent data sets. Because at least 32% of the empirically-sovereign states in Europe and the Middle East are missing from these data sets, the problem of missing states is more significant than that of non-state entities. Biases against small, dead, 19th century, and non-European states are characteristic of this problem as well.

[Table 2 about here]

As summarized in Table 3, the particular state years missing from the Polity, CNTS, Gleditsch and Ward (1999), and Fazal (2001) data are highly correlated with those missing from COW. Thus, I continue to focus on the COW data. The state years missing from that data fall into two categories: first, those missing for states considered to be COW system members in other years, and second, those missing for states whose existence COW never acknowledges. After discussing the biases evident in each category, I explore the implications of these biases for international and comparative political research.

[Table 3 about here]

State Years Missing for States Considered to be COW System Members in Other Years

As shown in Table 4, COW lacks some years for 20 European and 10 Middle Eastern states it considers to be system members in other years. Together, these state years make up 913 (29%) of the 3,153 state years missing from COW in these regions.

[Table 4 about here]

As expected, most (525 or 58%) of the 913 missing years are in the 19th century (but this is a slimmer majority than one might anticipate) and most (906 or 99%) of them are from the beginning of the states' lives. Contrary to expectation, however, just 309 (34%) of the missing state years are in the Middle East. Yet, as explained below, this figure understates the extent of Eurocentrism the COW data.

Fifty-nine percent (537) of the state years missing among the states COW counts as system members in other periods belong to states with populations of less than 500,000 in the last year they are missing from the data.⁶ This is roughly consistent with COW's criterion of 500,000 in every year of inclusion. Yet 482 (90%) of these missing years among small states belong to three small European states (Monaco from 1861 to 1992, Liechtenstein from 1816 to 1989, and San Marino from 1816 to 1991) that COW treats as system members in the 1990s, despite their populations of less than 35,000 in 1993.

⁶I use this indicator of population because more recent population statistics are easier to obtain, and they reduce the temporal bias of using an unchanging 500,000 cut-off for the last two centuries. To code, I relied on COW, CNTS, *The Statesman's Yearbook*, *Almanach de Gotha*, *The World Almanac*, McEvedy and Jones 1978, *Bulletin de la Commission Centrale de Statistique* 1866, and Jan Lahmeyer's population statistics website.

The other 41% (375) of missing years among sometime COW system members are among 23 states with populations over 500,000 in the last year of their exclusion from the data. Interestingly, although most (14) of these states are European, Middle Eastern states account for most of the missing state years. Moreover, all but one of the 309 missing Middle Eastern state years among sometime COW system members are in states with populations over the COW benchmark (the exception is Syria in 1945, for which I lack population data). Thus the Middle Eastern states whose missing years fall into this first category are larger than their European counterparts and are excluded for more years.

The seven large Middle Eastern states missing from the COW data for more than one year are (in descending order of population size in the last year they are missing from the COW data): Persia from 1816 to 1854 (1854 population between that of Sweden and Sicily), Egypt from 1816 to 1854 (1854 population akin to Belgium's), North Yemen from 1918 to 1925 (1925 population slightly larger than Switzerland's), Morocco from 1816 to 1846 (1846 population slightly smaller than the Papal States'), the Saudis from 1902 to 1926 (1926 population between Lithuania's and Norway's), Tunisia from 1816 to 1824 (1824 population larger than Hesse-Darmstadt's), and Oman from 1816 to 1970 (1970 population between that of Cypress and Kuwait).

Presumably, these Middle Eastern states are excluded from the COW system because they do not meet the criteria of recognition by both Britain and France. Yet most of them -- Persia from 1816 to 1854, Egypt from 1840 to 1855, Morocco from 1816 to 1847, Tunisia from 1816 to 1825, and Oman from 1839 to 1970 -- do meet Fazal's more relaxed juridical sovereignty criteria of recognition by either Britain or France for at least part of the periods in question (Fazal 2001, p. 88). Moreover, despite significant interactions with and pressures from European states, each of these states was empirically sovereign:

-- During the years Persia is missing from the COW data (1816 to 1854), it held its own in a war with the Turks (1821 to 1823) and resisted British efforts to modernize the Persian military (1838) (Langer 1980, pp. 894-5).

- In 1805, Egypt drove out the erstwhile Turkish governor and reclaimed its independence.

During the missing years (1816 to 1854), it had a powerful army that conquered territory from Sudan to the Persian Gulf (Langer 1980, pp. 865-867).
- At the beginning of North Yemen's missing years (1918 to 1925), the state gained its independence from the Ottomans. During the interwar years, it developed a consolidated administrative structure and worked to dislodge the British from Southern Yemen and the Saudis from Asir [*Encyclopedia Britannica*, "Yahya (Mahmud al-Mutawakkil)" and "Arabia: Yemen, The Age of Imperialism"].
- During Morocco's missing years (1816 to 1846), the state was ruled by a sultan who had his own army and successfully resisted incorporation into the Ottoman and French empires (*Encyclopedia Britannica*, "North Africa: Morocco, History").
- The Saudis were also empirically sovereign during the years they are missing from the COW data (1902 to 1926). In 1902, more than ten years after the conquest of their lands by the Rashidis, the Saudis retook Riyadh. By 1904, they had defeated both the Rashidis and the Ottomans. After that, they maintained "the fiction of Ottoman overlordship" in hopes of balancing British power while extending their holdings on the Arabian peninsula (*Encyclopedia Britannica*, "Arabia: Saudi Arabia, Ibn Saud and the third Saudi state").
- Tunisia (missing from 1816 to 1824) was also an Ottoman province on paper. But until the establishment of the French protectorate in 1871, it successfully maintained its sovereign independence (*Encyclopedia Britannica*, "North Africa: Ottoman Rule in the Maghrib").
- Finally, Oman (missing from 1816 to 1970) is one of the longest-lived, empirically sovereign states in Europe and the Middle East. Although not recognized by COW until 1971, the Al Bu Sa-id dynasty, which has ruled Oman since 1749, was an important international actor in the Middle East and North Africa in the early to mid-1800s (*Encyclopedia*

Britannica, "Oman, history of: Restoration of Omani Rule" and "Al Bu Sa'id Dynasty"; see also Metz 1993, "Independence").

