

COMPARATIVE TAXATION DATASET

Elisabeth Anderson,¹ Monica Prasad,² Andre Nickow³

¹New York University Abu Dhabi, Division of Social Sciences

²Northwestern University, Department of Sociology and Institute for Policy Research

³Northwestern University, Department of Sociology

August 28, 2018

The construction of this dataset was made possible by NSF grant #0847725.

MYPE – Mid-Year Population Estimates measured in thousands. Source: International Historical Statistics (IHS), 2003.

The estimates relate to population actually present at mid-year, except where otherwise indicated.

CE – Collver's Estimates of Average Vital Rates for Quinquennia. Source: International Historical Statistics, 2003. Estimates are separated into Birth (**CEBIRTH**), Death (**CEDEATH**) and Infant Mortality (**CEINMORT**). These figures are from Collver, O. Andrew (1965), *Birth Rates in Latin America: New Estimates of Historical Trends and Fluctuations* (Berkeley Institute of International Studies, University of California, 1965).

INFMORT – Infant Mortality Rates. Source: International Historical Statistics, 2003. The data are for deaths of infants under one year old per 1,000 live births.

EAP – Economically Active Population measured in thousands by Major Industrial Groups. Source: International Historical Statistics, 2003. Industrial sectors are separated by gender; Male (M), Female (F) and Total (T). The groups include Agriculture, Forestry & Fishing (**AGRI**); Extractive (**EXTR**); Manufacturing (**MANU**); Construction (**CONS**); Commerce, Finance, etc. (**CMRC**); Transport & Communications (**TRAN**); Services (**SRVC**); and Others Occupied (**OTHR**).

UNEM – Unemployment. Source: International Historical Statistics, 2003. Values are separated by raw numbers in thousands (**UNEMNUM**) and by % of appropriate workforce (**UNEMPRCT**). These figures rely on a variety of different indicators of unemployment, as indicated in the footnotes. In most cases the statistics are averages of monthly, quarterly, 12-monthly or 4-quarterly annual figures and are for persons over the age of compulsory schooling.

ID – Industrial Disputes. Source: International Historical Statistics, 2003. Values are separated by number of strikes and lockouts (**IDSTRKE**), number of workers involved (in thousands) (**IDWORKRS**), and number of work days lost (in thousands) (**IDDAYS**). Except where otherwise noted, the number of workers involved and the days' work lost as a result of their involvement relate to everyone affected by a dispute, not just to those directly involved.

INDPRD – Indices of Industrial Production. Source: International Historical Statistics, 2003.

INDPRD01 includes manufacturing, mining, and usually electricity and gas. Values for several countries in North and South America are separated by manufacturing (**INDPRD11**) and mining

(**INDPRD21**). Except as indicated in the footnote or headings, **INDPRD11** relates to manufacturing production alone, excluding mining, construction, and usually utilities also. Note that the base years change quite frequently and that IHS crudely splices them together.

INDPRD02, **INDPRD12** and **INDPRD22** represent our recalculations of the IHS data to homogenize base years to the greatest extent possible. This was done by establishing a common base year and converting values of a base year outside the range of the set base year, but with some data within the set base year, by dividing by the ratio between the value for that base year and the value for the set base year.

Original and homogenized base years are given in the notes. *Some base years could not be homogenized due to an absence of overlapping data, so care should be taken to check the notes before using this data.*

Homogenization calculations for each country are as listed below, with **INDPRD12** = I' and **INDPRD11** = I and **INDPRD22** = A' and **INDPRD21** = A. Years in brackets indicate the original base year before the data was homogenized.

Argentina (base year = 1958)

[1958] For 1935 – 1991, I' = I

[1958] For 1935 – 1988, A' = A

Canada (base year = 1958)

[1938] For 1923 – 1944, I' = I/(234/75)

[1958] For 1945 – 1998, I' = I

[1938] For 1924 – 1944, A' = A/(95/34)

[1958] For 1945 – 1998, A' = A

Chile (base year = 1958)

[1938] For 1927 – 1935, I' = I/(89/39)

[1958] For 1936 – 1998, I' = I

[1938] For 1927 – 1935, A' = A/(81/68)

[1958] For 1936 – 1998, A' = A

Uruguay (base year = 1958)

[1958] For 1953 – 1998, I' = I

US (base year = 1958)

[1899] For 1870 – 1914, I' = I/(186/63)/(287/91)

[1938] For 1915 – 1944, I' = I/(287/91)

[1958] For 1945 – 1998, I' = I

[1899] For 1880 – 1914, A' = A/(197/61)/(150/83)

[1938] For 1915 – 1944, A' = A/(150/83)

[1958] For 1945 – 1998, A' = A

Australia (base year = 1970)

[1958] For 1957 – 1964, $I' = I/(142/76)$
[1970] For 1965 – 1974, $I' = I$
[1980] For 1975 – 1993, $I' = I/(89/110)$
[1990] For 1994 – 1999, $I' = I/(100/126)/(89/110)$

Japan (base year = 1970)

[1913] For 1874 – 1930, $I' = I/(250/30)/(352/47)$
[1955] For 1931 – 1964, $I' = I/(352/47)$
[1970] For 1965 – 1974, $I' = I$
[1980] For 1975 – 1993, $I' = I/(81/122)$
[1990] For 1994 – 1999, $I' = I/(100/148)/(81/122)$

New Zealand (base year = 1970)

[1938] For 1930 – 1948, $I' = I/(157/60)/(150/70)$
[1958] For 1949 – 1964, $I' = I/(150/70)$
[1970] For 1965 – 1974, $I' = I$
[1975] For 1975 – 1977, $I' = I/(85/118)$
[1990] For 1994 – 1999, data could not be homogenized

Austria (base year = 1990)

[1913] For 1880 – 1912, data could not be homogenized
[1937] For 1923 – 1949, $I' = I/(123/37)/(300/100)$
[1963] For 1950 – 1993, $I' = I/(300/100)$
[1990] For 1994 – 1999, $I' = I$

Belgium (base year = 1990)

[1937] For 1901 – 1949, $I' = I/(100/60)/(205/100)$
[1963] For 1950 – 1993, $I' = I/(205/100)$
[1990] For 1994 – 1999, $I' = I$

Denmark (base year = 1990)

[1937] For 1927 – 1949, $I' = I/(137/51)/(229/100)$
[1963] For 1950 – 1993, $I' = I/(229/100)$
[1990] For 1994 – 1999, $I' = I$

Finland (base year = 1990)

