

Correlates of War Project Trade Data Set Codebook

Version 2.01

Codebook dated September 30, 2008

Katherine Barbieri, (Katherine.barbieri@sc.edu)

Omar M. G. Keshk, (keshk.1@osu.edu)

Brian M. Pollins (pollins.1@osu.edu)

The following guide provides a description of the Correlates of War Project's Trade Data. The data set includes annual dyadic and national trade figures for states within the international system, as defined by the Correlates of War Project. The data set covers the period 1870-2006 and includes four data files: (1) dyadic trade statistics; (2) national trade statistics; (3) and (4) supplementary information about dyadic and national trade statistics. The data are distributed in a flat text format, with variables separated by commas. Given the .csv extension, the files will open in Microsoft Excel by default. Users should avoid using Excel to read the dyadic file, since the number of observations far exceeds the maximum number of rows that some versions of Excel will read. You should be able to import the text file into most statistical software packages. Missing values in the main data sets are assigned a value of -9. Variable names, procedures, source codes, and descriptive statistics are provided below for each data file.

These data are available at the COW website, <http://correlatesofwar.org>. Please report any errors and direct any questions to the authors. Please cite and refer to the data set as the following:

Barbieri, Katherine, Omar Keshk, and Brian Pollins. 2008. Correlates of War Project Trade Data Set Codebook, Version 2.01. Online: <http://correlatesofwar.org>.

This codebook has 6 sections:

- 1) Dyadic Trade Data
- 2) National (Monadic) Trade Data
- 3) Main References
- 4) Supplementary Information
- 5) Supplementary References
- 6) Version/Data History

Dyadic Trade Data

Our data base builds upon and integrates data from previous trade data projects. The majority of the post-WWII data were obtained from the International Monetary Fund's *Direction of Trade Statistics* (2007 CD-ROM Subscription and hard copy versions for various years). These data were supplemented with data from BKP 1.0 (Barbieri, Keshk, and Pollins, 2005), Barbieri's International Trade Dataset, Version 1.0 (Barbieri 2002), and data from the Republic of China (ROC), Bureau of Foreign Trade. The pre-WWII data were obtained from Barbieri.¹

The variables appearing in the file `dyadic_trade_2.01.csv` are described in Table 1.

Table 1. Variables in the Correlates of War Dyadic Trade Data Set file

(file `dyadic_trade_2.01.csv`)

Variable	Description
<code>ccode1</code>	Correlates of War Country Code for State A
<code>ccode2</code>	Correlates of War Country Code for State B
<code>year</code>	Observation year
<code>importer1</code>	Name of country A
<code>importer2</code>	Name of country B
<code>flow1</code>	Imports of Country A from Country B in current US millions of dollars
<code>flow2</code>	Imports of Country B from Country A in current US millions of dollars
<code>source1</code>	source of data for flow1 variable (see table below)
<code>source2</code>	source of data for flow2 variable (see table below)
<code>Bel_Lux_alt_flow1</code>	Original Belgium and Luxembourg trade values (see notes below)
<code>Bel_Lux_alt_flow2</code>	Original Belgium and Luxembourg trade values (see notes below)
<code>China_alt_flow1</code>	Original The People's Republic of China trade values (see notes below)
<code>China_alt_flow2</code>	Original The People's Republic of China trade values (see notes below)
<code>version</code>	Data Set Version Number

Procedures for obtaining and reporting data

We first generated all import and export trade data organized by pairs of economic entities, using the International Monetary Fund's *Direction of Trade Statistics* on CD-ROM (*DOTS*). Monthly Subscription for 2007-2008 and *Direction of Trade Statistics Historical Data* CD-ROM for 1948-1980. The historical data were merged with the post-1980 data (dropping any redundant 1980 values). Details about the IMF's trade statistics are available on line (see IMF, 1993, 2008).

For any given state, the IMF generates an entry of that state's imports from and exports to a list of its partners. The partner may be a state or non-state actor. The IMF data include, but are not limited to, sovereign states, non-state territories, regional trading groups, geographic regions, level of development grouping (e.g., developing states), and "the world." At this stage, for the dyadic data each observation consists of a given state matched with a given partner. We used the instances in which a given state was partnered with the world to obtain national import and

¹ Supplementary notes and data files are available for the pre-WWII era (section 4 of this document). Those data files contain trade in local currency, exchange rates, and source notes.

export values.

Next, we match the reports of the importer and exporter. For any directional flow of trade between two states (e.g., East to West flows or North to South flows), the importer and exporter should both report a given value. We then have two flows, Flow 1 and Flow 2, each with an importer and exporter report attached to that flow. We begin by matching the importer and exporter reports for Flow 1 and the importer and exporter reports for Flow 2.

The COW dyadic data set reports trade between sovereign states. The matrix of pairs of sovereign states was generated using EUGENE (Bennett and Stam, 2000, 2007). We compared the Correlates of War list of states with the IMF economy list (Correlates of War, 2005). The IMF uses multiple names, spellings, and mixed case combinations for the same country. This appears to be a product of which state is reporting the trade. It makes it extremely difficult to merge the COW state names with the IMF country names. We had to visually inspect many of the cases. In some cases, a country was divided after civil war, reunited, or transformed into a new entity. At times, the IMF retains all the representations of a given economic/political entity, but typically only assigns data to those years in which the state exists.

For example, the IMF includes the USSR in its data base for the period 1948-2006; the Russian Federation from 1980-2006, and Russia from 1980-2006. There are no data for the USSR after 1991 and there are no data for the Russian Federation and “Russia” until after 1991. The problem of multiple names is compounded when one state reports Russia to be its partner and another state lists the Russian Federation as its partner. Typically, Russia uses that term to refer to its own imports and exports to other states, but other states tend to use the term Russian Federation.

