Expanded Trade and GDP Data

KRISTIAN SKREDE GLEDITSCH
Department of Political Science
University of California, San Diego

Many sources of economic data cover only a limited set of states at any given point in time. Data are often systematically missing for some states over certain time periods. In the context of conflict studies, economic data are frequently unavailable for states involved in conflicts, undermining the ability to draw inferences of linkages between economic and political interactions. For example, simply using available data in a study of trade and conflict and disregarding observations with missing data on economic variables excludes key conflicts such as the Berlin crisis, the Korean War, the Cuban Missile Crisis, and the Gulf War from the sample. A set of procedures are presented to create additional estimates to remedy some of the coverage problems for data on gross domestic product, population, and bilateral trade flows.

Comparative international research typically claims to pertain to a population of all states in the international system. Many sources of economic data, however, cover only an incomplete set of the states in the world at any given point in time.[1] Even more seriously, many data sources appear to systematically lack data from some states over certain time periods.

A cursory examination reveals that some types of states are more likely to lack valid observed economic data than others. Gross domestic product (GDP) coverage is relatively comprehensive for the Organization of Economic Cooperation and Development (OECD) member states, but data are often lacking for developing countries and socialist economies. Similarly, bilateral trade flow data tend to be poor—if at all available—for smaller, developing countries that are far apart or pairs of antagonistic states. Such missing data can induce a variety of problems that may affect inferences in the line of empirical research examining whether trade reduces conflict (see Oneal and

1. The universe in international studies is typically taken to be all states in the world within some time period, but the population of these states is rarely explicitly defined. In this study, I rely on the list outlined by Gleditsch and Ward (1999) for the population of states since 1816, including the supplementary list of microstates with fewer than 250,000 inhabitants. Because most economic data are available only after World War II, I limit myself here to the post-1945 period.

AUTHOR'S NOTE: The version of the expanded trade and GDP data described in this article (version 2.0, completed April 11, 2002) are available through http://www.yale.edu/unsy/jcr/jcrdata.htm. I am grateful to Nathaniel Beck, Erik Gartzke, Håvard Hegre, Michael Hiscox, Kosuke Imai, Scott Kastner, Gary King, Richard Tucker, and Michael Ward for comments and advice. I am in particular indebted to John Oneal, who meticulously compared his data with these estimates and pointed out numerous issues. The Human Security Program at Harvard University provided financial support.

JOURNAL OF CONFLICT RESOLUTION, Vol. 46 No. 5, October 2002 712-724
DOI: 10.1177/002200202236171

Russett [1999] and Schneider and Barbieri [1999] for overviews of recent debates). Data are likely to be missing for many dyads that are interconnected, but not necessarily capitalist or democratic, such as the socialist Council for Mutual Economic Assistance (CMEA) states. Thus, the only highly interdependent states observed in a sample may well be OECD-type countries. Similarly, many dyads involving developing countries or East-West dyads with little trade and high levels of conflict often have missing data. Perhaps most troubling for conflict research, countries involved in conflict often lack economic data, and key conflicts may drop out of the sample.

This article is structured as follows: I first review more formally why missing data can create problems in empirical research. I then outline a set of procedures to produce estimates for missing observations in data on GDP and bilateral trade flows. Finally, I illustrate some consequences of missing data in a simple example examining the relationship between trade and conflict.

THE MISSING DATA PROBLEM

Most research ultimately relies on some form of comparison among observations within a population of interest. In research in international relations, the universe is frequently clearly finite and quite small. Thus, the investigator may try to study the entire universe rather than a random sample. Missing data on particular variables, however, will lead to attrition from the population data. There are, in principle, no problems in generalizing from a subsample to the larger population if the smaller sample is a true random sample of the population. The resulting subsample researchers are left with after deleting observations with missing data from a population, however, may often differ systematically from the population at large. Unfortunately, the processes or mechanisms that induce missing data often make certain types of cases more likely not to be observed than others. As a result, missing data problems can lead to systematically biased samples. It is well known that comparisons relative to a nonrandom sample that differs from the population in some systematic fashion can yield misleading inferences (see Rubin [1976] for an extended discussion).

