

Datasets for empirical analysis referred to in the paper
"Do Constitution Cause Large Governments? Quasi-experimental
evidence"

by Torsten Persson and Guido Tabellini (2001)

Description of the Variables in EEA Data.xls and EEA Data.dta

ctrycd = IMF country code.

country = Country name.

GEOGRAPHIC AND INSTITUTIONAL EXPLANATORY VARIABLES:

colonial origins

col_(es, fr, or uk): dummy variable taking the value 1 if the country has been a colony of Spain (or Portugal) (ES), France (FR) or United Kingdom (UK) for a significant time and 0 otherwise.
Source: Wacziarg (1996).

coloth: residual variable allowing for colonizers other than those mentioned above. It is computed as $coloth = evercolony - col_es - col_uk - col_fr$. For those countries that had more than one significant colonizer (i.e. Israel, Jordan, Luxembourg, Namibia, Papua New Guinea, Philippines, Singapore and South Africa) this variable takes the value 1 regardless the fact that col_(es, fr or uk) may take the value 1 too (i.e. one of the two colonizers is Spain, France or England).
Source: Wacziarg (1996).

t_indep: years of independence of the country since 1748. They are computed subtracting the year of independence from 1998 (the year of independence is included in the computation of t_indep). 250 (i.e. 1998-1748) is the default time of independence for non-colonies Countries. (Note we considered 250 even for the USA).
Source: Wacziarg (1996).

Colonial exposition:

$col_espa = col_es * (250 - t_indep) / 250$

$col_uka = col_uk * (250 - t_indep) / 250$

col_othafr = (col_oth+col_fr) * (250-t_indep)/250

Population

prop65: percentage of population over the age of 65 in the total population.

Constructed as the population over 65, divided by the total population (both reported by WDI) and multiplied by 100.

Source: World Development Indicators CD-Rom 1999.

prop1564: percentage of population between 15 and 64 years old in the total

population. Constructed as the population between 15 and 64, divided by the total population (both reported by WDI) and multiplied by 100.

Source: World Development Indicators CD-Rom 1999.

ECONOMIC AND SOCIAL EXPLANATORY VARIABLES:

gdp(blns): gross domestic product as defined in the Introduction of International Financial Statistics.

Source: IFS January 2001 (for years 1993-98), and IFS Yearbook 1998 (for years 1990-92).

rgdph: real GDP per capita in constant dollars (international prices, base year 1985)

Data through 1992 are taken from the Penn World Table 5.6, while data on the period 1993-1998 are taken from the Easterly's series available on the World Bank's web

site. For those countries (namely Barbados, Bulgaria, Cyprus, Czech Rep., Fiji, Honduras, Hungary, Latvia, Nicaragua, Papua New Guinea, Slovak Rep., Sri Lanka, St. Vincent & G., Thailand, Turkey) for which observations on the last few periods are

not available, data are calculated from the previous rgdph and GDP per capita growth rates. GDP per capita growth rates are computed on the basis of Easterly's

methodology as $[(1+GDP\ growth)/(1+POP\ growth)]-1$ * 100 relying on the series

GDP at market prices (constant 1995 US \$) and Population, total of the world

Development Indicators.

Definitions of:

GDP at market prices (constant 1995 US \$): GDP measures the total

EEA Data.txt

output of goods and services for final use occurring within the domestic territory of a given country, regardless of the allocation to domestic and foreign claims. Gross domestic product at purchaser prices is the sum of gross value added by all resident producers in the economy plus any taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. The residency of an institution is determined on the basis of economic interest in the territory for more than a year. Data are in constant 1995 U.S. dollars. Dollar figures for GDP are converted from domestic currencies using 1995 official exchange rates. For a few countries where the official exchange rate does not reflect the rate effectively applied to actual foreign exchange transactions, an alternative conversion factor is used.

Population, total: Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. Refugees not permanently settled in the country of asylum are generally considered to be part of the population of their country of origin.

For those Countries for which rgdph is not available (i.e. Belarus, Estonia, Germany, Russia and Ukraine), data have been imputed in the following way:

where:

USAtPW is the U.S. real GDP per capita in constant dollars, international prices 1985

base year at time t (Source Penn World Table);

USAtWDI is the U.S. GDP at market prices (constant 1995 US \$) divided by the total

population at time t; therefore, it represents the U.S real GDP per capita in constant

dollars 1995 at time t obtained by the combination of the two series here above that

belong to the WDI (Source WDI - CD ROM 2000);

ZtWDI, as above, is the real GDP per capita in constant dollars 1995 at time t for

Country Z obtained dividing the GDP at market prices (constant 1995 US \$) by the total

population (Source WDI - CD ROM 2000).

Thereby, we impute the real GDP per capita in constant dollars, international prices-

1985 base year at time t for Country Z multiplying the available WDI series - real

GDP per capita in 1995 constant dollars - by the american price

EEA Data.txt

ratio (i.e. 1\$ of 1995 corresponding to x\$ of 1985).

Sources: Penn World Tables - mark 5.6 (PW); Easterly's series on www.worldbank.org; The World Bank's World Development Indicators (WDI).