We identified similar problems with multiple or changed names for Yugoslavia and Serbia and Montenegro, Korea and South Korea, Germany and West Germany, and Yemen and Yemen Arab Republic. In each case, we isolated those countries and insured that the appropriate data were recorded for the appropriate entity in the appropriate year. We made sure our state list corresponded with the COW state membership list.

We generated a non-directed dyadic data set. Each combination of states has only one entry. In each case, two states report the same flow of goods in one direction. Thus, we have two states reporting Flow 1 and Flow 2. We have four trade values: Flow 1, Importer report (IR); Flow 1, Exporter Report (ER); Flow 2, Importer Report (IR); and Flow 2, Exporter Report (ER). The data are arranged, so the lower number in the COW country list appears in the first column when data for Flow 1 and Flow 2 are merged. For example, the US-Canada dyad in 1989 would have one row, with four trade flows. Dyadic trade consists of goods flowing in two directions--from Canada to the US and from the US to Canada. Flow1_IR would provide the value of US imports from Canada, as reported by the US; while flow1_ER would be Canadian exports to the US, as reported by Canada. Similarly, flow2_IR represents Canadian imports from the US, as reported by Canada; and flow2_ER is the value of US exports to Canada, as reported by the US.

The IMF contains many missing trade values. Once we exhausted the data from the importing state's reports, we substituted the missing import values with available figures provided by the exporting state. In other words, when the importer report for Flow 1 was missing, we relied upon the exporter report for Flow 1. Next, we followed the same procedure to replace observations with import values of zero.

After using the current IMF CD-ROMs, we relied upon trade data from earlier IMF tapes (1996) to fill in import values that were missing or coded as zero. Next, we used Barbieri's trade data set to fill in additional missing values. Barbieri relies upon both IMF and non-IMF sources for her trade data. Finally, there were a few special cases in the data set.

First, the IMF does not report trade data for Taiwan, so we acquired those data from multiple sources. For the years 1949-1988, we obtained data when available from United Nations *Yearbook of International Trade Statistics* (1951-1984), APEC Study Center City University of Hong Kong Data Bank (2004), and Republic of China's (ROC), Council for Economic Planning and Development (2002, 2004). For the years 1989-2006, data was obtained from the ROC's Bureau of Foreign Trade.

Second, the IMF reported one aggregate value for Belgium and Luxembourg's trade. This changed in 1996. For the pre-1996 values, we include the IMF's original values, plus disaggregated data based on the relative size of the GDP of each country. To disaggregate the Belgium-Luxembourg combined reports, we first obtained the nominal GDP for Belgium and Luxembourg, respectively (World Bank, 2005). We generated a ratio value of the smaller to higher GDP values (i.e., Luxembourg to Belgium). We multiplied Luxembourg's dyadic trade figure by this ratio to produce its trade values. For Belgium, we multiplied the total trade figures by one minus the ratio value above. The original figures are placed in separate columns in our data set (i.e., *Bel_Lux_alt_flow1* and *Bel_Lux_alt_flow2*).² While this is not an ideal solution for disaggregating the trade figures, we would recommend using the adjusted figure rather than the aggregate figure that combines Belgium and Luxembourg's trade.

We found that the trade figures for Macao and Hong Kong continued to be reported separately from Mainland China after these areas were unified in 1998. We combined these figures to produce our measures of China's dyadic trade. Once again, we include the original figures in separate columns in our data set (*China_alt_flow1* and *China_alt_flow2*). The average percentage difference between the aggregated figures for China, Macao, and Hong Kong and the disaggregated figures is approximately 27% for the period 1999-2006.

Finally, after following the procedures above, we coded any remaining missing data points as missing data. This is one way in which our data set differs from that of some scholars. Some scholars opt to assume missing data indicate zero trade or that the trade conforms to a given linear trend that could be filled in through interpolation or extrapolation techniques. We believe these are dangerous and faulty assumption (See Barbieri, Keshk, & Pollins 2007).

In Table 2, we provide a summary of the sources and procedures we used to produce each data point. The corresponding numbers appear in our data set. This gives scholars using the data set the opportunity to alter the decision rules and rely upon a different set of figures.

² Another strategy to adjust these figures would be to use available disaggregated data for the post 1996 period and generate a ratio for each dyad and apply those to the pre-1996 period.

Table 2. Sources, Procedures, and Codes

Source Codes	Sources and Procedures
1	Barbieri Version 1 (see Barbieri 2002, appendix A; also see “ Supplementary Information ” section of this codebook).
2	IMF import reports, in c.i.f. values (IMF, 2007)
3	Missing import value replaced with the exporter’s trade report, in f.o.b. values (IMF, 2007).
4	Zero import trade value replaced with the exporter’s trade report, in f.o.b. values (IMF, 2007).
5	Missing import value replaced with importer’s report, in c.i.f. values (IMF, 1992 tapes)
6	Zero import trade value replaced with importer’s report, in c.i.f. values (IMF, 1992 tapes)
7	Missing import value replaced with exporter’s report, in f.o.b. values (IMF, 1992 tapes)
8	Zero import value replaced with exporter’s report, in f.o.b. values (IMF, 1992 tapes)
9	Missing values replaced with Barbieri trade values
10	Zero values replaced with Barbieri trade values
11	Belgium-Luxembourg Data 1948-2006 (see notes below)
12	Taiwan Data from multiple sources (1952-1988) ³
13	Data obtained from Aggregating China, Hong Kong and Macao
-9	Missing

Table 3 provides a summary of the percentage of observations impacted by the different steps described above, as a share of all the possible dyadic observations for the years 1870-2006.

