WORLD TOTALS IN WID: CORE TERRITORIES, CORE COUNTRIES, AND CORE MACRO AND DISTRIBUTIONAL VARIABLES, 1820-2023

ROWAIDA MOSHRIF
GASTON NIEVAS
THOMAS PIKETTY
ALICE SODANO
LUCAS CHANCEL

TECHNICAL NOTE N°2024/02

SEPTEMBER 2024

WORLD INEQUALITY LAB

World Totals in WID: Core Territories, Core Countries and Core Macro and Distributional Variables, 1820-2023

Rowaida Moshrif **Gaston Nievas** **Thomas Piketty** **Alice Sodano** **Lucas Chancel**

First version: June 3, 2024

September 26, 2024

Abstract

This technical note presents the world's territorial definitions that are currently used in WID.world. These definitions depend on the series and the time period: for historical series (1820-2023) we refer to the world as the sum of 33 core territories (24 countries + 9 subregions); for the annual series covering recent decades (1970-2023) we refer to the world as the sum of 216 core countries. By construction, all core variables (macro and distributional) are systematically available for these core territories/countries, and all core variables at the world level are equal to the sum of these variables for all core territories/countries according to both definitions. More precisely, we ensure that core variables are always available for all core territories in 1820, 1850, 1880 and every ten years from 1900 to 1970 and annually for all core countries for years 1970-2023. This implies that in some cases we will rely on simple assumptions to fill in some gaps and correct for the changes in borders and jurisdictions. This is explained in this technical note, which will be regularly updated.

1 Introduction

In this technical note, we delve into the definitions of territories as presented in WID.world. Territorial definitions in WID.world are contingent upon the temporal dimension. For historical analyses spanning from 1820 to 2023, the world is conceptualized as the aggregation of 33 core territories, comprising 24 countries alongside 9 subregions. In contrast, the more contemporary annual series, spanning from 1970 to 2023, delineate the world as the summation of 216 core countries.

We define a group of core macro and distributional variables based on their importance for comparative analysis. We guarantee the presence of core variables for all regions and territories for years 1820, 1850 and 1880, and subsequently at 10 years intervals, spanning from 1900 to 1970. Subsequently, for the period from 1970 onwards, core macro variables are accessible for all core countries. However, achieving this temporal consistency necessitates the utilization of assumptions to address data gaps, a process that we denote in detail later in this note.

2 How to aggregate to world totals?

2.1 Aggregate to world totals (core countries)

We provide an example of how to ensure aggregating world totals using the command 'wid' from STATA[1]:

```
global corecountries `" "AD" "AE" "AF" "AG" "AI" "AL" "AM" "AO" "AR" "AT" "AU" "AW" "AZ"
     "BA" "BB" "BD" "BE" "BF" "BG" "BH" "BI" "BJ" "BM" "BN" "BO" "BQ" "BR" "BS" "BT" "BW
    " "BY" "BZ" "CA" "CD" "CF" "CG" "CH" "CI" "CL" "CM" "CN" "CO" "CR" "CU" "CV" "CW" "
    CY" "CZ" "DE" "DJ" "DK" "DM" "DO" "DZ" "EC" "EE" "EG" "ER" "ES" "ET" "FI" "FJ" "FM"
    "FR" "GA" "GB" "GD" "GE" "GG" "GH" "GI" "GL" "GM" "GN" "GQ" "GR" "GT" "GW" "GY" "HK"
     "HN" "HR" "HT" "HU" "ID" "IE" "IL" "IM" "IN" "IQ" "IR" "IS" "IT" "JE" "JM" "JO" "JP
    " "KE" "KG" "KH" "KI" "KM" "KN" "KP" "KR" "KS" "KW" "KY" "KZ" "LA" "LB" "LC" "LI" "
    LK" "LR" "LS" "LT" "LU" "LV" "LY" "MA" "MC" "MD" "ME" "MG" "MH" "MK" "ML" "MM" "MN"
    "MO" "MR" "MS" "MT" "MU" "MV" "MW" "MX" "MY" "MZ" "NA" "NC" "NE" "NG" "NI" "NL" "NO"
     "NP" "NR" "NZ""'
global corecountries `" $corecountries "OM" "PA" "PE" "PF" "PG" "PH" "PK" "PL" "PR" "PS"
     "PT" "PW" "PY" "QA" "RO" "RS" "RU" "RW" "SA" "SB" "SC" "SD" "SE" "SG" "SI" "SK" "SL
    " "SM" "SN" "SO" "SR" "SS" "ST" "SV" "SX" "SY" "SZ" "TC" "TD" "TG" "TH" "TJ" "TL" "
    TM" "TN" "TO" "TR" "TT" "TV" "TW" "TZ" "UA" "UG" "US" "UY" "UZ" "VC" "VE" "VG" "VN"
    "VU" "WS" "YE" "ZA" "ZM" "ZW""'

wid, indicators(npopul) age(992) popul(i) years(1970/2023) clear

gen corecountry = .
foreach c of global corecountries {
    replace corecountry = 1 if country == "`c'"
}
```