[1937] For 1900 – 1949, $I' = I/(147/41)/(346/100)$
[1963] For 1950 – 1993, $I' = I/(346/100)$
[1990] For 1994 – 1999, $I' = I$

France (base year = 1990)

[1913] For 1870 – 1913, $I' = I/(57/56)/(112/43)/(225/100)$
[1937] For 1919 – 1949, $I' = I/(112/43)/(225/100)$
[1963] For 1950 – 1993, $I' = I/(225/100)$
[1990] For 1994 – 1999, $I' = I$

Germany (base year = 1990)

[1913] For 1870 – 1913, $I' = I/(100/84)/(90/28)/(226/100)$

[1937] For 1914 – 1949, $I' = I/(90/28)/(226/100)$

[1963] For 1950 – 1993, $I' = I/(226/100)$

[1990] For 1994 – 1999, $I' = I$

East Germany (base year = 1963)

[1937] For 1946 – 1949, $I' = I/(79/23)$

[1967] For 1950 – 1988, $I' = I$

Ireland (base year = 1990)

[1937] For 1926 – 1949, $I' = I/(152/55)/(464/100)$

[1963] For 1950 – 1993, $I' = I/(464/100)$

[1990] For 1994 – 1999, $I' = I$

Italy (base year = 1990)

[1913] For 1870 – 1913, $I' = I/(100/57)/(109/29)/(250/100)$

[1937] For 1914 – 1949, $I' = I/(109/29)/(250/100)$

[1963] For 1950 – 1993, $I' = I/(250/100)$

[1990] For 1994 – 1999, $I' = I$

Netherlands (base year = 1990)

[1937] For 1925 – 1949, $I' = I/(129/42)/(213/100)$

[1963] For 1950 – 1993, $I' = I/(213/100)$

[1990] For 1994 – 1999, $I' = I$

Norway (base year = 1990)

[1937] For 1909 – 1949, $I' = I/(137/51)/(219/100)$

[1963] For 1950 – 1993, $I' = I/(219/100)$

[1990] For 1994 – 1999, $I' = I$

Spain (base year = 1990)

[1913] For 1870 – 1913, $I' = I/(100/63.2)/(97.7/32)/(397/100)$

[1929] For 1914 – 1949, $I' = I/(97.7/32)/(397/100)$

[1963] For 1950 – 1993, $I' = I/(397/100)$

[1990] For 1994 – 1999, $I' = I$

Sweden (base year = 1990)

[1913] For 1870 – 1913, $I' = I/(100/46)/(155/55)/(198/100)$

[1937] For 1914 – 1949, $I' = I/(155/55)/(198/100)$

[1963] For 1950 – 1993, $I' = I/(198/100)$

[1990] For 1994 – 1999, $I' = I$

Switzerland (base year = 1990)

[1963] For 1958 – 1993, $I' = I/(194/100)$

[1990] For 1994 – 1999, $I' = I$

UK (base year = 1990)

[1913] For 1870 – 1914, $I' = I/(93.7/57.4)/(120.7/66)/(171/100)$

[1937] For 1915 – 1949, $I' = I/(120.7/66)/(171/100)$

[1963] For 1950 – 1993, $I' = I/(171/100)$

[1990] For 1994 – 1999, $I' = I$

OUTPETRL – Output of Crude Petroleum. Source: International Historical Statistics, 2003. Values are in thousands of metric tons unless stated otherwise. We have standardized all values in **OUTPETR2** into thousands of metric tons. Output from oil shale is not included. Prior to 1950, France, Italy, and U.K. produced trivial amounts not shown here.

OUTNATGS – Output of Natural Gas. Source: International Historical Statistics, 2003. Values are in millions of cubic meters until 1966, then Petajoules subsequently. We have standardized all values in **OUTNATG2** as follows: for Australia and Japan after 1960, the unit is millions of petajoules. For all other countries/years, the unit is millions of cubic meters. The statistics normally refer to usable natural gas from gasfields only and do not include methane from coalfields.

OUTELEC – Output of Electrical Energy. Source: International Historical Statistics, 2003. Values are in gigaWatt hours, except where otherwise indicated. We have standardized all values in **OUTELEC2** into gigawatt hours. Except where otherwise indicated, statistics relate to gross output and include electricity consumed by power stations. The negligible pre-1895 production figures for Italy are not given.

EXTRD – External Trade Aggregates in Current Values. Source: International Historic Statistics, 2003. Values are separated by imports (**EXTRDIMP**) and exports (**EXTRDEXT**). Units (both monetary and numerical) vary from country to country. We have standardized numerical units into millions in **EXTRDIM2** and **EXTRDEX2**. Except where otherwise indicated, statistics are for merchandise trade only: imports for domestic consumption, and exports of domestic origin plus re-exports of commodities originally for domestic consumption. Imports are normally valued c.i.f., and exports f.o.b. **EXIM%GFD**, **EXIM%GDP**, and **EXIM%GNP** are EXTRDIM2 as a percentage of GFDGDP, IHSGDP, and IHSGNP, respectively. **EXEX%GFD**, **EXEX%GDP**, and **EXEX%GNP** are EXTRDEX2 as a percentage of GFDGDP, IHSGDP, and IHSGNP, respectively.

MONEY – Money Supply. Source: International Historical Statistics, 2003. Values are separated by M1 (**MONEYM1**) and M2 (**MONEYM2**). M1 is notes and coin in the hands of the public plus residents' demand deposits. M2 includes M1 plus time and savings deposits. Non-residents, and banks, deposits are not always excluded in earlier years. Breaks (indicated in notes where possible) may thus indicate improved or changed data collection methods. For Europe, changes in reporting systems and covered institutions are particularly frequent. In IHS, two figures are sometime given for the year in which the break occurred. We put the second figure in a note. We have standardized all numerical units into millions in **MONEYM12** and **MONEYM22**. **MOM1%GFD**, **MOM1%GDP**, and **MOM1%GNP** are MONEYM12 as a percentage of GFDGDP, IHSGDP, and IHSGNP, respectively. **MOM2%GFD**, **MOM2%GDP**, and

MOM2%GNP are **MONEYM22** as a percentage of **GFDGDP**, **IHSGDP**, and **IHSGNP**, respectively.