Table 3. Breakdown of Observations Affected

Source Codes	Sources and Procedures	Share of Observations Flow 1	Share of Observations Flow2
1	Barbieri Version 1 (see Barbieri 2002, appendix A; also see “ Supplementary Information ” section of this codebook).	19,389 (2.63%)	19,268 (2.62%)
2	IMF import reports, in c.i.f. values (IMF, 2007)	429,756 (58.40%)	415,586 (56.48%)
3	Missing import value replaced with the exporter’s trade report, in f.o.b. values (IMF, 2007).	34,564 (4.70%)	49,353 (6.71%)
4	Zero import trade value replaced with the exporter’s trade report, in f.o.b. values (IMF, 2007).	14,598 (1.98%)	20,787 (2.82%)
5	Missing import value replaced with importer’s report, in c.i.f. values (IMF, 1992 tapes)	1,443 (0.20%)	1,037 (0.14%)

³ Taiwan data for 1951-1969 are from United Nations; national data for 1971-1972 are from the APEC Study Center; data for 1973-1988 are from the ROC, Council of Economic Planning and Development; and data for 1989-2006 are from the ROC, Bureau of Foreign Trade.

6	Zero import trade value replaced with importer's report, in c.i.f. values (IMF, 1992 tapes)	4,403 (0.60%)	2,729 (0.37%)
7	Missing import value replaced with exporter's report, in f.o.b. values (IMF, 1992 tapes)	371 (0.05%)	418 (0.06%)
8	Zero import value replaced with exporter's report, in f.o.b. values (IMF, 1992 tapes)	2,915 (0.40%)	616 (0.08%)
9	Missing values replaced with Barbieri trade values	248 (0.03%)	384 (0.05%)
10	Zero values replaced with Barbieri trade values	1,110 (0.15%)	1,320 (0.18%)
11	Belgium-Luxembourg Data 1948-2006 (see notes below)	9,244 (1.26%)	2,946 (0.40%)
12	Taiwan Data from multiple sources (1952-1988)	4,453 (0.61%)	4,237 (0.58%)
13	Data obtained from Aggregating China, Hong Kong and Macao	1,524 (0.21%)	1,524 (0.21%)
-9	Missing	211,829 (28.79%)	215,642 (29.31%)

**Descriptive Statistics for the Dyadic Trade Data
(from Stata 9.2; missing values excluded)**

```
. sum ccode1 ccode1 year
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ccode1	735847	290.8787	205.7359	2	987
ccode2	735847	581.6965	220.317	20	990
year	735847	1976.672	27.17725	1870	2006

```
. sum flow1 if flow1!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
flow1	524018	134.5041	2155.169	0	315362.2

```
. sum flow2 if flow2!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
flow2	520205	123.733	2312.541	0	1198527

```
. sum source1 if source1!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
source1	524018	2.459456	1.780396	1	13

```
. sum source2 if source2!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
source2	520205	2.365948	1.465838	1	13

```
. sum bel_lux_alt_flow1 if bel_lux_alt_flow1!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
bel_lux_al~1	14192	269.9532	1636.336	0	32816.5

```
. sum bel_lux_alt_flow2 if bel_lux_alt_flow2!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
bel_lux_al~2	14194	283.9805	1692.964	0	29157.1

```
. sum china_alt_flow1 if china_alt_flow1!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
china_alt_~1	1470	3117.385	15866.98	0	305788

```
. sum china_alt_flow2 if china_alt_flow2!=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
china_alt_~2	1470	1783.779	7391.551	0	118444

National (Monadic) Trade Data

Table 4 below provides a list of the variables appearing in the in national_trade_2.0.csv data file.

Table 4. Variables in the COW National Trade File

(file national_trade_2.0.csv)

Variable	Description
ccode	Correlates of War Country Code
country	Name of country
year	Observation year
imports	Total Imports of Country in current US millions of dollars
exports	Total Exports of Country in current US millions of dollars
alt_imports	Original People's Republic of China total import values and Belgium/Luxembourg to0tal import values (see discussion above)
alt_exports	Original People's Republic of China total export values and Belgium/Luxembourg to0tal export values (see discussion above)
source1	source of data for imports variable (see table 2)
source2	source of data for exports variable (see table 2)
version	Data Set Version Number

Descriptive Statistics for the National Trade Data (from Stata 9.2; missing values excluded)

```
.sum ccode year
Variable |      Obs      Mean    Std. Dev.      Min      Max
-----+-----
ccode    |    12147    406.0749    256.6998         2     990
year     |    12147   1960.776     36.1416    1870    2006

.sum imports if imports~=-9
Variable |      Obs      Mean    Std. Dev.      Min      Max
-----+-----
imports  |    11608   12047.66   60332.99         .03  1919260

.sum exports if exports~=-9
Variable |      Obs      Mean    Std. Dev.      Min      Max
-----+-----
exports  |    11608   11651.63   52929.52         .01  1111050

.sum alt_imports if alt_imports~=-9
Variable |      Obs      Mean    Std. Dev.      Min      Max
-----+-----
alt_imports |      102   72676.02  126374.8    1942.8   791793

.sum alt_exports if alt_exports~=-9
Variable |      Obs      Mean    Std. Dev.      Min      Max
-----+-----
alt_exports |      102   77062.25  145430.5    1640.2  969284

.sum source1 if source1~=-9
Variable |      Obs      Mean    Std. Dev.      Min      Max
-----+-----
source1   |    11985    1.382061    .492394         1         3
```