[1] For the list of country codes, see https://wid.world/codes-dictionary/#country-code

```
keep if (corecountry == 1 | country=="WO") & year >= 1970

gen total=value if country=="WO"

replace value=. if country=="WO"

collapse (sum) sum=value total, by(year)

gen diff=(total-sum)/total
```

As this computer code illustrates, the world population is exactly equal to the sum of core countries population throughout the 1970-2023 period[2]. The same applies to national income and other core macro variables. For distributional variables, the world distribution (as well as the various regional distributions) is exactly equal to the aggregation of the core countries distribution, using the distributional aggregation code available in the WID.world computer codes GitHub platform (https://github.com/WIDworld/wid-world).

2.2 Aggregate world totals (core territories)

Here we provide an example of how to ensure aggregating world totals at the level of core territories using the command 'wid' from STATA:

```
global coreterritories `" "RU" "OA" "CN" "JP" "OB" "DE" "ES" "FR" "GB" "IT" "SE" "OC" "
    QM" "AR" "BR" "CL" "CO" "MX" "OD" "DZ" "EG" "TR" "OE" "CA" "US" "AU" "NZ" "OH" "IN"
    "ID" "OI" "ZA" "OJ" "'


wid, indicators(npopul) age(992) popul(i) years(1970/2023) clear

gen coreterritory = .
foreach c of global coreterritories {
    replace coreterritory = 1 if country == "`c'"
}
keep if keep if (coreterritory == 1 | country=="WO") & (year==1820 | year==1850 | year
    ==1880 | year==1900 | year==1910 | year==1920 | year==1930 | year==1940 | year==1950
     | year==1960 | year >= 1970)

gen total=value if country=="WO"

replace value=. if country=="WO"

collapse (sum) sum=value total, by(year)

gen diff=(total-sum)/total
```

[2]Due to rounding mistakes and data storage issues, totals and sums might not always be exactly equal, but we ensure percentage differences are always less than 10^{-4}, i.e. 0.01%.

3 Core territories

3.1 Long-run historical series: 1820-2023

For the purpose of long-run series, WID divides the world into 8 regions and 33 core territories (24 countries + 9 subregions). The core aggregate variables include total population, adult population, total national income and average national income. Based on the paper by Chancel and Piketty (2021) followed by the technote by Chancel, Moshrif, Piketty, and Xuereb (2023), the historical macro variables were provided for years 1820, 1850, 1880, at 10 years interval from 1900 to 1970 and for all years from 1970 to 2023 (annual series). As these series were added to existent historical series provided by WID fellows, the macro series are now available at the yearly level from 1900 onwards.[3]

Table 1

The World as the Sum of 33 Core Territories: core macro variables (1820-2023)

Regions	Core Territories	Core Variables	Years
East Asia (3)	China, Japan, Other East Asia	**Total population** (npopul999i), **Adult population** (npopul992i), **Total national income** (mnninc999i), **Average national income** (anninc)	**1820**, **1850**, **1880**, **1900-2023 (annual)**
Europe (8)	Britain, France, Germany, Italy, Spain, Sweden, Other Western Europe, Eastern Europe		
Latin America (6)	Argentina, Brazil, Chile, Colombia, Mexico, Other Latin America		
Middle East/North Africa (4)	Algeria, Egypt, Turkey, Other Middle East/North Africa		
North America/Oceania (5)	USA, Canada, Australia, New Zealand Other North America/Oceania		
Russia/Central Asia (2)	Russia Other Russia/Central Asia		
South/South-East Asia (3)	India, Indonesia, Other South/Southeast Asia		
Sub Saharan Africa (2)	South Africa, Other Sub-Saharan Africa		

Interpretation. For the purpose of long-run series, WID divides the world into 8 regions and 33 core territories (24 countries + 9 subregions). The core aggregate variables include total population, adult population, total national income and average national income. All core variables are available for all core territories for years 1820, 1850, 1880, and for all years from 1900 to 2023 (annual series).

3.2 Recent period : 1970-2023

For the purpose of the more recent series on macro variables, WID divides the world into 216 core countries. The list of countries encompassing the world can be seen in Table 2 below, together with their geographical region.

[3]The missing data points for the core macro variables were linearly interpolated between the historical long-run series by Chancel and Piketty (2021) and original data points provided by WID fellows.