Thus, although these Middle Eastern states are not in the COW data, they were certainly part of the international-political system.

So were the 14 European states with populations over 500,000. But most of these states are excluded from the COW system for just one or two years. The three European states excluded for more than five years are illuminating:

- The Italian state of Parma, missing from COW from 1816 to 1850, is excluded from system membership because it did not meet COW's population threshold until 1850. Yet Parma's population in 1816 was more than ten times larger than the populations of Monaco, Liechtenstein, and San Marino in 1993, when they are considered to be part of the COW system.
- Serbia is excluded from 1867 to 1877, until its independence was recognized by Britain and France in the Treaty of Berlin (1878). Yet, after 1867, no Ottoman troops remained in Serbia. Moreover, although on paper Serbia remained under the "collective guarantee" of several European great powers (per the 1856 Treaty of Paris), those states never occupied Serbia or commandeered its foreign or domestic policy. Instead, under Prince Michael and the regents of Prince Milan, Serbia entered into treaties with Greece, revised its constitution, and declared war on Turkey (Langer 1980, p. 761).
- Greece is missing from the COW system from 1822 to 1827. Yet as territories were liberated during the war of independence from the Ottomans, Greek administrative structures were established. In 1822, these structures were codified by a constitution. Although there was considerable tension among the various Greek factions from 1822 on, there was clearly a Greek state prior to 1827, when Britain, Russia, and France pledged in treaties to mediate

between the Greeks and Ottomans, thus formally recognizing -- but hardly creating -- the new Greek state (*Encyclopedia Britannica*, "Greece: Factionalism in the emerging state").

A kind of Eurocentrism, specifically a Western European bias, is evident even in the exclusion of these large European states, which were all in erstwhile Hapsburg and Ottoman lands. Here again, the problem of according juridical recognition only to states recognized by Britain and France is clear.

This examination of the first category of states missing from the COW data -- states that COW considers to be system members in just some of the years they were empirically sovereign -- reveals several troubling aspects of COW's coding. First, the COW population rule is invoked in some cases (*e.g.*, Parma) but not in others (Monaco, Liechtenstein, and San Marino). Second, the number of excluded state years (913) is high relative to the number of state years in the COW data (5,354). Third, because just 58% of the missing years are from the 19th century, coping with the selection bias in the COW data may require more than eschewing analysis prior to 1900. Fourth, because the population rule is responsible for 59% (537) of the missing 913 years, the juridical recognition rule is presumably responsible for the others. Thus developing a complete accounting of the states in the international-political system (one that does not privilege the large over the small, the European over the non-European, or the 20th century over the 19th century) seems to require abandoning both of these rules and doing the hard work of coding empirical sovereignty.

State Years Missing from COW Among States Never Considered to be COW System Members

The second and larger category of European and Middle Eastern state years missing from COW is the result of COW's non-recognition of any of the state years among 69 states that were empirically sovereign from 1816 to 1994. These states, which are listed in Table 5, account for 2240 (71%) of the 3,153 state years missing from COW's European and Middle Eastern data.

[Table 5 about here]

In this category, unlike the first one, the vast majority (2,129 or 95%) of the missing 2,241 years are from the 19th century. Like their counterparts in the first category, however, most (47 or 68%) of the missing states are European. But here again, a comparison of population size and number of missing years indicates that the number of missing states fails to convey the Eurocentrism caused by COW's juridical sovereignty criteria.

Among the European states in this category are 29 small states (those that meet my more reasonable population cutoff of less than 500,000 in the last year of their lives), five large states, and 13 states for which I have no population data. The small states contribute 1343 (89%) of European missing years, while the large states contribute just 51 (3%) of them. By contrast, among the Middle Eastern states that are completely missing from the COW data are seven small states, ten large states, and five states for which I lack population data. Small states contribute just 404 (55%) of the missing Middle Eastern years, while large ones contribute 156 (21%) of them. Thus just 3% of the European states that meet my more reasonable population criterion are excluded from COW system membership, while 21% of Middle Eastern states meeting this criterion are excluded -- presumably because they were not recognized by both Britain and France.

This bias against Middle Eastern states is troubling for several reasons. First, Middle Eastern states are not disproportionately excluded from the COW data simply because they were "flashes in the pan." Although the average life spans of European and Middle Eastern states that are completely excluded from the COW data are almost identical (32.04 and 33.36 years per state, respectively), the average life spans of the large states in each region are very different (1.09 and 7.09 years per large state).⁷ On

⁷To calculate average life spans, I divided the number of missing state years by the number of missing states.

average, the Middle Eastern states endured more than six times longer than their European counterparts.