CENGOVEX – Total Central Government Expenditure. Source: International Historical Statistics, 2003. Monetary units for each value vary from country to country. Numerical units are in millions of stated currency unit, unless otherwise indicated. We have standardized all numerical units into thousand millions in **CENGOVX2**. Statistics are all listed in millions and derived from closed accounts, not budgets. Note that in countries with federal systems of government, expenditures of ‘local’ governments may be at least as important as central government expenditures. All types of central government budgetary expenditure, other than debt redemption, are included. However, extra-budgetary expenditures are also often included, and the types of outlay included vary. Changes are noted where possible. **CGEX%GFD**, **CGEX%GDP**, and **CGEX%GNP** are **CENGOVX2** as a percentage of **GFDGDP**, **IHSGDP**, and **IHSGNP**, respectively.

CGR – Total Central Government Revenue, With Some Main Tax Yields. Source: International Historical Statistics, 2003. Total revenue means total ordinary revenue exclusive of loan receipts, except where otherwise noted. Whether receipts from public enterprises are included (and, if so, whether net or gross) varies from country to country. Changes in consumption are noted where possible. Statistics are from the closed accounts, except where otherwise noted. Values are separated by Total Revenue (**CGRTOTL** and **CGRTOTL2**, the latter standardized by us into millions of stated currency), Customs (**CGRCSTMS** and **CGRCSTM2**, the latter standardized by us into millions of stated currency unit), Internal Revenue (**CGRINTRL** and **CGRINTR2**, in millions of stated currency), Excises (**CGREXCES** and **CGREXCE2**, the latter standardized by us into millions of stated currency), Sales Tax (**CGRSALES** and **CGRSALE2**, the latter standardized by us into millions of stated currency), Income Taxes Only (**CGRINCM** and **CGRINCM2**, the latter standardized by us into millions of stated currency), Income and Wealth Taxes (**CGRINCMW** and **CGRINCW2**, the latter standardized by us into millions of stated currency), Social Security Tax (**CGRSSTAX**, in thousand millions of stated currency and **CGRSSTAX2**, standardized by us into millions of stated currency), Direct Taxes (**CGRDIRCT** and **CGRDIRC2**, the latter standardized by us into millions of stated currency), VAT (**CGRVAT** and **CGRVAT2**, the latter standardized by us into millions of stated currency), Income and Property Tax (**CGRIPT** and **CGRIPT2**, the latter standardized by us into millions of stated currency), Salt and Tobacco Tax (**CGRSTT** and **CGRSTT2**, the latter standardized by us into millions of stated currency), Profit Tax (**CGRPROFT** and **CGRPROF2**, the latter standardized by us into millions of stated currency), Petroleum Tax (**CGRPETRT** and **CGRPETR2**, the latter standardized by us into millions of stated currency), Oil Tax (**CGROIL** and **CGROIL2**, the latter standardized by us into millions of stated currency), Consumption Tax (**CGRCONST** and **CGRCONS2**, the latter standardized by us into millions of stated currency), Tobacco Monopoly (**CGRTOBAM** and **CGRTOBAM2**, in millions of stated currency), Tobacco and Petroleum Monopoly (**CGRPTM** and **CGRPTM2**, the latter standardized by us into millions of stated currency), Automobile Tax (**CGRAUTO** and **CGRAUTO2**, the latter standardized by us into millions of stated currency), Land and Income and Property Tax (**CGRLIPT** and **CGRLIPT2**, the latter standardized by us into millions of stated currency), Turnover Tax (**CGRTURNO** and **CGRTURN2**, the latter standardized by us into millions of stated currency), Salt and Tobacco Monopolies (**CGRSTM** and **CGRSTM2**, in millions of

stated currency), Registration Tax (**CGRREGIT** and **CGRREGI2**, the latter standardized by us into millions of stated currency), and Turnover and Transactions Taxes (**CGRTRNTT** and **CGRTRNT2**, the latter standardized by us into millions of stated currency). **CGRT%GFD**, **CGRT%GDP**, and **CGRT%GNP** are CGRTOTL2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRC%GFD**, **CGRC%GDP**, and **CGRC%GNP** are CGRCSTM2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRI%GFD**, **CGRI%GDP**, and **CGRI%GNP** are CGRINTRL as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRE%GFD**, **CGRE%GDP**, and **CGRE%GNP** are CGREXCE2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRS%GFD**, **CGRS%GDP**, and **CGRS%GNP** are CGRSALE2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRM%GFD**, **CGRM%GDP**, and **CGRM%GNP** are CGRINCM2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRW%GFD**, **CGRW%GDP**, and **CGRW%GNP** are CGRSINCW2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRX%GFD**, **CGRX%GDP**, and **CGRX%GNP** are CGRSSTX2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRD%GFD**, **CGRD%GDP**, and **CGRD%GNP** are CGRDIRC2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRV%GFD**, **CGRV%GDP**, and **CGRV%GNP** are CGRVAT2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRIPT%GFDGDP**, **CGRIPT%IHS GDP**, and **CGRIPT%IHS GNP** are CGRIPT2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRZ%GFD**, **CGRZ%GDP**, and **CGRZ%GNP** are CGRSTT2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRP%GFD**, **CGRP%GDP**, and **CGRP%GNP** are CGRPROF2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRR%GFD**, **CGRR%GDP**, and **CGRR%GNP** are CGRPETR2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRN%GFD**, **CGRN%GDP**, and **CGRN%GNP** are CGRCONS2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRB%GFD**, **CGRB%GDP**, and **CGRB%GNP** are CGRTOBAM as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRL%GFD**, **CGRL%GDP**, and **CGRL%GNP** are CGRPTM2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRU%GFD**, **CGRU%GDP**, and **CGRU%GNP** are CGRAUTO2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRY%GFD**, **CGRY%GDP**, and **CGRY%GNP** are CGRLIP2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRSTM%GFDGDP**, **CGRSTM%IHS GDP**, and **CGRSTM%IHS GNP** are CGRSTM as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRO%GFD**, **CGRO%GDP**, and **CGRO%GNP** are CGRTRN2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRA%GFD**, **CGRA%GDP**, and **CGRA%GNP** are CGRSTM as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. **CGRG%GFD**, **CGRG%GDP**, and **CGRG%GNP** are CGRREGI2 as a percentage of GFDGDP, IHS GDP, and IHS GNP, respectively. Numerical and currency units vary from country to country. Values are in millions of stated currency unit, except as otherwise indicated. **CGRQ%GFD** is CGRTRNT2 as a percentage of GFDGDP. **CGRQ%GDP** is CGRTRNT2 as a percentage of IHS GDP. **CGRQ%GNP** is CGRTRNT2 as a percentage of IHS GNP.