Researchers often ignore missing data problems and proceed by listwise deletion—that is, working with the sample that remains after omitting each case or row of the data matrix with missing observations. It is implicitly assumed that the resulting sample constitutes a representative sample of the population. However, when the attrition due to missing data is not random, simply going with the available data when faced with missing data can lead researchers astray. King et al. (2001) show that a missing data problem may often be a more serious problem than omitted variable bias.

APPROACHES TO MISSING DATA

Missing data are not observed or recorded, and researchers are often reluctant to make any assumptions about these cases. However, simply omitting missing observations as they arise is effectively a decision to limit the sample according to data avail-

ability. The risk of making some potentially incorrect or unwarranted assumptions may be less serious than submitting to the perils of listwise deletion.

Even though missing data are unknown in that they have not been observed by some reporting agency collecting the data, researchers are usually not entirely ignorant about the nonobserved characteristics. In many cases, researchers possess information and insights that can be applied to make informed guesses. This article shows how one set of estimates can be derived for two economic sources of data: the Penn World Tables (PWT) and the International Monetary Fund's (IMF) Direction of Trade (DOT) data.

There is considerable work in statistics on estimation techniques with missing data. These typically take advantage of observed data on other variables to estimate missing observations on particular variables (see King et al. [2001] for a review). Although these approaches certainly may be applied to missing data problems for trade and GDP data, the choice of what variables to include in the imputation model—especially whether to include political variables associated with conflict—can be seen as influencing answers to research questions. Without good predictor variables, a statistical imputation model will produce point estimates associated with considerable uncertainty and large variances. As a first cut, it seems valuable to see how much additional information can be extracted from the existing data and alternative sources in a transparent manner before resorting to multiple imputation techniques. This article takes a pragmatic first step in trying to find some initial improved estimates.

EXPANDED GDP DATA

BASE DATA: THE PWT

GDP is by far the most common indicator of a state's resources or economic wealth. The PWT data (Summers and Heston 1991), produced by the Center for International Comparisons at the University of Pennsylvania, is often considered the most theoretically attractive GDP measure. The PWT data display a set of national accounts denominated in a common set of prices in a common currency so that real quantity comparisons can be made, both between countries and over time. The population figures are in 1,000s. The GDP per capita figures included in this revised version are real figures in constant U.S. dollars (base 1985) as well as nominal figures in U.S. dollars in current international prices.

The PWT mark 5.6 data are available from 1950 to 1992. Unfortunately, about 25% of the observations in the population are missing in the PWT data. Quite a few countries in the international system are not included in the PWT data altogether or included only for a short time period.[2]

To address the case of countries not included, I have generated a set of GDP per capita and population estimates based on the figures reported in the Central Intelligence Agency's (CIA's) *World Factbook* (1998). I classified a country with no

2. There are no within-series missing observations in the PWT data.

TABLE 1
Gross Domestic Product and Population Data Categories

Code	*Data Origin*	*Share of Observations*
0	Observed data from the PWT	76.16
1	Lags and leads based on first nonmissing observations, deflated to current prices	11.14
2	Estimate based on figures from the *World Factbook*	12.70

SOURCE: For the Penn World Table (PWT) data, see Summers and Heston (1991); for the *World Factbook* data, see Central Intelligence Agency (1998).

recorded data as falling into one of six categories depending on its size and economic characteristics. For each of the categories, I chose a putatively similar reference country for which there is good coverage in the PWT 5.6 data. The categories and their respective reference countries were as follows: (1) developing socialist economies—Ethiopia, (2) small developing countries—Jamaica, (3) small industrialized countries—Switzerland, (4) industrialized socialist economies—Hungary, (5) developing African states—Guinea, and (6) large developing states—Pakistan. For each of the states with missing data, I used the CIA estimate to find the proportional size of its GDP per capita and population relative to the reference country. These proportions are shown in Tables A1 through A6 in the appendix.

These CIA estimates are available for only a single year. In this version of the data, the relative proportions are assumed to remain constant over the time period. One could improve the estimates by adjusting the proportions over time using data from older versions of the *World Factbook*.

The relative share of missing data varies considerably over time in the PWT data and ranges from about 12% missing in 1985 to more than 50% in 1992. Such missing leads or tail parts of series were filled in by estimates based on the first/last nonmissing observations. These estimates assume that real GDP per capita remains the same for the lead/tail part. The figures were deflated to current prices using a U.S. GDP deflator.

Table 1 displays the codes that identify the origin of observations in the GDP per capita and population data.

