

Appendix A

NOTES ON TRADE DATA

The trade database project began in 1991 as part of my doctoral research, but remains ongoing. The International Trade Database employed in this analysis is a revised and expanded version of the database described and employed in my dissertation research (Barbieri 1995). The database includes dyadic trade-flow figures and national import and export figures. These data are available at the Peace Science Society (International) website at <<http://pss.la.psu.edu/intsys.htm>>.

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.¹ Those that were available are included in the analyses.² The data set only includes reported statistics derived from trade reports; no estimation techniques were employed to replace missing values, as discussed later.

Missing data for exchange rates, cases including only partial information about dyadic flows or national totals, and reports containing unrealistic values are coded as missing values. Unrealistic values include those cases where dyadic trade exceeds the total trade of one of the nations in the dyad. This situation can arise from a number of factors including erroneous trade reports, inconsistencies between partner reports, and 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 cor-

responding 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, I did not interpolate the figures to derive a figure for 1899, but counted that point as missing data. My rationale for refraining from interpolation and extrapolation techniques resides in my belief that many historical reports already include cases where states or publishers have employed these techniques to replace missing values. Additional data manipulation would introduce further error by assuming that trade conformed to a particular trend, when missing data may instead correspond 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 impossible to distinguish missing information 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 transit trade. In most cases states report separate categories for special and general trade. However, in some cases states fail to distinguish between special and general trade or only report 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 me 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 began with the major powers and was 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 source.

When a discrepancy arose between trading partners, I 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 their 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 are more readily available for all independent states. From 1900 through 1945 data were first recorded by the figures reported by the importing nation. When entering data into the trade matrix in a given year, I proceeded through the matrix by country, entering all reported import values with each trading partner. After completing all states' records of imports, I proceed through the matrix once again by country and supplemented the missing data of dyadic flows with the information provided by exporting states. Again, it was necessary at times to rely on one state for both the import and export values within a 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 importers' or exporters' reports permits me to utilize information derived from each state in a given dyad. This diminishes the error introduced if one state has a tendency to underreport or overreport 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 variation in the importer and exporter's reports are expected. For example, the International Monetary Fund (1991) estimates that figures for c.i.f. are 10 percent greater than those reported in f.o.b. Relying on either the export or import flow will improve the consistency in the range of values reported by each state for one particular flow.

Aggregated Data

One difficulty imposed in the project resides in the problem of aggregated trade figures. In some instances a country combines the value of its trade with two or more nations. However, when trade with two or more states is explicitly identified, 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, I examined 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 was subtracted from the aggregated value provided in Britain's trade report. Thus, information was 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-World War II 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. For the pre-World War II 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, it may be possible to estimate the share of the union's trade conducted by Luxembourg in order to include Luxembourg into the analysis. At this time, little information was found to resolve this problem.

Another case of aggregated trade figures is more problematic and remains unresolved. In both the pre-World War II and post-World War II periods, national trade statistics generally include a category entitled "Trade with Other Countries" in which the values of transactions with minor trading partners are 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 studies 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.

Currency Conversion

Most data were collected in local currency and converted to current U.S. dollars. The majority of exchange rates used for conversions were taken from the *Polity II* project (Gurr, Jagers, and Moore 1989). 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.³ Several problems arose when converting trade figures from local currencies to U.S. dollars. The primary problem was the lack of available exchange rates for many states. In many instances trade data reports were available, but exchange rates were unavailable. 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 compilers of the data, when no name is given, the exchange rates should correspond to the national currencies reported in the *Statesman's Yearbook*. Yet, in some cases with a missing currency name, the state in question has multiple currencies. Thus, it was necessary to compile figures for exchange rates 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, I 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 Latin American states, the value 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 U.S. 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 the nation experienced a change in currency or whether

there were real variations attributable to hyperinflation or 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.

Sources for exchange rates include Bidwell (1970), *Currency Conversion Tables: A Hundred Years of Change* (London: Rex Collings); the *Statesman's Yearbook* (1870–1940); and U.S. Department of Commerce (1920–39), *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.

Pre–World War II 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, I 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 are too scarce to make any meaningful analysis possible. Thus, the project begins with the years following the unification of Germany and Italy.

The Statesman's Yearbook (1870–1940) 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, reading the country profile often identified information about multiple currencies operating within a given state or about revaluation of a currency.

Data for developing states are more readily available in tabular form after 1885. For the period 1873–85 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, DC: 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: Gale); 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 *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–World War II Trade Data

For the post–World War II 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–World War II 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 their annual publications. Missing data were investigated and supplemented with the International Monetary Fund's *International Financial Statistics* (Washington, DC: IMF Statistics Department, Monthly 1956–98) and *The Direction of Trade Statistics Yearbook* (Washington, DC: Real Sector Division, IMF Statistics Department, Quarterly 1956–98). Missing data for the post–World War II period have not all been investigated. This project remains ongoing.

Other sources for the post–World War II period 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 1950–96); Europa Publications, *The Europa Year Book*, vol. 1, *Europe* (London: Europa, Annual 1960–96); M. C. Kaser and E. A. Radice, eds. (1985), *The Economic History of Eastern Europe, 1919–1975*, vol. 1, *Economic Structure and Performance between the Two Wars* (Oxford: Clarendon); Paul Maher (1972), *Soviet and East European Foreign Trade, 1946–1969* (Bloomington: Indiana University Press); State Statistical Bureau of the People's Republic of China, Renze Zhang, ed., *China Statistical Yearbook* (New York: Praeger, 1990); Wray O. Candilis (1968), *The Economy of Greece, 1944–66*. (New York: Praeger). Rodger A. Clarke and Dubravko J. I. Matko (1983), *Soviet Economic Facts, 1917–1981* (New York: St. Martin's); Zbigniew M. Fallenbuchl, ed. (1976), *Economic Development in the Soviet Union and Eastern Europe*, vol. 2 (New York: Praeger); David Lascelles (1976), *Comecon to 1980* (London: 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); 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).⁴

Coding Rules for Trade Data

In my early research I treated zero and missing values as missing data. I now generally include zero values only if these values have been investigated and multiple sources have revealed little or no trade between states. 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 such a large percentage of overall dyadic trade figures. Some researchers treat missing values as zero values, but I would 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, one cannot assume that a missing datum implies no trade. As mentioned previously, I have found the same tendency with the zero values reported by the IMF, where zero values may simply mean missing values.

In discussions with other scholars collecting or employing trade data, I 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.