WHLPRCI1 – Wholesale Price Indices. Source: International Historical Statistics, 2003. In a few noted cases, different indices have been spliced together to give a rough indicator of the long-term movement of prices. **WHLPRCI2** represents a recalculation of WHLPRCI1 to

homogenize the base years to the greatest extent possible. This was done by establishing a common base year and converting values of a base year outside the range of the set base year, but with some data within the set base year, by dividing by the ratio between the value for that base year and the value for the set base year. Original and homogenized base years are given in the notes. *Some base years could not be homogenized due to an absence of overlapping data, so care should be taken to check the notes before using this data.*

Homogenization calculations for each country are as given below, with $WHLPRCIN2 = I'$ and $WHLPRCIN1 = I$. Years in brackets indicate the original base year before the data was homogenized.

Argentina (base year = 1980)

[1938] For 1913 – 1964, $I' = I/(12399/100)/(3218/100)/(16075/100)$
[1964] For 1965 – 1975, $I' = I/(3218/100)/(16075/100)$
[1975] For 1976 – 1980, $I' = I/(16075/100)$
[1980] For 1981 – 1984, $I' = I$
[1984] For 1985 – 1990, $I' = I/(100/23002)$
[1990] For 1991 – 1993, $I' = I/(100/8611405)/(100/23002)$
[1995] For 1994 – 1998, data could not be homogenized

Canada (base year = 1980)

[1900] For 1870 – 1938, $I' = I/(163/100)/(501/62)$
[1938] For 1939 – 1975, $I' = I/(501/62)$
[1980] For 1976 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Chile (base year = 1980)

[1938] For 1913 – 1959, $I' = I/(10177/100)/(2357/100)/(40067/100)/(1782/100)$
[1959] For 1960 – 1972, $I' = I/(2357/100)/(40067/100)/(1782/100)$
[1972] For 1973 – 1975, $I' = I/(40067/100)/(1782/100)$
[1975] For 1976 – 1980, $I' = I/(1782/100)$
[1980] For 1981 – 1993, $I' = I$
[1990] For 1994 – 1998, $I' = I/(100/643)$

Uruguay (base year = 1980)

[1970] For 1963 – 1975, $I' = I/(1515/100)/(860/100)$
[1975] For 1976 – 1980, $I' = I/(860/100)$
[1980] For 1981 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

US (base year = 1980)

[1910 – 1914] For 1870 – 1938, $I' = I/(115/100)/(432/65)$
[1938] For 1939 – 1975, $I' = I/(432/65)$
[1980] For 1976 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Australia (base year = 1985)

[1938] For 1871 – 1938, $I' = I/(100/31)/(100/22.7)$

[1970] For 1939 – 1969, $I' = I/(100/22.7)$

[1985] For 1970 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Japan (base year = 1985)

[1869] For 1870 – 1928, $I' = I/(522/83)/(9638/32)/(100/48.6)$

[1938] For 1929 – 1948, $I' = I/(9638/32)/(100/48.6)$

[1970] For 1949 – 1969, $I' = I/(100/48.6)$

[1985] For 1970 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

New Zealand (base year = 1985)

[1938] For 1891 – 1938, $I' = I/(100/28)/(100/15.6)$

[1970] For 1939 – 1969, $I' = I/(100/15.6)$

[1985] For 1970 – 1993, $I' = I$

Austria (base year = 1953)

[1867-1877] For 1870 – 1909, data could not be homogenized

[1929] For 1914 – 1948, $I' = I/(275/42)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Belgium (base year = 1953)

[1913] For 1870 – 1913, data could not be homogenized

[1929] For 1914 – 1948, $I' = I/(289/94)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Denmark (base year = 1953)

[1891-1900] For 1875 – 1914, $I' = I/(145/97)/(236/74)$

[1929] For 1915 – 1948, $I' = I/(236/74)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Finland (base year = 1953)

[1929] For 1913 – 1948, $I' = I/(1009/63)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

France (base year = 1953)

[1901-1910] For 1870 – 1914, $I' = I/(118/17)/(1699/65)$

[1929] For 1915 – 1948, $I' = I/(1699/65)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Germany (base year = 1953)

[1913] For 1870 – 1914, data could not be homogenized

[1929] For 1915 – 1944, data could not be homogenized

[1953] For 1948 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Ireland (base year = 1953)

[1953] For 1938 – 1993, $I' = I$

[1995] For 1994 – 1999, data could not be homogenized

Italy (base year = 1953)

[1913] For 1870 – 1914, $I' = I/(96/20)/(5437/104)$

[1929] For 1915 – 1948, $I' = I/(5437/104)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Netherlands (base year = 1953)

[1929] For 1901 – 1948, $I' = I/(196/74)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Norway (base year = 1953)

[1913] For 1891 – 1914, $I' = I/(105/69)/(188/67)$

[1929] For 1915 – 1948, $I' = I/(188/67)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Spain (base year = 1953)

[1913] For 1870 – 1913, $I' = I/(100/58)/(450/57)$

[1929] For 1914 – 1948, $I' = I/(450/57)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Sweden (base year = 1953)

[1913] For 1870 – 1913, data could not be homogenized

[1929] For 1914 – 1948, $I' = I/(188/72)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1993 – 2000, data could not be homogenized

Switzerland (base year = 1953)

[1914] For 1870 – 1921, $I' = I/(222/142)/(165/102)$

[1929] For 1922 – 1948, $I' = I/(165/102)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

UK (base year = 1953)

[1913] For 1870 – 1914, $I' = I/(101/72)/(193/73)$

[1929] For 1915 – 1948, $I' = I/(193/73)$

[1953] For 1949 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

CONPRCI1 – Consumer Price Indices. Source: International Historical Statistics, 2003. Cost-of-living or retail price indices are given where consumer price indices are unavailable. In some noted cases, indices have been spliced together to roughly indicate the long-term movement of prices. **CONPRCI2** represents a recalculation of CONPRCI1 to homogenize the base years. This was done by establishing a common base year and converting values of a base year outside the range of the set base year, but with some data within the set base year, by dividing by the ratio between the value for that base year and the value for the set base year. Original and homogenized base years are given in the notes. *Some base years could not be homogenized due to an absence of overlapping data, so care should be taken to check the notes before using this data.*

Calculations for each country are as listed below, with CONPRCI2 = I' and CONPRCI1 = I . Years in brackets indicate the original base year before the data was homogenized.