Second, the non-recognition of these states was affected by technological and political variables that should be of interest to political scientists. Ninety percent (140) of the 156 missing years among large Middle Eastern states are from the 19th century, when transportation and communication technologies were less conducive to bilateral diplomatic recognition than they were in the 20th century, and before international organizations provided an institutional architecture for multilateral recognition. Moreover, 38% (59) of the missing years among these states occurred between the Treaty of Paris (1856), in which the European great powers pledged to respect the territorial integrity of the Ottoman Empire (Langer 1980, p. 774), and the end of World War I, when the European great powers defeated Turkey and began to redraw the Middle Eastern map.⁸

Third, and most important, each of these states was empirically sovereign. Among the five largest of the ten large Middle Eastern states that are completely missing from the COW data were two states in Sudan and three in Saudi Arabia.⁹ The first Sudanese state (1816 to 1821) was a kingdom without regional rivals until the Egyptians conquered it in 1821 (*Encyclopedia Britannica*, "Sudan: The Funj"). The second Sudanese state emerged in 1882, when Mahdists wrested the city of Kordofan from Egyptian control. During its life (1882 to 1898), the Mahdists consolidated power into a Muslim state and conducted an expansionist foreign policy despite Egyptian resistance and punitive British and French expeditions. This state remained empirically-sovereign until 1898, when the British conquered it to forestall French control of the Nile (*Encyclopedia Britannica*, "Sudan, The Mahdiah"). In 1821, Sudan's population was between that of Sardinia (Italy) and the Netherlands. In 1898, it was more populous than Portugal.

⁸Despite these obstacles, two of these large Middle Eastern states -- Algeria from 1816 to 1830 and Sudan in 1886 -- do meet Fazal's more relaxed juridical sovereignty criteria of recognition by either the United Kingdom or France for at least part of the periods in question (Fazal 2001, p. 87).

⁹For details and documentation on the other missing states, refer to the Appendix.

The three 19th century Saudi states that are completely missing from the COW data were also empirically sovereign. The first Saudi state was alive from 1816 (when my data begins and areas outside of a few garrison towns and off the pilgrim route were only nominally under Ottoman control) to 1818, when Egypt forced the surrender of Wahhabi leaders and posted troops to the main towns. The second Saudi state (1824 to 1838) emerged when Saudis recaptured Riyadh and expelled the Egyptian garrison. By 1830, other Egyptian garrisons had fallen, as had a number of nomadic tribes; in 1838, Egypt retook Riyadh, captured King Faysal, and installed a new leader. In 1841, the third Saudi state emerged when the Egyptian garrison in Riyadh was once again expelled, and Faysal resumed his reign. This state endured until 1891, when the Rashidis conquered Wahhabi lands and Saudi leaders fled to Kuwait (*Encyclopedia Britannica*, "Arabia; History Since the 7th Century, The Wahhabis"; "Arabia: Saudi Arabia, History"; and "Mulaydah, Battle of al-"). Throughout the 19th century, the Saudi population approximated that of Switzerland.

As in the first category, Eurocentrism is evident not just in the number of missing Middle Eastern states. It is also evident in the Western Eurocentrism of excluding large, Eastern European states from the COW system. The affected states, which were in erstwhile Hapsburg, Russian, or Ottoman lands, are Hungary in 1849, Azerbaijan from 1918 to 1920, Georgia from 1918 to 1921, Armenia from 1918 to 1922, and Montenegro from 1878 to 1915.

Eurocentrism is not the only worrisome bias evident in the states COW never acknowledges. The exclusion of states below the population threshold of 500,000 is also cause for concern because of the large number of states (36) and state years (1,747) eliminated by this coding rule.

Finally, as expected, there is a strong bias against dead states. As shown in Table 5, thirty-two of the European and Middle Eastern states never acknowledged by COW died by conquest, while 35 died by union. By contrast, just 39 of the COW system members in this region died by conquest from 1816 to 1994, while 12 died by union. Thus COW system members experienced just 45% of the conquests and

74% of the unions in Europe and the Middle East from 1816 to 1994.

Summary of Findings about States Missing from the COW Data

My findings about the states missing from the COW data are summarized in Table 6. Four points stand out. First, because the 3,153 missing state years reflect the lives of 99 states, the average state missing from the data is absent for 32 years. Thus, the problem is not just a year or two here and there. Second, 2,654 (84%) of the 3,153 missing years are from the 19th century. Thus there is a clear bias against historical states.

[Table 6 about here]

Third, 2,284 (72%) of the missing years are excluded from the COW data based on population size. Yet COW excludes a number of large Middle Eastern and Eastern European states and includes a number of states below its own population threshold. As shown in Table 7, for 13 states comprising 219 state years, COW does not follow its own population coding rules. Instead, it admits certain small states on an *ad hoc* basis. Of these state years, 167 (76%) are in Western Europe, while 52 (24%) are in the Middle East. Although these percentages are similar to the proportion of European and Middle Eastern state years in my SSAD data on empirically-sovereign states (73% and 26%, respectively), the addition of these small states is cause for concern, both because it is unsystematic and because all but one of the state years included among small states are in the 20th century, which reinforces the temporal bias of the COW data.

[Table 7 about here]

Finally, 582 (18%) of the 3,153 state years missing from the COW data are among states whose most recent population exceeded 500,000 persons. These cases are apparently excluded due to COW's juridical sovereignty coding rule. This rule falls harder on Middle Eastern states (464 or 80% of the 582 missing years) than European ones (118 or 20% of the missing years). Moreover, as discussed, large Eastern European states are more often excluded than their Western European cousins.

As shown in Table 8, the temporal effects of COW's population rule -- and its *ad hoc* application -- are evident in that, in 1816, fifty-one percent of states in Europe and the Middle East for which I have population data had populations of less than 500,000, yet none of these small states is acknowledged by COW. By contrast, in 1994, just 12% of states had populations of less than 500,000, and all of such states are COW system members. The temporal effects of COW's juridical sovereignty rule (and its Eurocentrism) are also evident in that just three empirically-sovereign states were excluded from the COW system in 1994,¹⁰ whereas 52 were excluded in 1816. Of those excluded in 1816, thirty-seven had populations below 500,000. Eight had populations over that benchmark. All were in the Middle East.