```
.sum source2 if source2~=-9
```

Variable	Obs	Mean	Std. Dev.	Min	Max
source2	11985	1.384814	.4930454	1	3

References - Main

- APEC Study Center, City University of Hong Kong Data Bank. 2004.
<http://fbweb.cityu.edu.hk/hkapec/DataBank/Taiwan-Trade-Balance.pdf>
- Barbieri, Katherine. 2002. *The Liberal Illusion: Does Trade Promote Peace?* Ann Arbor, University of Michigan Press.
- Barbieri, Katherine and Omar Keshk. 2008. "Getting an Accurate Picture of Trading Relationships." Unpublished Manuscript.
- Barbieri, Katherine, Omar Keshk, and Brian M. Pollins. 2003. "The Correlates of War Trade Data Set: A Progress Report." Paper presented at the Annual Meeting of the Peace Science Society (International), Ann Arbor, Michigan, November 13-16.
- Bennett, D. Scott, and Allan Stam. 2000. "EUGene: A Conceptual Manual." *International Interactions* 26:179-204. Website: <http://eugenesoftware.org>.
- Bennett, D. Scott, and Allan Stam. 2007. EUGene: Expected Utility Generation and Data Management Program. Version 3.201, October 29. <http://www.eugenesoftware.org/>.
- Correlates of War Project. 2005. "State System Membership List, v2004.1." Online, <http://correlatesofwar.org>.
- International Monetary Fund. 1993. "A Guide to Direction of Trade Statistics" on CD-Rom and at <http://www.imf.org/external/pubs/cat/longres.cfm?sk=154.0>.
- International Monetary Fund. *Direction of Trade Statistics: Historical Database and Browser* on CD- ROM.
- International Monetary Fund. 2007. *Direction of Trade Statistics on CD-ROM*. Monthly Subscription.
- International Monetary Fund. 2008. "Data and statistics: Direction of Trade Statistics Consistency between partner country data." <http://www.imfstatistics.org/dot/DOTConsi.htm>
- Republic of China (ROC) Bureau of Foreign Trade Website. 2008. Data source, Directorate General of Customs, Ministry of Finance. <http://cus93.trade.gov.tw/english/FSCE/FSC0011E.ASP>
- Republic of China (ROC), 2002, 2004. Council for Economic Planning and Development, *Taiwan Statistical Data Book*.
<http://www.cepd.gov.tw/english/file/2004DataBook.pdf?CFID=87002&CFTOKEN=50209707>
- STATA 9.2. (2007) Statistics/Data Analysis. College Station, TX: StataCorp.
- United Nations, Statistical Office, Department of Economic Affairs. *Yearbook of International Trade Statistics*. New York: United Nations, 1951-1984.

World Bank. 2005. World Development Indicators, On-Line Version. World Bank, Washington DC.

Supplementary Information

Data in the COW trade data set identified as having a source of “Barbieri version 1” refers to Barbieri’s International Trade Database, Version 1.0, described in detail in Appendix A of Barbieri (2002). The following describes files and procedures relating to that original data set. In particular, the COW trade data includes two supplementary data files that contain more detailed information about the original sources of those data points, including the original data collected in local currency along with exchange rate information, sources, and additional notes, when relevant.

Two files (Microsoft Excel format) contain supplementary background material and original source information:

- 1) “dyadic_trade_supplement.xls” contains supplementary information on the dyadic trade data;
- 2) “national_trade_and_exchange_rate_supplement.xls” contains supplementary information on total national trade data, and exchange rate information.

Codes and variables contained in those supplementary files are as follows:

Dyadic Trade Supplementary file (dyadic_trade_supplement_2.01.xls)

Variable	Description
Year	Observation Year
Importer	Correlates of War Project’s State Abbreviation for Importer
Exporter	Correlates of War Project’s State Abbreviation for Exporter
Value	Value in reported currency
Currency	Currency associated with reported value
Exp	Exponential value for report
Source	Abbreviation of source listing, see notes for sources
Note	Elaboration when needed
Xrate	Exchange rate used for flow calculation
Flow1	Value in Millions of US dollars
Version	Data Set Version Number

National Trade/Exchange Rate Supplementary file

(national_trade_and_exchange_rate_supplement_2.01.csv)

Variable	Description
Year	Observation Year
St Abb	Correlates of War Project’s State Abbreviation for Importer
StNum	Correlates of War Project’s State Country Code
Tot(\$mil)	Total trade is US millions
TotIm	Total national imports in reported currency
TotEx	Total national exports in reported currency
Currency	Currency associated with reported value
E	Exponential value for report
XRate	Gurr rate supplemented by Barbieri calculations
GXR	Gurr exchange rate from Polity II
Bidrate	Bidwell Rate, see source listing
NewRt	Exchange rate used for flow calculation
NewExpRt	Exchange rate for exports, when different from import rate
KatRate	Barbieri calculated the rate
Revised	Elaboration when needed

Source	Abbreviation of source listing, see notes for sources
Note	Elaboration when needed
TotRimp	Total national imports in Millions of US dollars
TotRExp	Total national exports in millions of US dollars
TotTrade	Total Imports plus Exports in Millions of US dollars
Version	Data Set Version Number

Procedures and additional details of those original Barbieri International Trade Database files and the COW Trade Data supplementary data files are as follows.

OVERVIEW OF DATABASE PROJECT

Data were collected for all sovereign states within the interstate system, as defined by the Correlates of War (COW) Project for the period 1870-1992. Few observations are available for the years corresponding to World War I and World War II. The data set only includes reported statistics derived from trade reports; no estimation techniques were employed to replace missing values, as discussed below.

Missing data for exchange rates, cases including only partial information about dyadic flows or national totals, and reports containing unrealistic values are included as missing values do not appear in the final COW data files, but may be present in the supplementary files. Unrealistic values include those cases where dyadic trade exceeds the total trade of one of the nations in the dyad. This situation may arise from a number of factors that include erroneous trade reports, inconsistencies between partner reports, dramatic fluctuations in the exchange rates that are applied to the trade figures (e.g., national totals may be reported in USD, while dyadic figures were transformed with exchange rates that can vary widely from one year to the next).

COLLECTION PROCEDURES AND ENTERING DATA