Argentina (base year = 1980)

[1913] For 1900 – 1938, $I' = I/(120/100)/(3167/100)/(773/100)/(137/0.49)$

[1938] For 1939 – 1959, $I' = I/(3167/100)/(773/100)/(137/0.49)$

[1959] For 1960 – 1969, $I' = I/(773/100)/(137/0.49)$

[1969] For 1970 – 1975, $I' = I/(137/0.49)$

[1980] For 1976 – 1984, $I' = I$

[1984] For 1985 – 1990, $I' = I/(100/17462)$

[1990] For 1991 – 1993, $I' = I/(100/10716641)/(100/17462)$

[1995] For 1994 – 2000, data could not be homogenized

Canada (base year = 1980)

[1913] For 1900 – 1938, $I' = I/(129/100)/(375/66)$

[1938] For 1939 – 1975, $I' = I/(375/66)$

[1980] For 1976 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Chile (base year = 1980)

[1913] For 1913 – 1927, data could not be homogenized

[1938] For 1928 – 1959, $I' = I/(11238/100)/(870/100)/(37037/6.6)$

[1959] For 1960 – 1969, $I' = I/(870/100)/(37037/6.6)$

[1969] For 1970 – 1975, $I' = I/(37037/6.6)$

[1980] For 1976 – 1993, $I' = I$

[1995] For 1994 – 2000, data could not be homogenized

Uruguay (base year = 1980)

[1938] For 1929 – 1964, $I' = I/(2020/100)/(13849/11)$

[1964] For 1965 – 1975, $I' = I/(13849/11)$
[1980] For 1976 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

US (base year = 1980)

[1913] For 1870 – 1938, $I' = I/(142/100)/(383/65)$
[1938] For 1939 – 1975, $I' = I/(383/65)$
[1980] For 1976 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

US – Puerto Rico (base year = 1980)

[1943] For 1941 – 1975, $I' = I/(311/76)$
[1980] For 1976 – 1993, $I' = I$

Australia (base year = 1985)

[1938] For 1870 – 1938, $I' = I/(100/25)/(100/24.9)$
[1970] For 1939 – 1969, $I' = I/(100/24.9)$
[1985] For 1970 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Japan (base year = 1985)

[1938] For 1922 – 1946, $I' = I/(1453/8.2)/(100/36.9)$
[1970] For 1947 – 1969, $I' = I/(100/36.9)$
[1985] For 1970 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

New Zealand (base year = 1985)

[1938] For 1907 – 1938, $I' = I/(100/29.8)/(100/17.5)$
[1970] For 1939 – 1969, $I' = I/(100/17.5)$
[1985] For 1970 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Austria (base year = 1970)

[1929] For 1914 – 1948, $I' = I/(422/43)/(142/85)$
[1953] For 1949 – 1965, $I' = I/(142/85)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Belgium (base year = 1970)

[1913] For 1870 – 1913, data could not be homogenized
[1929] For 1914 – 1948, $I' = I/(326/95)/(125/84)$
[1953] For 1949 – 1965, $I' = I/(125/84)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Denmark (base year = 1970)

[1913] For 1870 – 1914, $I' = I/(103/60)/(174/80)/(147/73)$
[1929] For 1915 – 1948, $I' = I/(174/80)/(147/73)$
[1953] For 1949 – 1965, $I' = I/(147/73)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Finland (base year = 1970)

[1929] For 1914 – 1948, $I' = I/(696/70)/(178/66)$
[1953] For 1949 – 1965, $I' = I/(178/66)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

France (base year = 1970)

[1914] For 1870 – 1914, $I' = I/(100/17)/(1664/63)/(166/81)$
[1929] For 1915 – 1948, $I' = I/(1664/63)/(166/81)$
[1953] For 1949 – 1965, $I' = I/(166/81)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Germany (base year = 1970)

[1913] For 1870 – 1914, data could not be homogenized
[1929] For 1915 – 1944, data could not be homogenized
[1953] For 1945 – 1965, $I' = I/(128/88)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

East Germany (base year = 1960)

[1960] For 1950 – 1989, $I' = I$

Ireland (base year = 1970)

[1929] For 1914 – 1948, $I' = I/(181/79)/(143/77)$
[1953] For 1949 – 1965, $I' = I/(143/77)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Italy (base year = 1970)

[1913] For 1870 – 1914, $I' = I/(100/22)/(4734/86)/(156/88)$
[1939] For 1915 – 1948, $I' = I/(4734/86)/(156/88)$
[1949] For 1949 – 1965, $I' = I/(156/88)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Netherlands (base year = 1970)

[1913] For 1880 – 1914, $I' = I/(100/71)/(166/76)/(146/78)$
[1929] For 1915 – 1948, $I' = I/(166/76)/(146/78)$
[1953] For 1949 – 1965, $I' = I/(146/78)$

[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Norway (base year = 1970)

[1913] For 1901 – 1913, data could not be homogenized
[1929] For 1914 – 1948, $I' = I/(164/74)/(147/79)$
[1953] For 1949 – 1965, $I' = I/(147/79)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Spain (base year = 1970)

[1929] For 1914 – 1935, data could not be homogenized
[1953] For 1939 – 1965, $I' = I/(217/78)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Sweden (base year = 1970)

[1914] For 1870 – 1914, $I' = I/(100/59)/(154/177)/(149/81)$
[1929] For 1915 – 1948, $I' = I/(154/177)/(149/81)$
[1953] For 1949 – 1965, $I' = I/(149/81)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

Switzerland (base year = 1970)

[1914] For 1890 – 1914, $I' = I/(100/63)/(139/96)/(127/84)$
[1929] For 1915 – 1948, $I' = I/(139/96)/(127/84)$
[1953] For 1949 – 1965, $I' = I/(127/84)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

UK (base year = 1970)

[1851] For 1870 – 1871, $I' = I/(120/113)/(100/61)/(134/77)/(141/80)$
[1914] For 1872 – 1914, $I' = I/(100/61)/(134/77)/(141/80)$
[1929] For 1915 – 1948, $I' = I/(134/77)/(141/80)$
[1953] For 1949 – 1965, $I' = I/(141/80)$
[1970] For 1966 – 1993, $I' = I$
[1995] For 1994 – 2000, data could not be homogenized

PTS – Pupils and Teachers in School. Source: International Historical Statistics, 2003. Values are all in thousands and are separated by Pupils (**PU**) and Teachers (**TE**), then further separated into a (primary schools) (**PRM**), b (secondary schools) (**SEC**), and All Schools (**ALL**). The definition of the type of school, the time period to which the figures apply, and which pupils are covered varies. Changes in the scope and nature of the statistics are noted where possible. In general, vocational, technical, or teacher-training institutions are not included. **PTSTEPR2** and **PTSTESE2** have been standardized into thousands by us.