[Table 8 about here]

As shown in Table 9, my discussion of the biases in the COW data is confirmed by bivariate OLS regression of *State age*,¹¹ *Population*, *Area*, and *Year* against *COW system membership*. These regressions indicate that differences between the European and Middle Eastern state years included in and excluded from COW's data are statistically significant.¹² The states included in the COW data are significantly

¹⁰These states are the Serb Republic of Bosnia and Hercegovina, the Croatian Republic of Herzeg-Bosna, and Chechnya. At present, I lack population data for each of these states.

¹¹State age is a count of the number of years a state was alive from either 1816 or their date of birth (whichever came first) to 1994 or the year of their death (whichever is later).

¹²It is also a significant predictor of whether a state is in Europe or the Middle East, but because most of the state years I have identified in these regions are in Europe, more research is needed to determine whether the

more likely to be old, large, and alive in the 20th century than states in the larger universe of empirically-sovereign European and Middle Eastern states from 1816 to 1994. Thus, based on COW system membership alone, it is possible to predict several key attributes of states.

[Table 9 about here]

Research Implications of My Findings About Missing States

Without coding the missing state years for democracy, war participation, alliances, international trade, and other variables, it is difficult to say whether my findings about the selection bias in the COW data will undermine the findings of the many studies that use this data to evaluate the causes and effects of domestic- and international-political outcomes. Yet my qualitative and quantitative findings about COW's biases against small, young, dead, 19th century, and non-European states, combined with the large number of studies and data sets using the COW data, suggest that scholars must grapple with several issues in evaluating past research and developing new research designs.

First, because COW excludes so many state years among small states -- 27% of the state years in Europe and the Middle East from 1816 to 1994 -- scholars should either adopt lower population criteria or be explicit about the fact that their findings pertain only to the largest states in the international-political system. Yet until the juridical sovereignty rule is abandoned (admitting large, excluded Middle Eastern states to the data) and COW's neglect of its own population criterion is reversed (removing small European states), it will be difficult to claim the latter.

Second, the exclusion of small, 19th century states means that analyses of historical trends in state size may wrongly identify a peak in state size between 1876 and 1885 (Lake and O'Mahony 2002, p. 5). Moreover, the failure to acknowledge so many conquests -- and thus to remove small states from the

Eurocentric bias is statistically significant.

system and add their territories to those of their conquerors -- may mask an increase in state size during periods of war, especially wars in offense-dominant eras, when conquest is most likely (Adams 2003/04). Thus it is premature to conclude that there has been an “incredible rise and decline in average state size in modern history” (Lake and O’Mahony 2002, p. 38).

Third, because the COW data are biased against young and 19th century states, COW system membership may paint an erroneous picture of when contemporary European and Middle Eastern states emerged and consolidated. Thus comparative-political studies of nation-building that use COW data (as well as Polity and CNTS data, whose membership is highly correlated with COW’s) should be conducted and interpreted with care.

Fourth, because COW understates the number of historical Middle Eastern states, it creates the false impression that there were few states in the Middle East before European colonialism -- and thus that colonialism could be justified as a civilizing mission or that instability in post-colonial Middle Eastern states could be explained by political inexperience. The bias against Middle Eastern states is also troubling because the Middle Eastern states excluded from the data tend to be larger than European ones. For the sake of both historical accuracy and contemporary understanding, it is vital to redefine membership in the international-political system based on empirical, not juridical, sovereignty. Until that is done, scholars must remember (and remind their readers) that there is a good chance that their findings pertain only to the international and comparative politics of Western Europe or the global diplomacy of Britain and France in the 19th century and League of Nations and United Nations members in the 20th century.

Fifth, because COW system members experienced just 45% of the conquests and 74% of the unions in Europe and the Middle East from 1816 to 1994, the COW data create the false impression that states rarely die.¹³ Moreover, the biases in the COW data affect the significance of variables in models of

¹³International relations theorists of all stripes assume that state death is rare. Kapstein, for example, claims that “relatively few states go out of business” (1995, p. 771), while Wolfers asserts that “survival has only exceptionally been at stake, particularly for the major Powers” (1952, p. 488), and Wendt says that “[t]he death rate

conquest and union.¹⁴ Finally, because conquest entails war, the COW data understate the amount of war in the international system. This suggests that much of the literature on the on the historical incidence of war and peace may be wrong.

Even if the states excluded from the COW data did not reflect systematic biases, the large number of missing state years (37% of the state years among empirically sovereign states in Europe and the Middle East) would be cause for concern. After all, although (most) researchers know that states are (usually) included in COW if they pass the juridical sovereignty and population tests, the COW system membership data is often treated as an exhaustive list of the states that existed after 1816. That it is really just a sample of states is of great importance to scholars studying the incidence of war, peace, democratization, and other international and comparative political outcomes. And, of course, that is what most researchers are doing with the data.

When missing state years are coded and included in empirical analyses, the number of state years will rise, and the significance of previous findings at the global, regional, dyadic, and national levels of analysis may change. Scholars who use COW system membership to define their universe of cases should at least acknowledge this possibility. Better yet, they will evaluate the validity of their findings by comparing the results of models using global, COW data from 1816 to 1994 to models using COW data since the late 19th or early 20th century (to assess the significance of COW's temporal bias), and by

of states is almost nil" (1999, p. 279). Even Waltz, whose structural-realist theory of international politics suggests that the absence of an international sovereign makes state survival difficult, especially for weak states, concurs. According to him, "The death rate among states ... is remarkably low" (1979, p. 137-8).

