

README file for exptradegdpv4.1

First release: 21 July 2004  
This version: 20 August 2004

Kristian Skrede Gleditsch  
Department of Political Science  
University of California, San Diego  
La Jolla, CA 92093-0521, USA  
kgleditsch@ucsd.edu

This archive contains the following files.

- 1) trade\_dd.asc
- 2) trade\_udd.asc
- 3) pwt\_v61.asc
- 4) cowfilter.pl
- 5) uddtrade\_cow.asc
- 6) pwt\_cow.asc
- 7) expopen.asc
- 8) expdata.asc

See below for additional information on files 1-6 and origin codes for the trade and GDP data

The data themselves are described in greater detail in the accompanying paper:

Gleditsch, Kristian S. 2002. "Expanded Trade and GDP data."  
Journal of Conflict Resolution 46(5):712-24.

The Perl code to generate these data is available in a separate archive.

Users are warned that the code is only sparsely documented and may be platform dependent. The current data were generated on a system running RedHat Linux 7.2

For descriptions of the variables and variable names used in files, see the below explanations of the

Brief summary of changes from Version 3.0 to 4.0/4.1

- 
- 1) Version 4.0 incorporates new estimates for trade with Taiwan, which is not included altogether in the IMF data. These estimates are taken from two sources, 1971 to 2002 data from the Directorate General of Budget Accounting and Statistics Executive Yuan of the ROR (Taiwan), and 1952 to 1974 data from the Statistical Yearbook of ROC (Taiwan). These data were collected and converted into \$US by Zhiming Chen and Bruce M. Russett of Yale University.
  - 2) Version 4.0 fixes a problem in Version 3.0 of the data where some of the missing initial leads for certain observations were replaced by interpolated values rather than deflated estimates based on the first-non

missing flows

3) IMF lists values for Vietnam, without distinguishing whether values refer to North and South Vietnam before unification. The hard copy of the IMF yearbook for 1977, however, states on p. 262 that "values through 1974 refer to the former South Vietnam". In this new version, it is therefore assumed that all data for Vietnam up through 1974 refer to South Vietnam. I am grateful to Bruce M. Russett for locating this information in the hard copies of the IMF yearbook.

4) The GDP data have been updated with base data from the new 6.1 version of the Penn World Tables

5) Version 4.1 corrects a problem in the 4.0 version of the GDP data that led to incorrect estimates for many countries not in the PWT 6.1 data with numerical numbers higher than 450

6) The 20 August 2004 version of the 4.1 data introduce two corrections to the file file expdata.asc. The 21 July 2004 did not incorporate the changes to the 4.1 version of the GDP data noted under 5) or data on n gdp and population for the years 1998-2000.

[Changes from Version 2.1 to 3.0 are summarized at the end of the file]

#### Variable descriptions

-----  
Note:

- The trade data are given in millions of current year US dollars
- The population figures are in 1000s
- The real GDP per capita figures are real figures in constant US dollars (base 1996)

acra/b	Three letter acronym for state A/B
numa/b	Numeric id code for state A/B
year	Year
expab	Export A to B
eabo	Export A to B origin code
impab	Import A from B
iabo	Import A from B origin code
expba	Export B to A
ebao	Export B to A origin code
impba	Import B from A
ibao	Import B from A origin code
popa	Population of A
rgdpca/b	Real GDP per capita of A/B
cgdpcab	Current GDP per capita of A/B
goa/b	Origin of GDP per capita estimate A/B
tottra/b	Total trade for country A/B

Table 1: GDP and population data categories

Code	Data origin
0	Observed data from the Penn World Tables data 6.1
-1	Observed data point from Penn World Tables data 5.6, with no corresponding point in PWT 6.1 (a)
1	Lags and leads based on first non-missing observations, deflated to current prices
2	Interpolated estimates (a)
3	Estimate based on figures from the CIA World Factbook

Notes: (b) Some observations have missing data on one but not all of the the three figures from the Penn World Tables. PWT 6.1 figures are used whenever available, but the -1 code is used whenever any of the figures are taken from the PWT 5.6 data

(a) The previous version had no missing observations within time series in the GDP data. This is not the case for these data. All of these gaps are between observations from PWT 5.6 (i.e., -1s) and PWT 6.1 (i.e., 0s)

Table 2: Export from A to B data categories

Code	Data origin	Share of observations
0	Observed data from IMF Direction of Trade Data	
1	Observed data from World Export Database	
1.1	Observed data from the Statiscal Yearbook of ROC (Taiwan)	
1.2	Observed data from Directorate General of Budget Accounting and Statistics Executive Yuan of the ROR (Taiwan)	
2	Estimate based on B's imports from A (observed data from DOT)	
3	Estimate based on A's import from B	
4	Interpolated estimate	
5	Lags and leads based on first non-missing observations, deflated to current prices	
6	Dyads with no observed data, assumed to be 0	

