

## Data Appendix

### Variables used in the Cross-Country Regression Analysis

#### I. Dependent Variables

##### GDP per capita:

PPP-adjusted GDP per capita, in 1950 and 1990 from Maddison (1995, Tables D1 and F4).

PPP-adjusted GDP per capita, in 1995 from World Bank (1998). For countries missing 1995 World Bank data, the data are from CIA (1996) (or from CIA, 1997, as noted): Afghanistan; Albania; Bosnia and Herzegovina (CIA 1997); Bhutan (CIA 1997); Brunei (CIA 1997); Cambodia; Cuba; Djibouti (CIA 1997); Equatorial Guinea (CIA 1997); Eritrea; French Guiana (CIA 1997); Iraq; Korea Dem. People's Rep.; Kuwait; Liberia; Libya Arab Jamahiriyy; Myanmar (CIA 1997); Somalia; Sudan; Taiwan (CIA 1997); Tanzania; The former Yug. Rep. of Macedonia; Yugoslavia (Serbia/Montenegro). Data for additional countries shown in Figure 1 are from CIA (1997).

##### GDP growth:

Instantaneous growth rate of PPP-adjusted GDP per capita from 1965 to 1990 from the Penn World Tables 5.6 (Summers and Heston, 1994).

#### II. Transport Cost and Market Proximity Measures

##### Lt100km:

The proportion of a country's total land area within 100 km. of the ocean coastline, excluding coastline in the arctic and sub-arctic region above the winter extent of sea ice (NGS, 1995). Calculated from digital coastlines in ArcWorld Supplement (ESRI, 1996a).

##### Lt100cr:

The proportion of a country's total land area within 100 km. of the ocean or ocean-navigable river, excluding coastline above the winter extent of sea ice and the rivers that flow to this coastline. Rivers were classified as ocean-navigable mainly according to descriptions in Rand McNally (1996), Britannica Online (1998), and Encyclopedia Encarta (1998). Precise information on our classification of river systems is available from the authors. Ocean-navigable rivers are displayed in Figure 9. Calculated from digital coastlines in ArcWorld Supplement database (ESRI, 1996a) and rivers in the ArcAtlas database (ESRI 1996b).

##### Pop100km:

The proportion of the population in 1994 within 100 km. of the coastline

(as defined for Lt100km). The data for population distribution in 1994 come from the first detailed world GIS population dataset (seen in Figure 2) described in Tobler, et al. (1995).

Pop100cr:

The proportion of the population in 1994 within 100 km. of the coastline or ocean-navigable river (as defined for Lt100cr). The population data are as for Pop100km.

CoastDensity:

Coastal Population/Coastal km<sup>2</sup> = (Population \* Pop100km)/(Land Area \* Lt100km). Units: persons per square kilometer.

InteriorDensity:

Interior Population/Interior = (Population \* (1-Pop100km))/(Land Area \* (1-Lt100km)). Units: persons per square kilometer.

Landlocked, not in Europe:

Indicator for landlocked country, excluding countries in Western and Central Europe (Austria, the Czech Republic, Hungary, the Former Yugoslav Republic of Macedonia, Slovakia, and Switzerland). Includes Eastern European countries of Belarus and Moldova.

LDistance:

The log of the minimum Great-Circle (air) distance in kilometers to one of the three capital-goods-supplying regions: the U.S., Western Europe, and Japan, specifically measured as distance from the country's capital city to New York, Rotterdam, or Tokyo.

CIF/FOB shipping cost margin:

The ratio of CIF import prices to FOB import prices as a measure of transport costs from IMF data (Radelet and Sachs, 1998).

### III. Other Geographical Variables

Tropicar:

The proportion of the country's land area within the geographical tropics. Calculated from ArcWorld Supplement database (ESRI, 1996a).