Stuart Bremer designed a program to generate all possible combinations of dyads within the interstate system for a given year. This was used to construct annual databases to enter the trade data. In all cases, trade figures are reported only when they were available for the given year corresponding to the matrix; no interpolation or extrapolation techniques were used. For example, if China's trade is absent in 1899, but present for 1898 and 1900, Barbieri did not interpolate the figures to derive a figure for 1899, but counted that point as missing data. Her rationale for refraining from interpolation and extrapolation techniques resides in her belief that many historical reports already include cases where states or publishers have employed estimation techniques to replace missing values. Therefore, additional data manipulation would introduce further error, by assuming that trade conformed to a particular trend, when missing data may instead corresponds to periods in which real changes in trade patterns took place. In addition, the absence of trade figures may represent a genuine cessation of trading relations. Since it is difficult to distinguish missing flows from absent flows, no values for missing data were assumed.

Special trade of commodities, rather than general trade, is reported. Special trade entails imports for home consumption and exports of domestically produced goods, while general trade includes special trade plus transit trade. In most cases states report separate categories for special trade and general trade. However, in some cases states fail to distinguish between special and general trade or report only general trade.

Trade with colonies is excluded from dyadic transactions and national totals for states possessing colonial territories. Although it would be interesting to include an analysis of trade between colonies and colonial powers, the limitations posed by unavailable data proved too great a hindrance to this goal.

In general, data reported by the importing nation are used for dyadic trade flow figures. However, the scarcity of trade data prior to 1900 often required Barbieri to rely on one nation's accounts for both import and export figures. Since systematic reporting appears to be correlated with level of development, particularly in the nineteenth century, for the years preceding 1885 data were entered for both imports and exports by nation, according to an approximation of their level of development. That is, the data of developed states were entered first and supplemented by that provided by less developed states. Data entry for this period begins with the major powers and is then supplemented by minor powers.

Checks for the accuracy of reports for the nineteenth century were conducted by examining the correlation of trade reports between trading partners. When one nation consistently reported figures similar to those reported by its partners, the state was deemed to be a credible reporter. When a discrepancy arose between trading partners, Barbieri relied on the reports of the state that consistently reported flows similar to those reported by the majority of its partners (i.e., where matched partner reports were more consistent). Figures for states that consistently overvalued or undervalued its dyadic flows were scrutinized when a drastic discrepancy in reports occurred. Here, again, the dyadic flow reported by the more reliable reporting state was used.

After 1900, and particularly after 1910, data were more readily available for all independent states. From 1900-1945 data were first recorded by the figures reported by the importing nation. When entering data into the trade matrix in a given year, Barbieri proceeded through the matrix by country, entering all reported import values with each trading partner. After completing all states' records of imports, she proceeded through the matrix once again by country and supplemented the missing data of dyadic flows with the information provided by the exporting states. Again, it was necessary at times to rely on one state for both the import and export reports in a given dyadic relationship. However, this was avoided whenever possible.

The figures provided by the importing nation are used for several reasons. First, relying on either the importer's or exporter's report permits us to utilize information derived from each state in a given dyad. This diminishes the error introduced if one state has a tendency to under-report or over-report its trade flows. Second, import figures were chosen rather than export figures, since they are generally considered more reliable and comprehensive than export records. The greater tendency for states to impose tariffs on imported, rather than exported, goods is believed to result in more rigorous efforts to adopt systematic measures to record import trade. Finally, relying on either import or export flows reduces the discrepancy of flows reported in c.i.f. (i.e., cost, insurance, freight or charged in full) and f.o.b. (i.e., free on board). The former includes costs of transferring the good to the point of consumption, while the latter excludes such costs. Imports are generally reported c.i.f. and exports f.o.b. Therefore, some variations in the importer's and exporter's reports are expected. For example, the International Monetary Fund (1991) estimates that figures for c.i.f. are 10% greater than those reported in f.o.b.

Aggregated Data

As in the post-WWII cases discussed in the main codebook, one difficulty imposed in the Barbieri project resides in the problem of aggregated trade figures. In some instances a country

combined the value of its trade with two or more nations. However, when trade with two or more states is explicitly named, it is possible to use a triangular method to disaggregate trade flows. For example, if Great Britain combines the value of trade conducted with Spain and Portugal into one total, one could then examine the value that Spain and Portugal report as trading with Britain. When only one of these states reports its trade with Britain, the total of the reporting state is subtracted from the aggregate value provided in Britain's trade report. Thus, information is available for Britain's trade with two of its partners, by combining information provided by a third party.

Unfortunately, in both the pre and post-WWII period, it is impossible to disaggregate the figures for the economic union of Belgium and Luxembourg. Both the League of Nations and the IMF provide the aggregated figure for the union until 1996. For the pre-WWII period trade statistics for the union are entered as Belgian trade, since Luxembourg accounts for a less significant share of the union's trade. This issue undoubtedly requires further attention in future research. For example, in the post-WWII period, where GDP figures were available, we estimated the share of the union's trade conducted by Luxembourg and Belgium.