TRADE DATA

The DOT data set (IMF 1997) consists of annual time series on the countries' exports and imports by partner countries in millions of current-year U.S. dollars. In practice, most social scientists obtain these data through the Inter-university Consortium for Political and Social Research (ICPSR) (deposited as study no. 7628). The data are grouped by dyad or pairs of countries. Any two countries A and B give rise to a dyad AB. In any given year, each dyad AB in principle has four trade flows: (1) exports from A to B, (2) imports by A from B, (3) exports from B to A, and (4) imports by B from A. In principle, exports from A to B should equal imports by B from A, and vice

versa. However, it is well known that these accounting identities often do not hold in recorded data on trade (Rozanski and Yeats 1994).[3]

According to the codebook provided by ICPSR, the DOT data are based on reports generated by IMF member states. However, the codebook also indicates that some data are "estimates derived from the reports of partner countries for non-reporting and slow-reporting countries" (p. 2). The data are expressed in millions of U.S. dollars in current international prices.[4] The data currently available cover 1948 to 1996. All exports are valued f.o.b. ("free on board"). Imports are generally valued c.i.f. ("customs and freight included"). For some countries, f.o.b. figures are given; in some cases, both c.i.f. and f.o.b. figures are included.[5] In this version, c.i.f. figures are used whenever available. An import code of 1 indicates that import figures are f.o.b.[6]

The DOT data contain observations for only about 40% of all the dyad years between 1948 and 1996. Coverage is particularly poor for developing and socialist states.

TRADE DATA IMPUTATIONS

A first step to improve the coverage of these data is to replace missing data with estimates from alternative data sources. The World Export Data (WED) (Faber and Nierop 1989) database has tried to compile data to fill in some of the gaps for export figures. These data are available for the years from 1948 to 1983. The WED data are particularly helpful in providing data on socialist states not included in the IMF data during the cold war period. All figures are f.o.b. in millions of U.S. dollars at current prices. According to the WED codebook,

> in case of exports from nations with centrally planned economies to other nations, we sticked [*sic*] to the IMF approach of converting the other nations' figures on annual imports (cost, insurance and freight included; cif) from nations with centrally planned

3. The trade data most familiar to political scientists are the data used by Oneal and Russett (e.g., 1999) in their studies on trade and conflict, which the authors have made publicly available. Their data are largely based on the same sources and include estimates for missing data based on rules quite similar to those applied to the expanded trade data presented here. However, Oneal and Russett assume no or zero trade flows for a much larger number of dyad years than is the case for these expanded data. Oneal and Russett assume no trade for a total of about 56% of the dyad years in their sample. The data outlined here provide alternative estimates for about 71% of the dyad years assumed to have no trade in the Oneal and Russett data. Although trade in these dyads tends to be much lower than in most dyads included in the original DOT data, many of these observations have nonzero trade flows. Some of the flows in an AB dyad have also been averaged in the Oneal and Russett data (personal communication with John Oneal).

4. According to the codebook provided with the ICPSR release, "data originally reported in national currencies are converted by the IMF with the exchange rates" (line rf) "published in the IMF's International Financial Statistics" (p. 2).

5. The ICPSR manual indicates that "a 10 percent adjustment have been applied" to f.o.b. figures (p. 2). The manual does not indicate whether these have been increased or decreased by 10%, but the former seems more likely. Many are skeptical of whether such conversions can be made in a consistent fashion for cross-country data. Because countries have different c.i.f.'s, applying a common proxy may be flawed. In addition, duties are imposed at different levels in different countries.

6. Note that the IMF reports trade for Belgium and Luxembourg jointly. I have here estimated individual trade shares proportional to the population shares of each country relative to the combined total (i.e., .96 for Belgium and .04 for Luxembourg). Trade between Belgium and Luxembourg was assumed to be proportional to the joint trade with the Netherlands, adjusting for population size.

TABLE 2
Recorded U.S.-Grenada Trade, 1981-1990

Year	*Exports US → GRN*	*Imports US ← GRN*	*Exports GRN → US*	*Imports GRN ← US*
1981	10.842[a]	0.0	0.285	10.842
1982	11.945[a]	0.0	0.432	11.945
1983	10.207[a]	0.0	0.527	10.207
1984	0	0.0	0.984	14.054
1985	0	0.0	0.719	17.164
1986	0	0.0	1.045	21.374
1987	0	0.0	1.497	23.824
1988	25.400	7.7	2.868	23.465
1989	27.900	8.1	2.221	30.240
1990	34.600	7.9	1.566	33.321

a. Indicates estimate based on reverse flow.

economies into export data. As conversion factor, we used a rough 0.9. The same procedure has been applied when no export data could be obtained but only import data.