¹⁴In testing multivariate event history models of conquest in Europe and the Middle East from 1816 to 1994 (Adams, article in progress), I have found that the significance of state age depends on whether I use the universe of states in COW (for which it is insignificant) or the universe of states in SSAD (for which it is significant at the .05 level). Similarly, in models of union (Adams, book manuscript in progress), great power decline just approaches significance when tested on COW system members but is significant at the .05 level when tested on all of the empirically-sovereign states in these Europe and the Middle East. Moreover, because I lack population data for 18 states that are completely missing from COW (12 of which died by conquest and 4 of which died by union), and because population size is a significant predictor of state vulnerability to conquest and union, the probabilities of conquest and union for small states are likely to be larger than models using COW data predict.

comparing the results of models using COW data for Europe and the Middle East from 1816 to 1994 to those using SSAD data (to evaluate the significance of COW's age, size, and Eurocentric biases).

Given the difficulty of obtaining data to code the political systems, war participation, and other attributes and activities of the large number of states missing from the COW system, one way to check the validity of past findings and ensure that of future ones is to generate a random sample of state years from the SSAD data on Europe and the Middle East, then concentrate on collecting data for the cases for which COW, Polity, CNTS, and other data sets lack data. This would reduce both data collection costs and the selection bias resulting from missing data.¹⁵ To further reduce data collection costs, COW, Polity, CNTS, and other data projects could combine efforts, collectively determining the size of the sample to be drawn and, once it is drawn, focusing their data collection and coding on that sample until missing data biases become less significant.

Ultimately, however, it will be necessary to inventory the empirically-sovereign states of the world in the last two centuries so that the regional sample can be compared to and supplanted by a random global sample and, later, a comprehensive data set of all of the empirically-sovereign states in the world. Bremer and Ghosn have recently argued that, in creating a data set of this nature, researchers should create an empirical sovereignty index (2003, p. 32). Whether this is a good idea will depend on how the index is defined and whether it is used to exclude states from the data. Because most researchers interested in outcomes affecting or affected by empirically-sovereign states will want to conduct their analyses on a universe of states defined according to Weberian criteria and because they will want to be able to compare their analyses to others, an effort should be made to have some threshold on the index be the tipping point for Weberian statehood. In my coding of empirically sovereign states in Europe and the Middle East, I have created such a threshold by stipulating that entities must have military control of at least some territory and the final word on foreign and domestic policy in the land. In other words, I use a three point

¹⁵On the under-utilization and benefits of sampling in international relations, see King and Zeng (2001a)

index in which only entities coded “3” (military control *and* foreign policy independence *and* domestic policy independence) are states.

As for the other selection criteria Bremer and Ghosn suggest (population, area, population density, political structure, cultural homogeneity, etc.), it is not clear that they really “add up” to a state. Certainly they are attributes of states, but is a small state less sovereign than a large one? Is a democracy less of a state than an autocracy? Is the United States less sovereign than Japan because the latter is more culturally homogeneous? Having such variables in the data would be useful because scholars could select the variables relevant to their research questions. But making state membership in a new, researcher-defined “system of states” contingent on coding criteria other than fundamental, empirical sovereignty indicators could bias analyses in new ways. Moreover, it would delay the day that scholars can evaluate the effects of COW’s selection bias on the significance of their findings.

Conclusion

Due to international anarchy, there is nothing to prevent states from participating in the “organized hypocrisy” of juridical sovereignty (Krasner 1999). As we have seen, they frequently do so. Withholding or reversing diplomatic recognition is a way for states to shun those with whom they are on the outs. Moreover, extending diplomatic recognition to non-state entities is a way to blow a legal smokescreen over conquests.

For scholars to participate in this hypocrisy is extremely problematic, both normatively and empirically. Yet we do so whenever we use juridical sovereignty to define the universe of states alive in particular times and places. Normatively, using juridical sovereignty as an indicator of state existence means that we accept the histories written by the victors, which turns our data into the European social

and King and Zeng (2001b).

register of states. Empirically, it means that our analyses and predictions are likely to be flawed because they are based on data that exclude a substantial number of small, young, dead, historical and non-European states. Thus traumatic episodes in international political history are forgotten and, perhaps, doomed to be repeated.

Despite years of assuming that conquest was uncommon and increasingly unlikely, perhaps because of the small number of states that “exit” the COW and Polity data, recent international-political history demonstrates that conquest remains a live issue. To fully understand this phenomenon, as well as the phenomenon of war it entails and the process of democratization it so often interrupts, we must stop pretending that dead states are alive, and we must stop ignoring states that are very much, if inconveniently, alive.

Table 1
European and Middle Eastern Non-State Entities Included in COW, 1816-1994¹⁶

<u>COW Code</u>	<u>Name</u>	<u>Years</u>	<u>No. of Years</u>
<u>EUROPE</u>			
310	Hungary	1945-1956	12
360	Romania	1945-1956	12
290	Poland	1945-1955	11
339	Albania	1915-1919	5
235	Portugal	1816-1819	4
220	France	1816-1817	2
345	Serbia	1916-1917	2
355	Bulgaria	1945-1946	2
269	Saxony	1867	1
315	Czechoslovakia	1969	1
329	Naples/Two Sicilies	1861	1
<u>MIDDLE EAST</u>			
660	Lebanon	1976-1994	19
630	Iran	1912-1920	9
645	Iraq	1942-1947	6
640	Turkey	1919-1922	4

¹⁶I use COW country codes as entity codes.

