

Trade Networks Datasets

General. The trade datasets are based on two data collections of exports and imports among states over the period of 1870-2008. Users of the trade networks should cite the following sources for the data.

Barbieri, Katherine and Omar Keshk. 2012. Correlates of War Project Trade Data Set Codebook, Version 3.0. Online: <http://correlatesofwar.org>.

Barbieri, Katherine, Omar M. G. Keshk, and Brian Pollins. 2009. "TRADING DATA: Evaluating our Assumptions and Coding Rules." *Conflict Management and Peace Science*. 26(5): 471-491.

Note: This dataset is labeled hereafter BKP.

Gleditsch, Kristian S. 2002. Expanded Trade and GDP Data. *Journal of Conflict Resolution*, 46(5): 712-724. Data available at: <http://weber.ucsd.edu/~kgledits/exptradegdp.html>. **Note: This dataset is labeled hereafter Gleditsch.**

The codebooks for both datasets are provided in the current menu.

Combination of Two Datasets

The trade networks data combine the two datasets mentioned above. In general, the two dataset use similar sources, so their figures are similar. However, there are substantial differences on several matters. First, BKP trade data cover the period 1870-2007. Gleditsch data cover the 1946-2002 period. Second, these dataset differ in terms of how these two sources treat missing data. BKP treat missing data as missing. They have up to 47% of their observations as missing data. Gleditsch interpolated missing data. This dataset contained less than 16% missing values. Third, BKP did not report trade data for any dyad during the two World Wars (the period of 1914-1918 and 1939-1947). In order to generate a combined trade network dataset I followed these steps.

1. I merged the two datasets
2. I employed the BKP data whenever this dataset had nonmissing values
3. I inserted non-missing data from Gleditsch into the BKP variables whenever Gleditsch's values were nonmissing
4. For dyads that did not fight each other during the two world wars, I interpolated trade values such that a given value of X for that dyad was calculated by:

$$x_t = x_{t-1} + \frac{dif_w}{yr_w + 1}$$

Where for WWI $dif_w = x_{1919} - x_{1913}$ and for WWII $dif_w = x_{1948} - x_{1938}$ and yr_w are the number of gap years between observations.

5. I generated relative export data by dividing the exports from state i to state j by the former's GDP. For states that had missing GDP data, I replaced the relative export variable by dividing the exports going from i to j by the sum of exports going out of state i for that year. Relative share data are also available.
6. Diagonal values (internal trade) are calculated by $x_{ii} = (GDP_i - \sum_{j \neq i} x_{ij}) / GDP_i$, which is the non-exported share of the GDP. In the cases of missing GDP data I defined the diagonals as

zero. This may be a problem in some analyses, so we have to be careful when we employ self-tie measures with the trade data.

7. All remaining missing data are assigned a value of zero. This covers about 26.5% of the cases. The general justification for that is that trade levels were probably negligible for most of these dyads. Missing value codes are available.

Datasets on the Webpage

The data in the webpage consists of the following **directed** network files:

1. Dyadictrade2012.csv: This is a valued network dataset in dyadic format.
 - a. Year (this is the network ID variable).
 - b. State i (this is the row label).
 - c. State j (this is the column label).
 - d. Exports. x_{ij} = the dollar export going from i to j as a share of i 's GDP (or the share of the total exports of state i if GDP data are missing for that state). (**Note:** self-ties are the share of i 's GDP not going to exports).
2. Dyadictrademat.csv: this is a matrix format of the dyadictrade2012.csv file. It contains a series of matrices of size $n_t \times n_p$ with n_t being the number of states that existed in year t . Its format is as follows:
 - a. Year (network ID provided on the top-left cell of each matrix).
 - b. First row—Column Labels—State j in the dyadictrade2012.csv file.
 - c. First column—Row Labels—State i in the dyadictrade2012.csv file.
 - d. Matrix entries—equivalent to the x_{ij} entry in dyadictrade2012.csv
3. Dichdydtrade2012.csv: This is binary version of the dyadictrade2012.csv file, with values defined as 1 if there was any level of trade between two states and zero otherwise.
4. dichdydtrademat.csv: A matrix version of dichdydtrade2012.csv

Codebooks. The trade codebooks are are placed on the alliance Webpage.