STUUNVST – Students in Universities. Source: International Historical Statistics, 2003. Generally, the statistics are for the school year beginning in the given calendar year. Institutions of higher education other than universities are sometimes included; where not, this is noted if possible. Undergraduate and graduate students are not distinguished. **STUUNVST2** standardized into thousands by us.

NAT – National Accounts Totals. Source: International Historical Statistics, 2003. Values are of GDP (**NATGDP** and **NATGDP2**, the latter standardized into current prices by us), GNP (**NATGNP** and **NATGNP2**, the latter standardized into current prices by us), NNP (**NATNNP**), GPFCE (**NATGPFCE** and **NATGPFCE2**, the latter standardized into thousand millions by us). GDP = gross domestic product; GNP = gross national product; NNP = net national product; and GPFCE = gross (private fixed) capital formation. All figures are at market prices unless otherwise noted. Capital formation statistics are for domestic capital only. Where revised figures at constant prices were not available to match revised figures at current prices, IHS used implicit deflators in the unrevised figures to derive new estimates for constant prices.

RL series includes various indicators of cultural heterogeneity by religion and language. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are separated by religion and language and given as percentages. Religion is split into Protestants (**RLPROTE**), Catholics (**RLCATHO**), Jews (which includes Israelites) (**RLJEWS**), and Other (which includes Greek Catholic, Greek Orthodox, Unknown, and Without Conf.) (**RLOTHERS**). Language is split into L1 (% of most spoken language) (**RLL1**), L2 (% of second-most spoken language) (**RLL2**), L3 (% of third-most spoken language) (**RLL3**), BiL (% of population that is bilingual and/or trilingual) (**RLBIL**), and NoL (number of languages spoken by at least 5% of the population) (**RLNOL**).

GGP series includes various measures of general government personnel. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are separated by General Administration Absolute Number (**GGPADMIN** and **GGPADMI2**, in which units were standardized by us), % in Judiciary (**GGPPRCJU**), % in Police (**GGPPRCPO**), General Administration in % of Total Population (**GGPGAPOP**), General Administration in % of Labour Force (**GGPGALAB**), General Administration in % of Total Personnel (**GGPGAPER**), Other Personnel in % of Total Personnel in Education (**GGPOPEDU**), Other Personnel in % of Total Personnel in Post (**GGPOPPST**), and Other Personnel in % of Total Personnel in Rail (**GGPOPRAI**). Number of personnel may be expressed in thousands or in absolute number (**GGPPERSN** and **GGPPERS2**, in which units were standardized by us). Total Personnel is also split into % of Total Population (**GGPPPOP**), and % of Total Labour (**GGPPLAB**).

TR series includes various measures of tax ratios. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are split into general government (GG) and central government (CG), then further split into GDP (GDP) and NDP (NDP). **TRGGGDP** is defined as general government taxes as a percentage of GDP, whereas **TRGGNDP** is defined as general government taxes as a percentage of NDP. Likewise, **TRCGGDP** is defined as central

government taxes as a percentage of GDP, while **TRCGNDP** is defined as central government taxes as a percentage of NDP.

CENTRLZN – Centralisation. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are central government taxes as a percentage of general government taxes.

GGT series includes various measures of General Government Taxes. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are listed in thousands and separated into total (**GGTTOTAL**) and tax categories in percent of total taxes. The latter is split into direct taxes (**GGTDIRECT**), income tax (**GGTINCME**), objective tax (**GGTOBJT**), inheritance (**GGTINHER**), indirect taxes (**GGTINDRT**), customs (**GGTCUSTM**), excises (**GGTEXCSE**), turnover taxes (**GGTURNNO**), and other ind. (**GGTOTHER**). Units for Total vary across countries. **GGTTOTAL%GFDGDP** is **GGTTOTAL** as a percentage of **GFDGDP**.

CGT series includes various measures of Central Government Taxes. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are listed in thousands and separated into total (**CGTTOTAL**) and tax categories in percent of total taxes. The latter is split into direct taxes (**CGTDIRECTX**) and indirect taxes (**CGTINDIR**). Direct taxes are split into income tax (**CGTINCTX**), property tax (**CGTPROT**), inheritance (**CGTINHER**), extra-ordinary tax (**CGTEXTRA**), assessed tax (**CGTASES**), land tax (**CGTLAND**), trade tax (**CGTTRADE**), corporation (**CGTCORP**), and other objective tax (**CGTOTOBJ**). Indirect taxes are split into customs (**CGTCUSTM**), excises (**CGTEXCS**), turnover tax (**CGTTURN**), and other (**CGTOTHER**). Units for Total vary across countries. **CGTTOTAL%GFDGDP** is **CGTTOTAL** as a percentage of **GFDGDP**.

PE series includes various measures of Public Expenditures. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. **PEGGCOST** and **PEGGCOS2**, in which units were standardized by us, are public expenditures by the general government in terms of absolute cost. **PEGG%GFDGDP** is public expenditures by the general government in terms of percent **GFDGDP**. **PEGG%FLORAGDP** is public expenditures by the general government in terms of percent **FLORAGDP**. **PEGGNDP** is public expenditures by the general government in terms of percent **NDP**. **PECGCOST** and **PECGCOS2**, in which units were standardized by us, is public expenditures by the central government in terms of absolute cost. **PECGGFDGDP** is public expenditures by the central government in terms of percent **GFDGDP**. **PECG%FLORAGDP** is public expenditures by the central government in terms of percent **FLORAGDP**. **PECGNDP** is public expenditures by the central government in terms of percent **NDP**. **PECCENTR** is public expenditures in terms of the percent of general expenditures spent by the central government, **PECREGNL** is public expenditures in terms of the percent of general expenditures spent by the regional governments, and **PECLOCAL** is public expenditures in terms of the percent of general expenditures spent by the local governments.

CGEM series includes various measures of central government expenditures by major categories. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-*

1975. **CGEMGDPT** is central government expenditures in major categories (defense, admin/justice, econ/environment, social services, and residual) terms of total % of GDP. **CGEMGDPD** is central government expenditures in the major category of defense in terms of total % of GDP. **CGEMGDPA** is central government expenditures in the major category of admin/justice in terms of total % of GDP. **CGEMGDPE** is central government expenditures in the major category of econ/environment in terms of total % of GDP. **CGEMGDPS** is central government expenditures in the major category of social services in terms of total % of GDP. **CGEMGDPR** is central government expenditures in the major residual category in terms of total % of GDP. **CGEMEXPD** is central government expenditures in the major category of defense in terms of % of total expenditures. **CGEMEXPA** is central government expenditures in the major category of admin/justice in terms of % of total expenditures. **CGEMEXPE** is central government expenditures in the major category of econ./environment in terms of % of total expenditures. **CGEMEXPS** is central government expenditures in the major category of social services in terms of % of total expenditures. **CGEMEXPR** is central government expenditures in the major residual category in terms of % of total expenditures.