Table 2
European and Middle Eastern State Years Missing From Various Data Sets, 1816-1994¹⁷

Data Set	No. of European and Middle Eastern State Years in this Data Set	No. (Percentage) of European and Middle Eastern State Years Missing from this Data Set
State Survival and Death (SSAD)	8507	
Correlates of War (COW)	5354	3153 (37%)
Polity	5343	3164 (37%)
Cross-National Time Series (CNTS)	5066	3441 (40%)
Gleditsch and Ward	5750	2757 (32%)
Fazal	5546	2961 (35%)

¹⁷The figures in the second column for COW and the other, non-SSAD data sets are not the total numbers of state years in these regions, which are inflated by inclusion of non-state entities. Instead, they are the number of state years in the SSAD data (*i.e.*, the number of state years among empirically sovereign states) that appear in these data sets.

Table 3
**Correlations Between European and Middle Eastern State Years Missing from COW
and Those Missing from Other Data Sets, 1816-1994**

Data Set	Pearson Correlation (r)
COW - Polity	.817**
COW - CNTS	.823**
COW - Gleditsch	.870**
COW - Fazal	.926**

N = 8507

** Correlation is significant at the .01 level (two-tailed)

Table 4
European and Middle Eastern State Years Missing from COW
Among States Considered to be COW System Members in Other Years, 1816-1994¹⁸
(SSAD Data)

State Code	State Name	Missing Years	N Years Missing	Population Last Msg. Yr	N Years in 19th Century	N Years in 20th Century
<u>EUROPE</u>						
290.00	Poland	1918	1	25834		1
211.00	Belgium	1944	1	8361		1
310.00	Hungary	1918	1	7874		1
344.00	Croatia	1991	1	4784		1
343.00	Macedonia	1991-1992	2	2004		2
349.00	Slovenia	1991	1	1966		1
271.00	Wuerttemberg	1871	1	1797	1	
240.00	Hanover	1837	1	1686	1	
267.00	Baden	1871	1	1447	1	
345.00	Serbia	1867-1877	11	1384	11	
275.00	Hesse-Darmstadt (Hesse-Grand-Ducal)	1868-1871	4	823	4	
350.00	Greece	1822-1827	6	753	6	
339.00	Albania	1913	1	650		1
335.00	Parma	1816-1850	35	501	35	
280.00	Mecklenburg-Schwerin	1816-1842	27	495	27	
332.00	Modena	1816-1841	26	462	26	
212.00	Luxemburg	1918-1919	2	261		2
221.00	Monaco	1861-1992	132	31	40	92
223.00	Liechtenstein	1816-1989	174	28	85	89
331.00	San Marino	1816-1991	176	24	85	91
<u>MIDDLE EAST</u>						
651.00	Egypt	1936	1	16008		1
630.00	Persia	1816-1854	39	7500	39	
600.00	Morocco	1912	1	4940		1
651.00	Egypt	1816-1854	39	4463	39	
678.00	North Yemen/ Yemen Arab Republic	1918-1925	8	3950		8
600.00	Morocco	1816-1846	31	2875	31	
670.00	Saudi Arabia	1902-1926	25	2063		25
616.00	Tunisia	1816-1824	9	800	9	
698.00	Oman	1816-1970	155	654	85	70
652.00	Syria	1945	1			1

¹⁸For states in countries acknowledged by COW at some point from 1816 on, I use COW country codes as SSAD state codes. Populations are in thousands.

Table 5
European and Middle Eastern State Years Missing from COW
Among States Never Considered to be COW System Members, 1816-1994¹⁹
(SSAD Data)

State Code	State Name	Missing Years	N Years Missing	Population Last Msg. Yr	N Years in 19th Century	N Years in 20th Century	Conquest/ Union
	<u>EUROPE</u>						
310.00	Hungary	1849	1	10807	1		1
373.00	Azerbaijan	1918-1920	3	4615		3	1
372.00	Georgia	1918-1921	4	2799		4	1
371.00	Armenia	1918-1922	5	2159		5	2
345.10	Montenegro	1878-1915	38	516	23	15	1
277.90	Nassau	1816-1866	51	466	51		1
278.10	Oldenburg (formerly Holstein-Oldenburg)	1816-1867	52	316	52		2
276.80	Hamburg	1816-1867	52	305	52		2
276.60	Brunswick-Wolfenbittel	1816-1867	52	303	52		2
278.90	Saxe-Weimar	1816-1867	52	283	52		2
212.00	Luxemburg	1867-1914	48	269	34	14	1
276.30	Anhalt-Dessau/Anhalt	1816-1867	52	197	52		2
278.70	Saxe-Gotha	1816-1825	10	186	10		2
278.80	Saxe-Meiningen	1816-1867	52	180	52		2
278.60	Saxe-Coburg/Saxe-Coburg-Gotha	1816-1867	52	169	52		2
278.50	Saxe-Altenburg (formerly Saxe-Hildburgshausen)	1816-1867	52	141	52		2
345.10	Montenegro	1816-1861	46	130	46		1
277.50	Lippe-Detmold	1816-1867	52	111	52		2

¹⁹For states in countries acknowledged by COW at some point from 1816 on, I use COW country codes as SSAD state codes. For other states, I assign codes based on their immediate neighbors (reflected in the part of the code before the decimal point) and alphabetical order (based on other missing states proximate to the same state). Thus because Montenegro borders Serbia (COW code 345) its state code is 345.10. Populations are in thousands. Codes in the Conquest/Union column indicate states that died by conquest (1), died by union (2), or remained alive in 1994 (0).