Malaria Index in 1966:

Index of malaria prevalence based on a global map of extent of malaria in 1966 (WHO, 1967), and the fraction of *falciparum* malaria. The fraction of each country's land area subject of malaria was calculated from digitized 1967 map shown in Figure 5 ("limited risk" areas excluded). The intensity of malaria is captured by the fraction of malaria cases that are the malignant *P. falciparum* species of malaria in 1990 (WHO, 1992). For African countries without 1990 *falciparum* data, we used the WHO (1997b) data (in which almost all African countries with malaria are described as "predominantly" *falciparum*, which we classified as 100%). The index is the product of the fraction of land area subject to malaria times the fraction of *falciparum* malaria cases.

Malaria Index in 1994:

Constructed in the same way as the malaria index for 1966, based on a global malaria map for 1994 (WHO, 1997a), and fraction of *falciparum* malaria in 1990.

Hydrocarbons:

Hydrocarbon deposits are the log of total BTUs per person of proven crude oil and natural gas reserves in 1993 from WRI (1996).

Southern Hemisphere:

Indicator for countries wholly below the equator, as well as Brazil, Democratic Republic of the Congo (Zaire), Republic of the Congo, Ecuador, Gabon, Indonesia, and Kenya.

Land Area:

Area in square kilometers from World Bank (1997), except for Taiwan and Mexico from CIA (1997), with submerged land subtracted out.

#### IV. Other Economic, Social, and Political Variables

Openness:

The proportion of years that a country is open to trade during 1965-90, by the criteria in Sachs and Warner (1995b). A country is considered to be open if it meets minimum criteria on four aspects of trade policy: average tariffs must be lower than 40 percent, quotas and licensing must cover less than 40 percent of total imports, the black market premium must be less than 20 percent, and export taxes should be moderate.

Public Institutions:

The quality of public institutions is based on an index created by Knack and Keefer (1995), which is itself an average of five indicators of the quality of public institutions, including (a) the perceived efficiency of the government bureaucracy, (b) the extent of government corruption, (c) efficacy of the rule of law, (d) the presence or absence of expropriation risk, and (e) the perceived risk of repudiation of contracts by the government. Each country is scored on these five dimensions on the basis of surveys of business attitudes within the countries. The subindexes on the five measures are then summed to produce a single, overall index that is scaled between 0 and 10.

New State:

The timing of national independence (0 if before 1914; 1 if between 1914 and 1945; 2 if between 1946 and 1989; and 3 if after 1989) from CIA (1996).

Socialism:

A variable equal to 1 if the country was under socialist rule for a considerable period during 1950 – 1995 based on Kornai (1992).

Life expectancy at birth, 1965:

Data from United Nations (1996).

Years of secondary schooling, 1965:

Data from Barro and Lee (1993).

Urbanization:

% of population living in urban areas, 1995, from World Bank (1998).

War-torn:

Indicator for countries that participated in at least one external war over the period, 1960-85, from Barro and Lee (1994), with additional countries classified by authors.

Population:

Total population in millions from World Bank (1997).

### **One degree by one degree population database:**

The data for population in 1994 come from the first detailed world GIS population dataset (seen in Figure 2) described in Tobler, et al. 1995. We aggregated the 5' by 5' cells to 1° by 1° cells creating approximately 14,000 observations. The world population distribution in 1800 comes from McEvedy and Jones (1978), mostly on a country-wide basis. The geographical data come from a variety of sources. Incidence of malaria for each 1° cell was digitized from a WHO (1997a) map for 1994. Distance of each cell from the coast was calculated from the ArcWorld coastal boundaries (ESRI, 1992). These boundaries were edited to remove the coasts north of the winter extent of sea ice as in Lt100km above. Ocean navigable rivers leading to the sea were classified as in Lt100cr as above. Inland rivers are rivers classified as navigable by ArcAtlas (ESRI, 1996b), but with no outlet to the sea, as well as rivers navigable to the sea, but not navigable by oceangoing vessels. Elevation data are derived from the ETOPO world elevation database (NOAA, 1988). Land used from rice-growing was derived from the ArcAtlas database on agriculture (ESRI 1996b). Soil depth and stream density (a count of the streams in each 1° cell from satellite data) come from NASA (Sellers, et al., 1997). Soil suitability for rainfed crops was derived from Digital Soils Map of the World (FAO, 1995). A classification of land areas into thirty-seven ecozones comes from the UNEP (Leemans, 1990).