The WED provides export data for an additional 11.4% of the total dyad years in the population.

In many cases, only some of the four annual flows between A and B have been recorded. In principle, exports from A to B should be the same as imports to B from A, and vice versa. Even though these accounting identities may not hold in the recorded data, an estimate based on the recorded data for the reverse flow seems better than treating observations as missing. All missing data have been substituted with the reverse flows whenever available.

On closer inspection, many of the trade flows of exactly zero in the DOT data seem problematic. To maintain a rectangular data structure, many missing observations appear to have been substituted with zeros. These structural zeros are probably better treated as missing observations rather than true zeros. Examine, for example, the recorded trade between the United States and Grenada displayed in Table 2.

The United States has no recorded data for exports to Grenada before 1984 (i.e., not even missing data entries in the original DOT data matrix). Because these observations are missing, Grenada's recorded figures for imports from the United States seem a logical substitute. Note, however, that U.S. exports to Grenada for 1984 through 1987 are given as zero but suddenly increase to $25.4 million in 1988. The recorded U.S. imports for Grenada, however, do not suddenly increase from zero to a substantial figure in 1988. The figures recorded for Granada's imports from the United States seem more plausible than the recorded zeros for the years 1984 to 1987, and the U.S. data prior to 1988 seem to have been inserted to create a rectangular data structure. Even though the 1984 to 1987 observations are not missing in the sense that they have been recorded as zeros by the IMF, an estimate based on the reverse flow seems more reliable than the "observed" data for such potential structural zeros.

In the expanded data, I have substituted all exact zeros with the reverse flow whenever this was not exactly zero. Missing data and suspicious zeros replaced with estimates based on the reverse flows amount to about 9% of the total annual export figures and almost 18% of the annual import figures. No assumption has been made about scaling export and import figures, but users can easily apply such conversions to these observations.

Although A's exports to some country B and its imports from B certainly need not be identical, one would generally expect these figures to be of similar magnitude. In the absence of other information, A's export to B is probably a reasonable estimate for A's imports from B and vice versa. About 6% of the population of export and import figures have been replaced by such estimates in this version of the data.

These imputations decrease the share of missing trade data for the population of dyad years from about 60% to about 33%. Linear interpolation of within time series missing data provides data for an additional 5% of the cases. Many dyad time series have spells of missing data at the beginning or end. Replacing these with estimates based on the first/last nonmissing observation, deflated to current international prices, yields data for an additional 3% of the population.

These estimates still leave 6,611 dyads that have no recorded data in any of the sources. Most of these are dyads that involve very small states or dyads of states separated by long distances. Trade in these dyads will in all likelihood be small or negligible. All these remaining dyads have been assumed to have no trade and set to zero. A separate file (notrade.asc) provided with these data identifies the dyads that have been assumed to have no trade.

Tables 3 and 4 identify the codes indicating the origin of the export and import observations as well as the shares from each category in the final expanded trade data.

DO MISSING DATA MATTER? TRADE AND CONFLICT EXAMPLES

Although a high proportion of missing data may matter in principle, is it likely to make a substantial practical difference in empirical research? Looking at some actual examples from research on trade and conflict provides one way to ascertain how missing data affect the sample and results inferred. Whether greater economic interdependence between states is likely to promote peace is one of the most prominent issues in current research in international relations. Although the controversy revolves around numerous other issues such as variable construction and model specification (see Oneal and Russett 1999; Schneider and Barbieri 1999), missing data problems may certainly be one of the factors contributing to the differences between published studies (cf., for example, Barbieri 1996; Oneal and Russett 1999). Economic interdependence is typically measured by a dependence ratio of the relative size of bilateral trade to a state's GDP, seen as an indicator of the importance of the dyadic trade to a state (see, for example, Russett and Oneal 2001). The value of such a ratio will be missing whenever data for one of the two bilateral flows (i.e., export and import) or GDP are missing.