Table 5 (con't)

State Code	State Name	Missing Years	N Years Missing	Population Last Msg. Yr	N Years in 19th Century	N Years in 20th Century	Conquest/ Union
	<u>EUROPE (con't)</u>						
276.50	Bremen	1816-1867	52	109	52		2
277.70	Mechlenburg-Strelitz	1816-1867	52	99	52		2
276.70	Frankfurt	1816-1866	51	89	51		1
278.30	Reuss-Schleiz (junior line)	1816-1867	52	88	52		2
279.20	Schwarzburg-Rudolstadt	1816-1867	52	75	52		2
279.30	Schwarzburg-Sondershausen	1816-1867	52	67	52		2
279.40	Waldeck-Pyrmont	1816-1867	52	57	52		2
277.60	Lubeck	1816-1867	52	48	52		2
276.10	Anhalt-Bernburg	1816-1853	38	45	38		2
278.20	Reuss-Greiz (senior line)	1816-1867	52	44	52		2
277.20	Hohenzollern-Siegmaringen	1816-1849	34	43	34		2
276.20	Anhalt-Cothen	1816-1847	32	40	32		2
279.10	Schaumburg-Lippe	1816-1867	52	31	52		2
276.90	Hesse-Homburg	1816-1866	51	27	51		2
290.00	Poland (Cracow)	1816-1827	12	26	12		1
277.10	Hohenzollern-Hechingen	1816-1849	34	21	34		2
359.00	Moldavian (Bessarabian) Republic	1918	1			1	2
369.10	Ukrainian National Republic	1918	1			1	1
230.10	Asturias	1937	1			1	1
337.20	Massa-Carrara	1816-1829	14		14		2
337.10	Lucca	1816-1847	32		32		2
372.20	Chechnya	1816-1859	44		44		1
369.10	Ukrainian National Republic	1918-1919	2			2	1
369.20	Western Ukrainian National Republic	1918-1919	2			2	1
372.40	Don Voisko	1918-1920	3			3	1
325.10	Fiume (Rijeka)	1920-1922	3			3	1
372.20	Chechnya	1991-1994	4			4	0

Table 5 (con't)

State Code	State Name	Missing Years	N Years Missing	Population Last Msg. Yr	N Years in 19th Century	N Years in 20th Century	Conquest/ Union
346.10	Serb Republic of Bosnia and Hercegovina	1992-1994	3			3	0
346.20	Croatian Republic of Herzeg-Bosna	1993-1994	2			2	2
<u>MIDDLE EAST</u>							
625.00	Sudan	1882-1898	17	5477	17		1
625.00	Sudan	1816-1821	6	5000	6		1
670.00	Saudis	1841-1891	51	2438	51		1
670.00	Saudis	1824-1838	15	2188	15		1
670.00	Saudis	1816-1818	3	2063	3		1
615.00	Algeria/Mascara	1832-1844	13	2027	13		1
615.00	Algeria	1816-1830	15	1500	15		1
670.10	Asir	1914-1920	7	1500		7	1
670.20	Hejaz	1916-1924	9	900		9	1
620.00	Tripolitania (Libya)	1816-1835	20	550	20		1
692.00	Bahrain	1816-1861	46	70	46		2
690.00	Kuwait	1816-1899	84	50	84		2
696.10	Sharjah/Al-Qawasim	1816-1892	77	35	77		2
694.00	Qatar	1913-1916	4	25		5	2
696.20	Abu Dhabi/Banu Yas	1816-1892	77	18	77		2
696.30	Dubai	1833-1892	60	18	60		2
694.00	Qatar	1816-1871	56	10	56		1
652.00	Syria	1920	1			1	1
670.10	Asir	1816-1872	57		57		1
678.00	Zaydi Imamate (San'a' or North Yemen)	1816-1872	57		57		1
670.30	Jabal Shammar/Rashidis	1869-1921	53		32	21	1
600.10	Confederal Republic of the Tribes of the Rif	1921-1926	6			6	1

Table 6
Summary of Findings about European and Middle Eastern States Missing from the COW System, 1816-1994²⁰

Region	N States	N State Years	N State Years 19th Century	N State Years Among States >500,000	N State Years Among States <500,000	N State Years Among States w/ No Pop.Data	N Conquests	N Unions
Europe - some yrs	20	604	322	67	537	0		
Europe - all yrs	47	1506	1443	51	1343	112	16	29
Total Europe	67	2110	1765	118	1880	112	16	29
ME - some yrs	10	309	203	308	0	1		
ME - all yrs	22	734	686	156	404	174	16	6
Total ME	32	1043	889	464	404	175	16	6
Total some years	30	913	525	375	537	1		
Total all years	69	2240	2129	207	1747	286	32	35
Grand Total	99	3153	2654	582	2284	287	32	35

²⁰In the region column, “Europe - some yrs” and “ME - some yrs” refer to states that COW acknowledges in some, but not all, years they are empirically sovereign. “Europe - all yrs” and “ME - all yrs” refer to states that are completely excluded from the COW system. Although four states in the “some year” category died during the years they are missing from the COW system, I do not include them in the conquest or union columns because researchers using the COW data would notice their deaths as they cease to be coded as COW system members in the year (or years) before their empirical deaths. These states are Baden, Wuerttemberg, and Hesse-Darmstadt, which all died by union in 1871 (last years in COW system 1870, 1870, and 1867 respectively) and Morocco, which died by conquest in 1912 (last year in COW system 1911).