CGE series includes various measures of central government expenditures which is broken down into finer categories. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. **CGETOTAL** is total central government expenditures. **CGETOTA2** is total central government expenditures with units standardized into millions of national currency by us. **CGETOTA%GFDGDP** is CGETOTA2 as a percentage of GFDGDP. **CGEEXPDF** is total central government expenditures in the category of defense **in terms of** % of total expenditures. **CGEEXPGA** is total central government expenditures in the category of general administration in terms of % of total expenditures. **CGEEXPJP** is total central government expenditures in the category of justice and police in terms of % of total expenditures. **CGEEXPES** is total central government expenditures in the category of economic services in terms of % of total expenditures. **CGEEXPTE** is total central government expenditures in the category of transportation and communication in terms of % of total expenditures. **CGEEXPSS** is total central government expenditures in the category of social security in terms of % of total expenditures. **CGEEXPPH** is total central government expenditures in the category of public health in terms of % of total expenditures. **CGEEXPHE** is total central government expenditures in the category of housing in terms of % of total expenditures. **CGEEXPES** is total central government expenditures in the category of education and science in terms of % of total expenditures. **CGEEXPDI** is total central government expenditures in the category of debt interest in terms of % of total expenditures.

GGEM series includes various measures of general government expenditures by major categories. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. **GGEMGDPT** is general government expenditures in major categories (defense, admin/justice, econ/environment, social services, and residual) terms of total % of GDP. **GGEMGDPD** is general government expenditures in the major category of defense in terms of total % of GDP. **GGEMGDPA** is general government expenditures in the major category of admin/justice in terms of total % of GDP. **GGEMGDPE** is general government expenditures in the major category of econ/environment in terms of total % of GDP. **GGEMGDPS** is general government expenditures in the major category of social services in terms of total % of GDP. **GGEMGDPR** is general government expenditures in the major residual category in terms of

total % of GDP. **GGEMEXPD** is general government expenditures in the major category of defense in terms of % of total expenditures. **GGEMEXPA** is general government expenditures in the major category of admin/justice in terms of % of total expenditures. **GGEMEXPE** is general government expenditures in the major category of econ./environment in terms of % of total expenditures. **GGEMEXPS** is general government expenditures in the major category of social services in terms of % of total expenditures. **GGEMEXPR** is general government expenditures in the major residual category in terms of % of total expenditures.

GGE series includes various measures of general government expenditures which is broken down into finer categories. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. **GGETOTAL** is total general government expenditures. **GGETOTA2** is total general government expenditures with units standardized into millions of national currency by us. **GGETOTA2%GFDGDP** is GGETOTA2 as a percentage of GFDGDP. **GGEEXPDF** is total general government expenditures in the category of defense **in terms of** % of total expenditures. **GGEEXPGA** is total general government expenditures in the category of general administration in terms of % of total expenditures. **GGEEXPJP** is total general government expenditures in the category of justice and police in terms of % of total expenditures. **GGEEXPES** is total general government expenditures in the category of economic services in terms of % of total expenditures. **GGEEXPTE** is total general government expenditures in the category of transportation and communication in terms of % of total expenditures. **GGEEXPSS** is total general government expenditures in the category of social security in terms of % of total expenditures. **GGEEXPPH** is total general government expenditures in the category of public health in terms of % of total expenditures. **GGEEXPHE** is total general government expenditures in the category of housing in terms of % of total expenditures. **GGEEXPES** is total general government expenditures in the category of education and science in terms of % of total expenditures. **GGEEXPDI** is total general government expenditures in the category of debt interest in terms of % of total expenditures.

ER series includes various types of social expenditures as a ratio of GDP. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. **ERSSEGDP** is social security expenditures as percentage of GDP. **ERSPEGDP** is social insurance and public health expenditures as percentage of GDP. **ERPBGDP** is pension benefits as percentage of GDP. **ERHBGDP** is health benefits as percentage of GDP. **ERUBGDP** is unemployment benefits as percentage of GDP. **ERFAGDP** is family allowance expenditures as percentage of GDP ().

FSSS series refers to financing shares in social security. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are separated into contributions by public authorities as percentage of total receipts (**FSSSPATR**) and contributions by employers as percentage of total receipts (**FSSSCETR**).

SIC series refers to social insurance coverage in terms of members active in various programs in terms of percentage of the labor force. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are separated into pension insurance (**SICPENI**), health insurance (**SICHLTHI**), occupational injuries insurance (**SICOINJI**), and unemployment insurance (**SICUNEMI**).

NCUR series refers to National Product at Current Prices. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are separated into GDP (m) (**NCURGDPM**), NDP (m) (**NCURNDPM**), NDP (f) (**NCURNDPF**), GDP(m) per capita (**NCURGMPC**), and NDP(f) per capita (**NCURNFPC**). GDP(m) per capita and NDP(f) per capita are in the same unit of currency, whereas the other categories are in the same units with regards to each other but in different units of the same currency.

NCON series refers to National Product at Constant Prices. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. Values are separated into GDP (m) (**NCONGDPM**), NDP (m) (**NCONNDPM**), NDP (f) (**NCONNDPF**), GDP(m) per capita (**NCONGMPC**), and NDP(f) per capita (**NCONNFPC**). GDP(m) per capita and NDP(f) per capita are in the same unit of currency, whereas the other categories are in the same units with regards to each other but in different units of the same currency.

OP series refers to origin and use of national product. Source: Flora, Peter (1987) *State, economy, and society in Western Europe, 1815-1975*. **OPGDPMAG** is percent of GDP originating from agriculture at market prices, **OPGDPMIN** is percent of GDP originating from industry at market prices, and **OPGDMSR** is percent of GDP originating from services at market prices. **OPGNMCPR** is percent of GNP used for private consumption at market prices. **OPGNMCPU** is percent of GNP used for public consumption at market prices. **OPGNMCTL** is percent of GNP used for total (public plus private) consumption at market prices. **OPGNMINF** is percent of GNP used for investment (net capital formation) at market prices. **OPGNMIGF** is percent of GNP used for investment (gross capital formation) at market prices. **OPGNMICS** is percent of GNP used for investment (changes in stocks) at market prices. **OPGNMITL** is percent of GNP used for investment (total investment) at market prices. **OPGNMEXP** is percent of GNP used for exports at market prices. **OPGNMIMP** is percent of GNP used for imports at market prices.