TABLE 3
Export from Country A to Country B Data Categories

Code	Data Origin	Share of Observations
0	Observed data from DOT	40.13
1	Observed data from WED	11.08
2	Estimate based on B's imports from A (observed data from DOT)	9.29
3	Estimate based on A's import from B	6.35
4	Interpolated estimate	5.19
5	Lags and leads based on first nonmissing observations, deflated to current prices	2.85
6	Dyads with no observed data, assumed to be 0	25.11

SOURCE: For the Direction of Trade (DOT) data, see International Monetary Fund (1997); for the World Export Data (WED), see Faber and Nierop (1989).

TABLE 4
Import by Country A from Country B Data Categories

Code	Data Origin	Share of Observations
0	Observed data from DOT, c.i.f. figures	38.91
1	Observed data from DOT, f.o.b. figures	3.99
2	Estimate based on B's exports to A (observed data from DOT or WED)	17.60
3	Estimate based on A's export to B	6.35
4	Interpolated estimate	5.19
5	Lags and leads based on first nonmissing observations, deflated to current prices	2.85
6	Dyads with no observed data, assumed to be 0	25.11

SOURCE: For the Direction of Trade (DOT) data, see International Monetary Fund (1997); for the World Export Data (WED), see Faber and Nierop (1989).
NOTE: c.i.f. = customs and freight included; f.o.b. = free on board.

There has been considerable speculation on the possible influences that missing economic data may have for research on trade and conflict. Because GDP coverage is relatively comprehensive for the OECD member states but often lacking for socialist economies that are not necessarily peaceful, the effects of interdependence may be overstated in the sample of available data. At the same time, because coverage for trade-flow data tends to be poor—if at all available—for many pairs of antagonistic states, the actual effect of trade may be larger than it would appear from the available sample. Because several and potentially counteracting biases from missing data might be imagined, the net consequences for our inferences from empirical studies are ambiguous.

For purposes of illustration, consider a simple logistic regression model of interdependence and conflict. The risk of conflict is often considered a function of the *less* dependent state in a dyad, typically operationalized by the smaller of the two depend-

TABLE 5
Logit Regression of the Impact of Dyadic Economic Dependence and Distance on the Likelihood of a Militarized Interstate Dispute

	Model 1			*Model 2*		
	Coefficient Estimate	*Standard Error*	*Z Score*	*Coefficient Estimate*	*Standard Error*	*Z Score*
Constant	2.113	0.217	9.74	3.681	0.129	28.47
Dependence low	–33.600	9.261	–3.63	–49.317	7.787	–6.33
Natural logarithm of distance between capital cities	–0.893	0.029	–30.25	–1.131	0.018	–64.47
n		92,096			394,057	
LR χ^2 ($df = 2$)		914.19			3,799.63	

ence ratios in a dyad. Because the propensity for trade between states is highly dependent on distance, I also include the natural logarithm of the distance between the capital cities of the two states in a dyad. Following the conventions in the literature, I use the presence of a militarized interstate dispute (MID) as an indicator of conflict between states (see Jones, Bremer, and Singer 1996). The list of states and dates of independence used by the Correlates of War (COW) project for compiling the MID data differs somewhat from that outlined in Gleditsch and Ward (1999). In this example, I include only observations for states and time periods included in the international system by the COW project.[7]

Table 5 shows the results of estimating such a logistic regression model on the observed economic data for the years 1950 to 1992 under the model 1 column. The model 2 column indicates the results for all states in the international system based on the expanded data. As can be seen, both the sample size and the results look quite different for the two models. The sample for model 1 encompasses slightly less than one-fourth of all the relevant dyad years in the period. The coefficient for the dependence ratio is negative and significant for both models, suggesting that higher levels of trade reduce the likelihood of conflict. Comparing the coefficient estimate for dependence in model 1 and model 2, however, reveals that the effect of trade dependence on the log odds of conflict is much greater in the full sample than in the sample constrained by the observed data. Although this model is very simplistic and lacks many key variables emphasized by researchers in this tradition, this comparison of the two samples demonstrates that missing data considerations can yield real differences in empirical results.

Listwise deletion of rows due to missing economic data will also determine whether observations from the conflict data are included in the final sample. There are

7. In particular, the COW system membership list after 1945 includes many states only after these have become members of the United Nations. Oman, for example, is not included until 1971. San Marino, often considered the world's oldest existing republic, is included as a system member by COW only after 1992. See Gleditsch and Ward (1999) for a more extended discussion of the differences.