Table 7
States with Populations of Less than 500,000 Included in the COW Data, 1986-1994²¹

State	Years	N Years Included	Population in First Yr. Incl.	Population in Last Yr. Incl.
EUROPE				
San Marino	1992-1994	3	20	24
Liechtenstein	1990-1994	5	28	31
Monaco	1993-1994	2	31	31
Andorra	1993-1994	2	62	65
Iceland	1944-1994	51	127	265
Luxemburg	1920-1940	21	261	295
Luxemburg	1944-1994	51	293	410
Malta	1964-1994	31	320	375
Modena	1849	1	499	499
MIDDLE EAST				
Qatar	1971-1994	24	122	475
Bahrein	1971-1989	19	222	489
Kuwait	1961-1965	5	322	471
United Arab Emirates	1971-1974	4	279	491

²¹Population data are in thousands.

Table 8
Comparison of States in SSAD and COW Data, 1816 and 1994

	SSAD DATA			COW DATA		
	N States	N < 500,000	% < 500,000	N States	N < 500,000	% < 500,000
1816	72	37	51%	20	0	0%
1994	68	8	12	65	8	12

Table 9
Bivariate OLS Regressions of COW System Membership
and State Age, Population, Area, and Year
(SSAD Data)

Dependent Variable	N	R²	Unstandardized B	Standard Error
Age	8,506	.013	9.134**	.873
Population	8,168	.096	16,567.105**	567.230
Area	5,750	.008	1,113.790**	160.132
Year	8,506	.244	58.420**	1.114

two-tailed tests of significance

* significant at the .05 level

** significant at the .01 level

Bibliography

- Adams, Karen Ruth. 2003/04. "Attack and Conquer? International Anarchy and the Offense-Defense-Deterrence Balance." *International Security*, Vol. 28, No. 3 (Winter).
- _____. Article in progress. "Conquering Myths: Testing Realist, Liberal, and Constructivist Arguments about State Vulnerability to Conquest."
- _____. Book manuscript in progress. *State Survival and Death: Domestic, International, and Technological Contexts*.
- Almanach de Gotha*. Various editions, 1822+. Gotha: Justus Perthes.
- Bremer, Stuart A. and Faten Ghosn. 2003. "Defining States: Reconsiderations and Recommendations." *Conflict Management and Peace Science*, Vol. 20, No. 1 (Spring), pp. 21-41.
- Bulletin de la Commission Centrale de Statistique*. 1866. Tome X. Brussels: Ministry of the Interior.
- Burant, Stephen R. 1989. "Soviet Influence: Loyalties and Control." *Hungary: A Country Study*. Washington, D.C.: Library of Congress Country Studies, GPO. Available at <http://lcweb2.loc.gov/frd/cs/cshome.html>.
- Encyclopedia Britannica*. <http://www.britannica.com>.
- Fazal, Tanisha. 2001. *Born to Lose and Doomed to Survive: State Death and Survival in the International System*. Stanford, CA: Stanford University. Ph.D. dissertation.
- Gleditsch, Kristian and Michael D. Ward. 1999. "A Revised List of Independent States Since the Congress of Vienna." *International Interactions*, Vol. 25, No. 4, pp. 393-413.
- Kapstein, Ethan. 1995. "Is Realism Dead? The Domestic Sources of International Politics." *International Organization* (Autumn).
- King, Gary and Langche Zeng. 2001a. "Logistic Regression in Rare Events Data." *Political Analysis*, Vol. 9, No. 2 (Spring), pp. 137-163.
- _____. 2001b. "Explaining Rare Events in International Relations," *International Organization*, Vol. 55, No. 3 (Summer), pp. 693-715.
- Krasner, Stephen D. 1999. *Sovereignty: Organized Hypocrisy*. Princeton, N.J.: Princeton University Press.
- Lake, David A. and Angela O'Mahony. 2004. "The Incredible Shrinking State: Explaining the Territorial Size of Countries." *Journal of Peace Research*, Vol. 48, No. 5, pp. 699-722
- Langer, William L. 1980. *An Encyclopedia of World History* 5th ed. Boston: Houghton Mifflin.
- Lahmeyer, Jan. Population statistics website. <http://www.library.uu.nl/wesp/populstat/populhome.html>.

- McEvedy, Colin and Richard Jones, *Atlas of World Population History* (New York: Facts on File, 1978).
- Metz, Helen Chapin. 1993. *Persian Gulf States: Oman*. Washington, D.C: Library of Congress Country Studies, USGPO.
- Russett, Bruce M., J. David Singer, and Melvin Small. 1968. "National Political Units in the Twentieth Century: A Standardized List." *American Political Science Review*, Vol. 62, No. 3 (September).
- Singer, J. David and Melvin Small. 1972. *The Wages of War, 1816-1965: A Statistical Handbook*. New York: John Wiley.
- Statesman's Yearbook*. Various editions, 1871+. New York: St. Martin's.
- Waltz, Kenneth N. 1959. *Man, the State, and War: A Theoretical Analysis*. New York: Columbia University Press.
- _____ 1979. *Theory of International Politics*. New York: McGraw Hill.
- Weber, Max. 1958. "Politics as a Vocation." H.H. Gerth and C. Wright Mills, eds. and trans. *From Max Weber: Essays on Sociology*. London: Oxford University Press.
- Wendt, Alexander. 1999. *Social Theory of International Politics*. Cambridge: Cambridge University Press.
- Wolfers, Arnold. 1952. "'National Security' as an Ambiguous Symbol." *Political Science Quarterly* 67:481-502.
- World Almanac* (later called the *World Almanac and Encyclopedia* and the *World Almanac and Book of Facts*). Various editions 1868-present. New York: Press Publishing Co.