Another case of aggregated trade figures is more problematic and remains unresolved. In both the pre-WWII and post-WWII periods, national trade statistics generally include a category entitled 'Trade with Other Countries.' in which the value of transactions with minor trading partners is aggregated, but no list of states contained in this category is provided. It is therefore impossible to distinguish nations that possess no trade ties from those who possess minor trade flows. Some scholars assume that the absence of a state's name from another nation's trade records denotes the absence of trade. Trade might actually exist, but be too minimal to include as a separate entry in the trade report. Since there are theoretical and methodological distinctions between zero trade and minor trade, missing data are treated as such, without substituting zero trade for absent trade figures. Clearly, the prospect for conflict to arise between states that have no ties whatsoever is significantly different from that which would exist between states that have some contact. Minor ties presuppose a relationship between states, while zero trade may indicate the absence of interstate interactions. It might also indicate the existence of tension that led to the cessation of trade. Unless cases are inspected individually, we would not advise making assumptions about whether or not trade exists when data are missing.

Currency and Exchange Rates

Most data were collected in local currency and converted to current US dollars. The majority of exchange rates used for the conversions were taken from the *Polity II* project (Gurr, Jagers, and Moore 1989). These are listed under Gurr exchange rate in the data set. The exchange rates listed in *Polity II* were originally collected by the Correlates of War Project, although some revisions were made by Gurr to account for dramatic discontinuities revealed in the time trends of individual exchange rates.⁴ Several problems arose when converting trade figures from local

⁴Barbieri acquired information about the *Polity II* exchange rates through a telephone conversation with Gurr (April 1994). He informed her that the majority of the *Polity II* exchange rates were originally collected by the COW Project. However, he and his colleagues changed some of the COW values (e.g., smoothing trends with sharp variations). They did not document which cases were changed. Philip Schafer (telephone conversation, April 1994), who collected exchange rate data for the COW Project, explained that the COW exchange rates were collected from *The Statesman's Yearbook*. He stated that when currency names were missing, one could assume that the exchange

currencies to US dollars. The primary problem was the lack of available exchange rate data for many states. In many instances trade data were available, but exchange rates were not. In addition, *Polity II* contains a variable that lists the name of the national currency to which the exchange rate is presumed to correspond. However, in many instances no currency name is given.

According to the original compilers of the COW exchange rate data, when no name is given for a given state's currency in a given year, the exchange rate should correspond to the national currency reported in the *Statesman's Yearbook*. Yet, in some cases where the currency name was missing, the corresponding state had multiple-currencies or multiple exchange rates. Thus, it was necessary for Barbieri to compile exchange rate figures for many countries. (The problems encountered in the exchange rate project were as extensive as those confronted with the trade data, but are only described briefly here). Upon further investigation, Barbieri found that some of the exchange rates reported in *Polity II* differed significantly from those found in two or more alternative sources.

Also, in some instances, particularly in cases of Latin American states, the values of import and export flows are reported in two different currencies. For example, silver pesos may be used for imports, while gold pesos are used for exports. This requires separate exchange rates for converting imports and exports into US dollar values.

A comprehensive assessment of exchange rates was undertaken and data were compiled to supplement missing and questionable exchange rate values. Data for each nation's exchange rate series were reviewed to identify dramatic departures from the general time series trend. These inconsistencies were investigated by reading country profiles to determine whether or not the nation experienced a change in currency or whether there were real variations attributable to hyperinflation of other trends. Wherever possible, multiple sources of exchange rates were compared to determine whether the values reported across sources were similar. In most cases, two reports were similar in value.

Exchange rate sources include Bidwell (1970) *Currency Conversion Tables: A Hundred Years of Change* (London: Rex Collings); the *Statesman's Yearbook* (1870-1940); and US Department of Commerce (1920-1939) *Foreign Commerce Yearbook*. When using the *Statesman's Yearbook* as a source of exchange rates, conversions were first made from local currency to British pounds and then converted from pounds to dollars, since this source reports local currency in terms of British pounds. Barbieri's calculations are listed under Katrate.

Pre-WWII Trade data

To acquire nineteenth and early twentieth century trade data, an exhaustive search of historical documents, including national almanacs, commerce dictionaries, and government documents, was undertaken. Initially, Barbieri sought to collect data dating back to 1816, to correspond with what most international relations scholars identify as the beginning of the current interstate system. However, trade data for the pre-1870 period were too scarce to make any

rate corresponded to the currency name listed as the national currency in the yearbook. Unfortunately, Barbieri identified many instances where the exchange rates reported in *Polity II* did not correspond to the rates appearing in *The Statesman's Yearbook*. It was unclear whether these discrepancies corresponded to Gurr's value adjustments or whether they represented reporting errors in the original COW database. Barbieri sought alternative sources for exchange rates when currency names were absent from *Polity II*.

meaningful analysis possible. Thus, the project begins with the years following the unification of Germany and Italy.

The Statesman's Yearbook (1870-1940) (SMYB in source code) was the primary source for trade figures for years preceding 1912. All volumes of the yearbook were used, since the trade figures for different countries do not always correspond with the almanac's publication dates. For example, the 1900 edition of *The Statesman's Yearbook* may contain values for China's trade for 1885 through 1887, but contain figures for the United States from 1898 through 1900. *The Statesman's Yearbook* contains country profiles that usually include tables of foreign trade figures. When these tables are not present, information was pieced together by reading entries related to a particular state's economic activities. For example, sections on economic activities may contain information about the country's total trade figures. In addition, references are often made to the share of trade conducted with a state's top trading partners. For example, an entry might explain that a third of the state's trade is conducted with a given partner. This information was used to derive the relevant dyadic trade figures from the total trade values. For the nineteenth century, many of the trade figures for developing states were obtained by this method.