TABLE 6
Comparison of Data for Organization for Economic Cooperation and Development (OECD) and Council for Mutual Economic Assistance (CMEA) Dyads

	OECD Dyads		*CMEA Dyads*	
Origin Category	*Number of Observations*	*Share*	*Number of Observations*	*Share*
0	7,349	76.32	304	13.41
1	59	0.61	58	2.56
2	1,678	17.43	1,739	76.71
3	324	3.36	14	0.62
4	0	0	59	2.60
5	113	1.17	58	2.56
6	106	1.10	35	1.54
Total	9,629	99.99	2,267	100.00

a total of 2,392 dyad years with MIDs among states in the international system over the period from 1950 to 1992. If one goes with simply the observed economic data,[8] as many as 1,552 of these dyad years fall out of the sample altogether due to missing data on one of the raw variables that go into the construction of the dependence scores. Some of the MIDs that drop out include key events in world politics such as the Berlin crisis, the Korean War, the Cuban Missile Crisis, and the Gulf War. Especially troubling is the fact that the attrition appears to be systematic. Dyads involving socialist and developing states are particularly likely to suffer from missing data.

Even though the sign of the coefficient for the dependence ratio does not change between the directly available sample and the expanded sample, the limitations imposed by simply going with the available data can be very severe in analyses of interactions in particular dyads or subgroups of states. Table 6 compares the availability of data for dyads composed of OECD members and dyads of CMEA members. As can be seen, although data are available for most of the population of OECD dyads (76.32%), only 13.41% of the CMEA dyad years are included in the IMF and PWT data. However, for a large number of the CMEA dyads, data are available from the alternative sources. Adding alternative estimates (i.e., origin categories 1 and 2) ensures data for about 95% of all the country years.

MIDs occur between OECD members in 44 dyad years and between CMEA members in 15 dyad years. However, none of the MIDs between CMEA states occur for observations available in the PWT or IMF data. A simple comparison of the available data could thus suggest that putatively independent CMEA states, unlike OECD states, have *no* conflict between each other. The results based on the expanded GDP and trade data look dramatically different. Given the total number of dyad years, dyads with two CMEA member states have a slightly higher incidence of MIDs than do dyads with two OECD members. A logistic regression of MIDs on the smaller dependence ratio and distance between capital cities indicates a negative, though not significant, coeffi-

8. No observations are missing for the data on capital distances between states.

cient estimate for the dependence ratio among dyads of OECD states but a significant, positive, coefficient estimate for the dependence ratio among dyads of two CMEA states.[9] This suggests that greater economic dependence among the communist and Soviet-dominated CMEA states during the cold war actually was positively associated with conflict. Merely using the available economic data for the CMEA countries would have left a sample without any conflicts between these states.

Much of the interest in the trade and conflict literature tries to address the prospects for peace in a world of increasing globalization and democratization. To untangle the potential effects of economic and political institutions on the likelihood of conflict, as well as where these may or may not apply, comprehensive data with the full range of the actual variation in such attributes are needed. The ability to say something about the prospects for peace in dyads involving countries undergoing political and economic transformations is clearly not bolstered if such cases tend to be deleted from analyses.

CONCLUSIONS

This article has outlined various coverage problems in common sources of economic data and demonstrated that these can have problematic consequences for empirical research. I have indicated a set of procedures for deriving estimates to expand existing data on GDP and economic trade flows. I release the data in the hope that they may be helpful to scholars. These estimates may well contain errors and could be problematic for many purposes. The current version of the data is intended as one of many steps toward improving data resources in international relations research. Comments and suggestions that may help improve the estimates in the data will be appreciated.

9. These results are not shown due to space considerations but are available on request.

APPENDIX
Gross Domestic Product (GDP) Per Capita and Population Estimates as Proportion of Reference Country

TABLE A1
Developing Socialist Economies Compared to Ethiopia

Country	*GDP Per Capita Proportion*	*Population Proportion*
Afghanistan	1.4286	0.4327
Cambodia	1.2500	0.1948
Cuba	2.7857	0.1859
Korea, People's Republic of	1.7857	0.3584
Vietnam, Democratic Republic of[a]	3.1607	1.2954
Vietnam, Republic of	1.5804	0.6477

a. Before 1975, North Vietnam is assumed to be approximately half the population of postwar Vietnam.