Table 3: Import by A from B data categories

Code	Data origin	Share of observations
0	Observed data from IMF Direction of Trade Data, c.i.f. figures	
1	Observed data from IMF Direction of Trade Data, f.o.b. figures	
1.1	Observed data from the Statiscal Yearbook of ROC (Taiwan)	
1.2	Observed data from Directorate General of Budget Accounting and Statistics Executive Yuan of the ROR (Taiwan)	
2	Estimate based on B's exports to A (observed data from DOT or WED)	
3	Estimate based on A's export to B	
4	Interpolated estimate	
5	Lags and leads based on first non-missing observations,	

deflated to current prices  
6 Dyads with no observed data, assumed to be 0

Supplementary information on files

1) trade\_dd.asc

-----

This file provides trade flows for directed dyads. Each pair of states A and B will give rise to two rows, one for the flow from A to B in that year, another for the flow from B to A in that year

2) trade\_udd.asc

-----

This file provides trade flows for undirected dyads. Each pair of states A and B will give rise to a single row only, with export and imports from A to B and exports and imports from B to A. The data are sorted on the numeric country id. Each dyad is included as AB whenever country number of A < country number of B, as BA otherwise.

3) pwt\_v6.asc

-----

This file contains GDP per capita and population data

4) cowfilter.pl

-----

Files 1-3 are based on the definition of the population of independent states in the international system after 1945 based on the criteria described in

Gleditsch, Kristian S., and Michael D. Ward. 1999.  
"A Revised List of Independent States Since 1816."  
International Interactions 25:393-413.

Note that the data here also incorporate the supplementary list of microstates with fewer than 250,000 inhabitants.

Observations for some states and time periods in these data will not be considered states in the international system as defined by the Correlates of War project's system membership list.

The file cowfilter.pl contains Perl code to generate files 5 and 6 below, which provide undirected trade flows and GDP data on the population of states and dyads as defined by the COW project. The code can easily be change to modify the directed dyad data

5) uddtrade\_cow.asc

-----

Alternate version of file 2, adhering to COW conventions

6) pwt\_cow.asc

-----  
Alternate version of file 3, adhering to COW conventions

7) expopen.asc

-----  
Total trade measures for each state and years, calculated from the expanded trade data, using all categories

8) expdata.asc

-----  
File with trade, gdp, population, and openness measures for all undirected dyads

Brief summary of changes from Version 2.1 to 3.0

-----  
1) Version 3.0 contains trade data for 1997-2000, which Professor Andrew G. Long of the Department of Political Science, University of Mississippi generously has made available.

Long collected data on trade figures valued in millions of current US dollars for 1997-2000 from the IMF Direction of Trade (DOT) Yearbook for 2002. The data were originally recorded in a spreadsheet file, containing Exports from A to B and from B to A and Imports by A from B and by B from A, along with a dichotomous variable recording whether the import data were valued f.o.b. or c.i.f. Gleditsch then modified the original Perl code to merge the updated 1997-2000 IMF DOT data onto the original 1948-1996 IMF DOT trade data, before carrying out the subsequent operations described in Gleditsch (2002). Note that trade data were not collected for any dyads involving certain microstates with very small populations, more specifically Barbados, Dominica, Grenada, St. Lucia, St. Vincent and the Grenadines, Antigua and Barbuda, St. Kitts-Nevis, Seychelles, Vanuatu, Marshall Islands, Palau, Federated States of Micronesia, and Western Samoa. As a consequence, the 1997-2000 figures reported for dyads involving these states are either category 5 (deflated values of the last non-missing observation) or 6 (observations set to 0).

The 1997 to 2000 data were physically entered by hand, and we cannot guarantee that the data are free from typographical mistakes. We ask that any user noticing any potential problems or mistakes in the 1997 to 2000 figures report these to Andrew Long (aglong@olemiss.edu) and Kristian Skrede Gleditsch (kgleditsch@ucsd.edu).

2) Version 3.0 of the data incorporate certain changes to the Gleditsch and Ward list of independent states that the expanded trade and GDP data are based on. The principal change is that the country label for the United Kingdom has changed

from UK to UKG in May 2003, in order to comply with the three letter acronym convention used for other states in these data. The other changes are minor changes to a country's date of independence.

See <http://weber.ucsd.edu/~kgledits/statelist.html> for additional information on the Gleditsch and Ward data on independent state and the full case description list.

3) The Perl code used to create version 2.1 of the data generated incorrect figures for some trade flows involving the United Kingdom. These problems were due to the two letter country code, and do not appear in Version 3.0 of the data.

4) The cowfilter.pl code has been updated to reflect the recent changes in the COW system membership list and its labels. Notably, the COW project has changed its label for Russia/Soviet Union from URS to RUS, changed the label for the Democratic Republic of Congo from ZAI to DRC, and changed the label for Rumania from RUM to ROM.