Two tasks:

First, need the Gleditsch river data transformed to yearly observations for each state a (I also want a separate dyadic dataset pairing the states). Note that I want this to 2013. So if the “contiguity end year” variable is 1993, transform it to “2013”. Not all states’ river contiguity ended in 1993. Some ended earlier, such as Russia, which became independent in 1991 and their contiguity as the successor state to the USSR (365) changed as they lost former republics, such as Belarus, the Baltics, etc.

Note that the variables I need are “state a”, “state a name”, “shared river”, and “total length” for each observation WHERE THE “SHARED RIVER” VARIABLE EQUALS 1. Drop the observations with the value “0” for “shared river”. Next, create a count variable for each entry of state a. In other words, I want to know how many states the US shares river contiguity with for each year (we have rivers on the border between Canada and Mexico since 1920 and 1831, respectively). So there would be a count variable of 2 for the US from 1920 onwards. However, note that you will have to copy and paste the values for “state b” in the same entry. So there is an entry for the US’s shared border with Mexico, but not for Mexico’s shared border with the US. So this will double the number of observations. The first dataset will need to look like the entry below. The other one is dyadic and need separate entries for the US/Can and the US/Mex, while also needing entries for Can/US and Mex/US.

Basically, I want two datasets. One that has a single entry for each state in the system that has river contiguity in a given year. Then, another dataset that has an entry for state a and state b in the same observation, showing how much of a river border they have with the other state and their total km of river border lengths in a given year.

Here is an example of the country-year observation dataset with the total river length variable (and not the dyadic yearly observation)

Year Statea statename totalriverlength totalnumberrivercontig

1920 2 USA 2760 2

1921 2 USA 2760 2

------------------------------------------------

2013 2 USA 2760 2

1920 20 CAN 1275 1

1921 20 CAN 1275 1

--------------------------------------------------

2013 20 CAN 1275 1

1831 70 MEX 1485 1 (river contiguity only with the US in 1831)

1832 70 MEX 1485 1

--------------------------------------------------

1868 70 MEX 1791 2 (river contiguity with the US and Guatemala in 1868)

------------------------------------------------

1981 70 MEX 1935 3 (river contiguity with US, Guat., and Belize)

For the next task, I need a variable to indicate whether a state is an ally with another state in a given year or not. Actually, to be safe, I need two dummy variables. One for if they are in any alliance and another for if they are in a defensive alliance with one another. So a dummy variable of “0/1”. In the dataset, the unit of observation is the directed country dyad year. This means that there is an entry for the US (ccode1-2) and UK (ccode2-200) for the year 1816 and then a separate entry for the UK (ccode1-200) and the US (ccode2-2) for the year 1816.

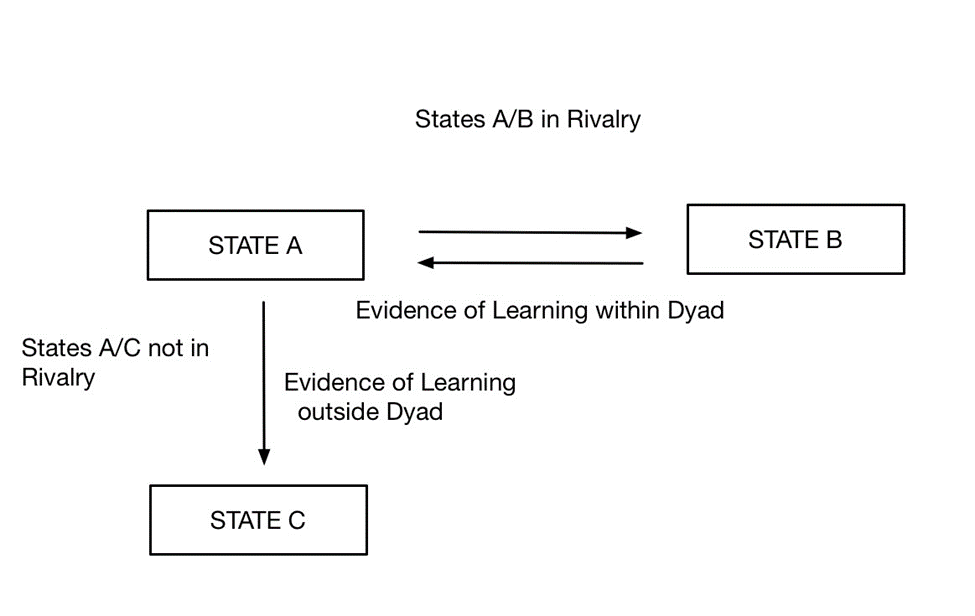
What I am doing in this paper is arguing that states that have a high level of conflict with another state (a rivalry) are more likely to use lower-levels of violence (MIDs) against states that they aren’t involved in a rivalry with. So, the US had a rivalry with the USSR from 1947 (ish)-1991. The theory is that the US and the USSR are more likely to have a MID with states they aren’t involved in a rivalry with. The problem is that my two variables of interest reflect whether or not the state has an ongoing rivalry at the time or had one in their past. What I need is a variable indicating whether ccode2 has a rivalry with the state that ccode1 has a rivalry with. As an example (I didn’t check to see if this actually happened, so bear with me), assume that the US (2) had a dispute with East Germany (265, the German Democ. Republic) in 1961. The questions I need answered is two-fold: did East Germany have 1.) a defensive alliance with any state the US had an ongoing rivalry with in 1961; 2.) any alliance with any state the US had an ongoing rivalry with in 1961? There are probably many ongoing rivalries the US had at that time, but the one that obviously springs to mind is the Soviet Union. So for this entry, there would be a “1” for the defensive alliance variable and a “1” for any alliance variable. Because East Germany was an ally of the Soviet Union in the Warsaw Pact (which was a defensive pact and, obviously, an alliance more generally). And the US had an ongoing rivalry with the Soviet Union at that time. It could be that there were multiple ongoing rivalries with the US at the time where the rivals were allies of East Germany. All I need to know is 0/1 if there was a rivalry going on at the time that the dispute occurred.

So in the data I provided to you, there is a yearly entry for each states’ alliances for each year. This data comes from the formal alliance data (<http://www.correlatesofwar.org/data-sets/formal-alliances>). Download the data. Use the alliance by directed yearly data. This has a separate entry for each alliance for each side. So the first alliance is one between the UK and Portugal (Id number 1). There are separate entries for both the UK and Portugal for each year the alliance is in effect. For building the dummy for if there is a defensive alliance, the “defense” variable is a dummy for if the alliance type is defensive between the states.

Next, I have included code for Thompson’s rivalry data. A dataset for this doesn’t actually exist because his approach is qualitative, rather than quantitative, if a state is in a rivalry with another. But the code does the trick. It shows if the first state in the code had a rivalry with the second state in the code and if so for what years. So, basically, to do this, I need to know if state a in the Thompson code had an ongoing rivalry (which I have already coded in the base dataset) and I need variables to know if state b in the base dataset had an alliance with state b in the Thompson data. Basically, is state a getting at state b (with whom it is in a rivalry) by going through state c (which is state b in the base dataset)

This should be two dummy variables: one for if state b in the base dataset had a defensive alliance with state a in the Thompson data and another for if there was any alliance between the two of them

So, my theory is:



The question I essentially want to answer is: does state c (represented as ccode 2 in base dataset) have either a defensive alliance or any kind of an alliance with state b (represented as the second entry in the Thompson data)? Because state A could be trying to “get at” state b by going through state c (if they are allies), therefore I need to control for it.