Reviewing each country's economic profile was a tedious process, but it proved valuable for providing additional information about deviations from the general norms of national reports. For example, information about multiple currencies operating within a given state or about the devaluation or revaluation of a currency was often identified by reading the country profile.

Data for developing states are more readily available in tabular form after 1885. For the period 1873-1885 U.S. Congressional records proved to be a useful source of trade data, in particular *U.S. Congress (House) Miscellaneous Documents* (1887), "Abstract of the Foreign Commerce of Europe, Australia, Asia and Africa, 1873-1885," United States Consular Reports, No. 85, October. (Washington: Government Printing Office). Data for this period were supplemented with other sources, including the *Statesman's Yearbook*; Mitchell (1982) *International Historical Statistics for Africa and Asia* (New York: New York University Press); Mitchell (1983) *International Historical Statistics for the Americas and Australasia* (Detroit, MI: Gale Research Company); and Wattenberg (1976) *Introduction and User's Guide to The Statistical History of the United States from Colonial Times to the Present* (New York: Basic Books).

For the period 1912-1938, the primary source used was the *League of Nations* (1912-1945) annual publications of *International Trade Statistics*, (Geneva: League of Nations). The title for this annual publication varies and includes the names *Memorandum on International Trade and the Balance of Payments* and *International Trade Statistics*. Data were taken from each annual volume. In addition, some data for 1935 and 1938 were derived from the League of Nations' (1942), *The Review of World Trade*. League of Nations' data were supplemented with data from *The Statesman's Yearbook* and Mitchell (1982, 1983).

Post-WWII Trade Data

For the post-WWII period, the majority of trade data are derived from the International Monetary Fund's (Inter-university Consortium for Political and Social Research, 1991) *Direction of Trade Statistics electronic tape*. Data were reconfigured from national accounts to dyadic trade flows. As in the pre-WWII case, reports of dyadic trade flows were derived from the importing countries' reported trade figures. When these figures were absent, the exporter's reports were used. The values that each state reports to import from each partner were added to derive the dyadic

total. Each state's total imports and export figures were combined to arrive at each nation's total trade.

In many instances, the electronic version of the IMF data tape reports trade flows as zero or missing, but these trade values are reported in the print version of the IMF's annual publications. Missing data were investigated and supplemented with The International Monetary Fund's *International Financial Statistics* (Washington, D.C.: IMF Statistics Department, Monthly 1956-1998) and *The Direction of Trade Statistics Yearbook*. (Washington, D.C: Real Sector Division, IMF Statistics Department, Quarterly 1956-1998).

Other supplementary material used to substitute the IMF's missing values, particularly for non-members of the IMF, include: Arthur S. Banks (1979) *Cross-National--Time-Series Data Archive* (Binghamton, NY: Center for Social Analysis); The United Nations *Yearbook of International Trade Statistics*, later renamed *International Trade Statistics Yearbook* (New York: Department of International Economic and Social Affairs Statistical Office, Annual 1950-1996); Europa Publications, *The Europa Year Book Vol. I. Europe*. (London: Europa Publications Ltd., Annual 1960-1996); M.C. Kaser and E.A. Radice, eds. (1985) *The Economic History of Eastern Europe 1919-1975, Vol. I: Economic Structure and Performance between the Two Wars* (Oxford: Clarendon Press); Paul Maher (1972) *Soviet and East European Foreign Trade, 1946-1969* (Bloomington: Indiana University Press); State Statistical Bureau of the Peoples Republic of China, Renze Zhang, ed., *China Statistical Yearbook* (New York: Praeger Publishers, 1990); Wray O.Candilis (1968) *The Economy of Greece 1944-66*. (New York: Fredrick A. Praege, Publishers); Rodger A Clarke and Dubravko J.I. Matko (1983) *Soviet Economic Facts, 1917-1981* (New York: St. Martin's Press); Zbigniew M Fallenbuchl, ed. (1976) *Economic development in the Soviet Union and Eastern Europe, Volume 2* (New York: Praeger Publishers); David Lascelles (1976) *Comecon to 1980* (London: The Financial Times); Marie Lavigne, ed. (1990) *The Soviet Union and Eastern Europe in the Global Economy* (Cambridge: Cambridge University Press); V. P. Gruzinov (1979) *The USSR's Management of Foreign Trade* (White Plains, NY: M.E. Sharpe, Inc); John B. Quigley (1974) *The Soviet Foreign Trade Monopoly* (Columbus: Ohio State University Press); William Nelson Turpin (1977) *Soviet Foreign Trade, Purpose and Performance* (Lexington, MA: Lexington Books); U.S. Department of Commerce, various publications including: *Selected U.S.S.R. and Eastern European Trade and Economic Data* (Bureau of East-West Trade, May, 1974); *Selected U.S.S.R. and Eastern European and Economic Data* (Bureau of East-West Trade, June, 1973); *Selected Trade and Economic Data of the Centrally Planned Economies* (Bureau of East-West Trade, May, 1975); *The Chinese Economy and Foreign Trade Perspective-1976* (Bureau of East-West Trade, June, 1977).⁵

Coding Rules for Trade Data

Barbieri included zero values in some instances where in-depth investigation revealed an absence of official trade ties. This, however, is an issue that requires further investigation. Differences in coding rules for dealing with missing data can lead to significant variations in trade data sets among researchers, particularly since missing data account for a large percentage of overall dyadic trade figures. As mentioned, some researchers treat missing values as zero values, but we caution researchers about adopting this rule. In some cases, data for country trade totals are missing, even if there is evidence that the state engages in trade. Thus, you can not assume that missing data implies no trade.

⁵Some of these sources were also used for the pre-WWII period.