TABLE A2
Small Developing Countries Compared to Jamaica

Country	*GDP Per Capita Proportion*	*Population Proportion*
Antigua and Barbuda	2.3939	0.0241
Brunei	5.1515	0.1218
Federated States of Micronesia	0.5333	0.0498
Kiribati	0.2424	0.0324
Marshall Islands	0.4394	0.0249
Maldives	0.5576	0.1131
Nauru	3.0303	0.0041
Tuvalu	0.2424	0.0041

TABLE A3
Small Industrialized Countries Compared to Switzerland

Country	*GDP Per Capita Proportion*	*Population Proportion*
Andorra	0.6818	0.0091
Liechtenstein	0.8712	0.0044
Monaco	0.9470	0.0044
San Marino	0.7576	0.0034

TABLE A4
Industrialized Socialist Economies Compared to Hungary

Country	*GDP Per Capita Proportion*	*Population Proportion*
Albania	0.2014	0.3304
Armenia	0.3649	0.3347
Azerbaijan	0.2216	0.7764
Belarus	0.7027	1.0212
Bosnia-Herzegovina	0.2324	0.3418
Croatia	0.6892	0.4592
Estonia	0.7432	0.1383
Georgia	0.0297	0.4973
Kazakhstan	0.4189	1.6518
Kyrgyzstan	0.2973	0.4463
Latvia	0.5541	0.2311
Lithuania	0.6622	0.3523
Macedonia	0.1419	0.1986
Moldova	0.2973	0.4380
Slovenia	1.3919	0.1934
Tajikistan	0.1338	0.5992
Turkmenistan	0.2203	0.4286
Ukraine	0.2973	4.8901
Uzbekistan	0.3378	2.3662

TABLE A5
Developing African States Compared to Guinea

Country	*GDP Per Capita Proportion*	*Population Proportion*
Equatorial Guinea	1.2712	0.0617
São Tomé and Principe	0.9322	0.0204
Zanzibar	0.6186	0.1035

TABLE A6
Large Developing States Compared to Pakistan

Country	*GDP Per Capita Proportion*	*Population Proportion*
Lebanon	2.2500	0.0258
Libya	3.3500	0.0361
Tibet	1.2395	0.0145
Yemen, People's Republic of	0.3700	0.0215

REFERENCES

Barbieri, Katherine. 1996. Economic interdependence: A path to peace or a source of interstate conflict? *Journal of Peace Research* 33:29-49.

Central Intelligence Agency. 1998. *World factbook*. Retrieved from http://www.cia.gov/cia/publications/factbook/index.html.

Faber, Jan, and Tom Nierop. 1989. *World export data, 1948-1983* (ICPSR no. 9116). Amsterdam: University of Amsterdam (producers); Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributors).

Gleditsch, Kristian S., and Michael D. Ward. 1999. A revised list of independent states since 1816. *International Interactions* 25:393-413.

International Monetary Fund. 1997. *Direction of trade statistics: Codebook* (ICPSR no. 7628). Washington, DC: International Monetary Fund.

Jones, Daniel M., Stuart A. Bremer, and J. David Singer. 1996. Militarized interstate disputes, 1816-1992: Rationale, coding rules, and empirical applications. *Conflict Management and Peace Science* 15:163-213.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review* 95:49-69.

Oneal, John R., and Bruce M. Russett. 1999. Assessing the liberal peace with alternative specifications: Trade still reduces conflict. *Journal of Peace Research* 36:423-42.

Rozanski, Jerzy, and Alexander Yeats. 1994. On the (in)accuracy of economic observations: An assessment of trends in the reliability of international trade statistics. *Journal of Development Economics* 44:103-30.

Rubin, Donald. 1976. Inference and missing data. *Biometrika* 63:581-92.

Russett, Bruce M., and John R. Oneal. 2001. *Triangulating peace: Democracy, interdependence, and international organizations*. New York: Norton.

Schneider, Gerald, and Katherine Barbieri. 1999. Special issue on trade and conflict. *Journal of Peace Research* 36:387-489.

Summers, Robert, and Alan Heston. 1991. The Penn World Table (mark 5): An expanded set of international comparisons, 1950-1988. *Quarterly Journal of Economics* 106:327-68.