$l_{yp} = \ln(\text{rgdph})$

trade: trade is the sum of exports and imports of goods and services measured as a share of gross domestic product.

Source: The World Bank's World Development Indicators (CD-ROM 2000).

Taiwan was completed using the same series from Easterly's dataset.

A few missing observations (i.e. some countries in some years) have been imputed using some Easterly's series according to this formula:

$$\frac{[\text{Exports of goods and services (BoP, current US\$)} + \text{Imports of goods and services (BoP, current US\$)}]}{\text{GDP at market prices (current US\$)}}$$

where:

Exports of goods and services (BoP, current US\$): Exports of goods and services comprise all transactions between residents of a country and the rest of the world involving a change of ownership of general merchandise, goods sent for processing and repairs, nonmonetary gold, and services. Data are in current U.S. dollars.

Imports of goods and services (BoP, current US\$): imports of goods and services comprise all transactions between residents of a country and the rest of the world involving a change of ownership of general merchandise, goods sent for processing and repairs, nonmonetary gold, and services. Data are in current U.S. dollars.

GDP at market prices (current US\$): GDP measures the total output of goods and services for final use occurring within the domestic territory of a given country, regardless of the allocation to domestic and foreign claims. Gross domestic product at purchaser values (market prices) is the sum of gross value added by all resident and nonresident producers in the economy plus any taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars. Dollar figures for GDP are converted from domestic currencies using single year official exchange rates. For a few countries where the official exchange rate does not reflect the rate effectively applied to actual

foreign exchange transactions, an alternative conversion factor is used.

Source: Easterly database available on line at www.worldbank.org

federal: federalism dummy. Source: Boix 2000.

POLITICAL EXPLANATORY VARIABLES:

maj: dummy variable taking the value 1 in presence of either a majority or a plurality rule, 0 otherwise. Only legislative elections are considered. Sources: Cox (1997), International Institute for Democracy and Electoral Assistance (1997), Quain (1999), and Kurian (1998).

pres: dummy variable taking the value 1 in presence of presidential regimes, 0 otherwise (Parliamentary). Only those regimes in which the confidence of the assembly is not necessary to the executive branch (even if the president is not chief of the executive, i.e. assembly-independent) are accepted under the definition of presidential form of government. Premier-presidential (semi-presidential like France) and president-parliamentary systems (like Ecuador) are excluded. Sources: Shugart and Carey (1992) and Kurian (1998).

gastil: average of gastil index for civil liberties and political rights. It is measured on one-to-seven scale with one representing the highest degree of freedom and seven the lowest. Countries whose combined averages for political rights and for civil liberties fall between 1.0 and 2.5 are designated "free", between 3.0 and 5.5 "partly free" and between 5.5 and 7.0 "not free". Source: Freedom House, Annual Survey of Freedom Country Ratings.

Size of Government

cgexp: central government expenditures as the % of GDP. Constructed as line 82 - Expenditures - under "Government Finance" in the IFS, divided by the GDP (line 99b) and multiplied by 100. Source: IFS January 2001 (for years 1993-98), and IFS Yearbook 1998 (for years 1990-92).

Note: when expenditure is missing or not up-dated in the

mentioned sources,
expenditure (if available) is taken from the GFSY (Government
Financial Statistics
Yearbook - IMF) under Summary Tables - Consolidated Central
Government - Total
Expenditure (C.II) or, under Summary Table - Budgetary Central
Government - Total
Expenditure (C.II) if the latter one is the only one available.

Regional dummies

africa: Regional dummy variable for African countries, taking the
value 1 if a country
is African, 0 otherwise. Source: Persson and Tabellini (1999).

asiae: Regional dummy variable for East Asian countries, taking
the value 1 if a
country is East Asian, 0 otherwise. Source: Persson and Tabellini
(1999).

laam: Regional dummy variable for Latin American countries,
taking the value 1 if a
country is Latin American, 0 otherwise. Source: Persson and
Tabellini (1999).

oecd: Dummy variable for OECD member countries, taking the value
1 if a Country
is an OECD member, 0 otherwise. Source: Persson and Tabellini
(1999).

euromide: Non-OECD European and Middle East country dummy.

REFERENCES FOR SOURCES

Boix, C. (2000), "Democracy, development and the public sector",
mimeo, University
of Chicago.

Cox, G. (1997), "Making Votes Count", Cambridge University Press.

International Institute for Democracy and Electoral
Assistance(1997), "Handbook of
Electoral System Design", Stockholm, Sweden.

Kurian, G. ed. (1998), "World Encyclopedia of Parliaments and
Legislatures", Fitzroy
Dearborn Publishers.

Persson, T. and G. Tabellini (1999), "The Size and Scope of
Government: Comparative
Politics with Rational Politicians, 1998 Alfred Marshall
Lecture", European Economic

Review 43: 699-735.

Quain, A. ed (1999), "The Political Reference Almanac", 1999/2000 Edition, Keynote Publishing Co. US. Available at www.polisci.com.

Shugart and Carey (1992), "Presidents and Assemblies: Constitutional Design and Electoral Dynamics", Cambridge University Press.

Wacziarg, R. (1996), "Information to Create Colonization Dummies", mimeo, Harvard University.