PPV series refers to political parties. Sources: Up to and including 1959: Flora, Peter et. Al. [Eds]. 1987. *State, economy, and society in Western Europe, 1815-1975: a data handbook in two volumes*. Chicago: St. James Press, Vol. 1 p. 99-151. 1960 and afterward: Huber, Evelyne, Charles Ragin and John Stephens. Comparative Welfare States dataset. Available online at <http://www.nsd.uib.no/macrodatabase/set.html?id=8&sub=1>
For Norway seats data 1882-1900: Online data from Norwegian Social Science Data Service, available at http://www.nsd.uib.no/polsys/index.cfm?urlname=storting&lan=&UttakNr=102&MenuItem=N1_1&ChildItem=&State=collapse

Pre-1960 data from Flora (1987) was transformed using the following method: Parties were classified according to the following categories: Left, Right, Center, Right Christian, Right Catholic, Center Catholic, Nationalist/Regionalist, Liberal, and Unknown. The percentage of votes received/seats controlled by parties in each category was then summed.

PPVVTURN is total votes in percentage of electorate.

PPVLEFTVOT is percentage of votes received by Left parties in the national legislature.

PPVRTVOT is percentage of votes received by Right parties in the national legislature.

PPVCNVOT is percentage of votes received by Center parties in the national legislature. **PPVRTCRV** is percentage of votes received by Right Christian parties in the national legislature. **PPVCNCRV** is Percentage of Votes Received by Center Christian Parties in the national legislature. **PPVRTCTV** is percentage of votes received by Right Catholic parties in the national legislature. **PPVCNCTV** is percentage of votes received by Center Catholic parties in the national legislature. **PPVNATVO** is percentage of votes received by Nationalist/Regionalist parties in the national legislature. **PPVLIBVO** is percentage of votes received by Liberal parties in the national legislature. **PPVUNKVO** is percentage of votes received by parties of unknown ideological orientation in the national legislature. **PPVOTHVO** is percentage of votes received by other parties in the national legislature. **PPVCNLFT** is percentage of votes received by Center/Left coalitions in the national legislature. **PPVCNRTV** is percentage of votes received by Center/Right coalitions in the national legislature. **PPVLFTCR** is percentage of votes received by Left Christian parties in the national legislature.

PPVLEFTS is percentage of seats in the national legislature controlled by Left parties. **PPVRTSEA** is percentage of seats in the national legislature controlled by Right parties. **PPVCNSEA** is percentage of seats in the national legislature controlled by Center parties. **PPVRTCRS** is percentage of seats in the national legislature controlled by Right Christian parties. **PPVCNCRS** is percentage of seats in the national legislature controlled by Center Christian parties. **PPVRTCTS** is percentage of seats in the national legislature controlled by Right Catholic parties. **PPVCNCTS** is percentage of seats in the national legislature controlled by Center Catholic parties. **PPVNATSE** is percentage of seats in the national legislature controlled by Nationalist/Regionalist parties. **PPVLIBSE** is percentage of seats in the national legislature controlled by Liberal parties. **PPVUNKSE** is percentage of seats in the national legislature controlled by parties of unknown ideological orientation. **PPVOTHSE** is percentage of seats in the national legislature controlled by other parties. **PPVLFCS** is percentage of seats in the national legislature controlled by Left Christian parties.

The following variables come from Huber, Evelyne, Charles Ragin and John Stephens, Comparative Welfare States Dataset, 2004. URL to data and codebook can be found at the following website: <http://www.unc.edu/~jdsteph/common/data-common.html>

LEFTCAB is left seats as a percentage of seats held by all government parties. **LEFTMAJ** is Left seats as a percentage of parliamentary seats needed to have a majority ($1/2 + 1$). **LTCABCUM** – Cumulative **LEFTCAB** score from 1946 to the year of the observation. For example, **LTCABCUM** for Australia 1948 = **LEFTCAB** 1946 + **LEFTCAB** 1947 + **LEFTCAB** 1948. **CNCAB** is center seats as a percentage of seats held by all government parties. **CNMAJ** is Center seats as a percentage of parliamentary seats needed to have a majority ($1/2 + 1$). **CNCABCUM** is cumulative **CNCAB** score from 1946 to the year of the observation. For example, **CNCABCUM** for Australia 1948 = **CNCAB** 1946 + **CNCAB** 1947 + **CNCAB** 1948. **CNCRAB** is Center, Christian seats as a percentage of seats held by all government parties. **CNCRMAJ** is Center, Christian seats as a percentage of parliamentary seats needed to have a majority ($1/2 + 1$). **CNCRUM** is cumulative **CNCRAB** score from 1946 to the year of the observation. For example, **CNCRUM** for Australia 1948 = **CNCRAB** 1946 + **CNCRAB** 1947 + **CNCRAB** 1948. **CNCTCAB** is Center, Catholic seats as a percentage of seats held by all government parties. **CNCTMAJ** is Center, Catholic seats as a percentage of parliamentary

seats needed to have a majority (1/2 plus 1). **CNCTCUM** is Cumulative CNCTCAB score from 1946 to the year of the observation. For example, CNCTCUM for Australia 1948 = CNCTCAB 1946 + CNCTCAB 1947 + CNCTCAB 1948. **RTCAB** is Right seats as a percentage of seats held by all government parties. **RTMAJ** is Right seats as a percentage of parliamentary seats needed to have a majority (1/2 plus 1). **RTCABCUM** is cumulative RTCAB score from 1946 to the year of the observation. For example, RTCABCUM for Australia 1948 = RTCAB 1946 + RTCAB 1947 + RTCAB 1948. **RTCRCAB** is Right, Christian seats as a percentage of seats held by all government parties. **RTCRMAJ** is Right, Christian seats as a percentage of parliamentary seats needed to have a majority (1/2 plus 1). **RTCRCUM** is Cumulative RTCRCAB score from 1946 to the year of the observation. **RTCTCAB** is Right, Catholic seats as a percentage of seats held by all government parties. **RTCTMAJ** is Right, Catholic seats as a percentage of parliamentary seats needed to have a majority (1/2 plus 1). **RTCTCUM** is cumulative RTCTCAB score from 1946 to the year of the observation. For example, RTCTCUM for Australia 1948 = RTCTCAB 1946 + RTCTCAB 1947 + RTCTCAB 1948.