In discussions with other scholars collecting or employing trade data, we found a number of areas in which scholars differ on decision rules regarding trade figures. This may be the source of discrepant empirical findings on issues such as trade and conflict, where different coding decisions lead scholars to assign different values to the degree of trade between states. Scholars employing trade data should think through the decision rules they adopt and consider how these rules might alter empirical findings. This is an area of research that requires further investigation.

Many of these issues are discussed by Barbieri, Keshk, and Pollins (2008).

Supplementary References

- Banks, Arthur S. 1979 *Cross-National--Time-Series Data Archive*. Binghamton, NY: Center for Social Analysis
- Barbieri, Katherine. 2002. *The Liberal Illusion: Does Trade Promote Peace?* Ann Arbor: University of Michigan Press.
- Candilis, Wray O. 1968 *The Economy of Greece 1944-66*. . New York: Fredrick A. Praeger, Publishers.
- Clarke, Rodger A. and Dubravko J.I. Matko. 1983. *Soviet Economic Facts, 1917-1981*. New York: St. Martin's Press
- Europa Publications. *The Europa Year Book Vol. I. Europe*. London: Europa Publications Ltd., Annual 1960-1996.
- Fallenbuchl, Zbigniew M, ed. 1976 *Economic development in the Soviet Union and Eastern Europe, Volume 2*. New York: Praeger Publishers.
- Gruzinov, V. P. 1979 *The USSR's Management of Foreign Trade*. White Plains, NY: M.E. Sharpe, Inc.
- Gurr, Ted Robert, Keith Jagers, and Will H. Moore. 1989. *Polity II Handbook*. Boulder: University of Colorado, mimeo.
- International Monetary Fund. 1991. *Direction of Trade Statistics electronic tape*. ICPSR 7628. Made available through the Inter-university Consortium for Political and Social Research. Direction of Trade Codebook. 8th ICPSR Release. Ann Arbor: Inter-university Consortium.
- The International Monetary Fund. *International Financial Statistics*. Washington, DC: IMF Statistics Department, Monthly 1956-1998.
- The International Monetary Fund. *The Direction of Trade Statistics Yearbook*. Washington, DC: Real Sector Division, IMF Statistics Department, Quarterly. 1956-1998.
- Kaser, M.C. and E.A. Radice, eds. 1985 *The Economic History of Eastern Europe 1919-1975, Vol. I: Economic Structure and Performance between the Two Wars*. Oxford: Clarendon Press.
- Lascelles, David. 1976 *Comecon to 1980*. London: The Financial Times.
- Lavigne, Marie, ed. 1990 *The Soviet Union and Eastern Europe in the Global Economy*. Cambridge: Cambridge University Press.
- League of Nations. 1912-1945. *International Trade Statistics*. Geneva: League of Nations.
- League of Nations. 1910-1940. *Memorandum on International Trade and the Balance of Payments and International Trade Statistics*. Geneva: League of Nations.
- League of Nations. 1942. *The Review of World Trade*. Geneva: League of Nations.

- Maher, Paul. 1972. *Soviet and East European Foreign Trade, 1946-1969*. Bloomington: Indiana University Press.
- Mitchell. 1982 *International Historical Statistics for Africa and Asia*. New York: New York University Press
- Mitchell. 1983 *International Historical Statistics for the Americas and Australasia*. Detroit, MI: Gale Research Company
- Quigley, John B. 1974 *The Soviet Foreign Trade Monopoly*. Columbus: Ohio State University Press.
- The Statesman's Yearbook. 1864-1990*. London: Macmillan and Company.
- Turpin, William Nelson. 1977. *Soviet Foreign Trade, Purpose and Performance*. Lexington, MA: Lexington Books.
- The United Nations Yearbook of International Trade Statistics*, later renamed *International Trade Statistics Yearbook*. New York: Department of International Economic and Social Affairs Statistical Office, Annual 1950-1996.
- United States. Bureau of the Census. 1975. *Historical Statistics of the United States, Colonial Times to 1970*. Washington, DC: U.S. Department of Commerce Bureau of the Census.
- United States. Bureau of East-West Trade. 1976. *The Chinese Economy and Foreign Trade Perspective*. Washington, DC: Department of Commerce, Domestic and International Business Administration, Bureau of East-West Trade.
- United States. Bureau of East-West Trade. 1975, 1976. *Selected Trade and Economic Data of the Centrally Planned Economies*. Washington: DC, Department of Commerce, Industry and Trade Administration, Bureau of East-West Trade.
- United States. Bureau of East-West Trade. 1973, 1974. *Selected U.S.S.R. and Eastern European Trade and Economic Data*. Washington, DC: Department of Commerce.
- United States Congress (House) Miscellaneous Documents. 1887. "Abstract of the Foreign Commerce of Europe, Australia, Asia and Africa, 1873-1885," United States Consular Reports, No. 85, October. Washington, DC: Government Printing Office.
- Wattenberg, Ben J. 1976. *Introduction and User's Guide to the Statistical History of the United States from Colonial Times to the Present*. New York: Basic Books, Inc.
- World Bank. 1995. *World Data Series* on CD-Rom.
- World Bank. International Economics Department. 1993. *World Tables of Economic and Social Indicators, 1950-1992 Codebook*. ICPSR 6159. Data supplied by the Inter-university Consortium for Political and Social Research.
- Renze Zhang, ed., 1990. *China, China Statistical Yearbook*. New York: Praeger Publishers.

Version/Data History

Version 2.0 was posted publicly on June 23,2008.

Version 2.01 was posted publicly on September 30, 2008, with changes to a handful of data points to remove decimal-place errors. The following changes were made:

File: dyadic_trade_2.01.csv (old version: dyadic_trade_2.0.csv):

1. Dyad 2-155 year 1925 flow 2 value was changed from 40998 to 40.998
2. Dyad 2-155 year 1931 flow 1 value was changed from 33290 to 33.29
3. Dyad 155-160 year 1925 flow 1 value was changed from 4456 to 4.456
4. Dyad 155-220 year 1930 flow 1 value was changed from 8487 to 8.487
5. Dyad 155-220 year 1931 flow 1 value was changed from 5033 to 5.033
6. Dyad 155-220 year 1932 flow 1 value was changed from 792 to 0.792

File: dyadic_trade_supplement_2.01.csv (old version: dyadic_trade_supplement.csv):

1. 1925 chl usa (155,2) changed “Exp” from 3 to 6 on 8/3/08; changed “Flow 1” from 88384 to 88.384
2. 1931 usa chl (2, 155) changed “Exp” from 6 to 3 on 8/3/08; changed “Flow 1” from 33290 to 33.29
3. 1925 chl arg (155, 160) changed “Exp” from 6 to 3 on 8/3/08; changed “Flow 1” from 4456 to 4.4576
4. 1930 chl frn (155, 220) changed “Exp” from 6 to 3 on 8/3/08; changed “Flow 1” from 8487 to 8.487
5. 1931 chl frn (155, 220) changed “Exp” from 6 to 3 on 8/3/08; changed “Flow 1” from 5033 to 5.033
6. 1932 chl frn (155, 220) changed “Exp” from 6 to 3 on 8/3/08; changed “Flow 1” from 792 to .792

File: national_trade_and_exchange_rate_supplement_2.01.csv also updated (old version: national_trade_and_exchange_rate_supplement.xls)

File: national_trade_2.0.csv is unchanged in release 2.01.