Table 2

The World as the Sum of 216 Core Countries (1970-2023)

Regions	Core Countries
East Asia (8)	China, Hong Kong, Japan, Macao, Mongolia, North Korea, South Korea, Taiwan
Europe (47)	Albania, Andorra, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Gibraltar, Greece, Guernsey, Hungary, Iceland, Ireland, Isle of Man, Italy, Jersey, Kosovo, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Moldova, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom
Latin America (43)	Anguilla, Antigua and Barbuda, Argentina, Aruba, Bahamas, Barbados, Belize, Bonaire, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Curaçao, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Montserrat, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Sint Maarten, Suriname, Trinidad and Tobago, Turks and Caicos Islands, Uruguay, Venezuela
Middle East/North Africa (20)	Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestinian Territories, Qatar, Saudi Arabia, Syria, Tunisia, Turkey, United Arab Emirates, Yemen
North America/Oceania (19)	Australia, Bermuda, Canada, Fiji, French Polynesia, Greenland, Kiribati, Marshall Islands, Micronesia, Nauru, New Caledonia, New Zealand, Palau, Samoa, Solomon Islands, Tonga, Tuvalu, United States, Vanuatu
Russia/Central Asia (11)	Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan.
South/South-East Asia (19)	Afghanistan, Bangladesh, Bhutan, Brunei, Cambodia, India, Indonesia, Laos, Malaysia, Maldives, Myanmar, Nepal, Pakistan, Papua New Guinea, Philippines, Singapore, Sri Lanka, Thailand, Timor-Leste, Vietnam
Sub Saharan Africa (54)	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Djibouti, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Republic of the Congo, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Tanzania, Togo, Uganda, Zambia, Zimbabwe

Interpretation. For the purpose of long-run series, WID divides the world into 8 regions and 216 core countries. The core aggregate variables include the variables listed on table 1, as well as foreign assets and liabilities and capital income inflows and outflows. They are available for all regions, core territories and core countries for all years 1970-2023 (annual series). The complete list of core macro variables available over the 1820-2023 is listed in table 1 above. By construction, all variables at the world level are equal to the exact sum of these variables for the 8 core regions, the 33 core territories and the 216 core countries. In particular, net foreign wealth and net foreign capital income are always equal to zero at the world level.

4 Core macro variables

While for the historical long run series we provide 4 core macro variables (total population, adult population, total national income and average national income), for the recent period core macro variables provided amount to 37 and are provided annually from 1970[4] to 2023 -and updated every year-. For a full list of the 37 variables, please refer to Table 3 below.

[4]Some macro variables such as population and GDP are available for many countries starting 1950.

We rely on official sources such as the World Bank[5], the International Monetary Fund[6] or the UN System of National Accounts[7] for most of national accounts data. See Blanchet and Chancel (2016) for additional methodological details (Blanchet and Chancel, 2016). We complement these with specific studies such as Piketty and Zucman (2014) or the Madison Project. Each series for each country constitutes a specific case and its source is well documented in the series metadata, we encourage the reader to refer to it in case of doubt.

In the cases where data cannot be found in any source, we first carry forward (or carry backward) the most recent available value as a share of the country's GDP and then, if no value is ever present, we impute it using the regional average of the value as a share of the countries' GDPs. We also rely on some other assumptions that are described below.

[5]https://data.worldbank.org/

[6]https://www.imf.org/en/publications/weo

[7]https://unstats.un.org/unsd/snaama/

Table 3

Core Macro Variables: Variables Description and Decomposition

WID Codes	Variables Description and Decomposition
nninc	*(=) National income*
gdpro	(+) Gross domestic product
confc	(-) Consumption of fixed capital
nnfin	(+) Net foreign income
finrx	(+) Foreign income received from the rest of the world
flcir	(+) Labor and capital income from the rest of the world
comrx	(+) Compensation of employees received from the world
pinrx	(+) Property income received from the rest of the world
fdirx	(+) Foreign direct investment income received from the world
ptfrx	(+) Portfolio and other income received from the rest of the world
ptdrx	(+) Debt income received from the rest of the world
pterx	(+) Equity income received from the rest of the world
ptrrx	(+) Reserves income received from the rest of the world
ptfrr	(+) Reinvested earnings on foreign portfolio investment, received from the rest
fsubx	(+) Subsidies on production received from the rest of the world
finpx	(-) Foreign income paid to the rest of the world
flcip	(+) Labor and capital income paid to the rest of the world
compx	(+) Compensation of employees paid to the rest of the world
pinpx	(+) Property income paid to the rest of the world
fdipx	(+) Foreign direct investment income paid to the rest of the world
ptfpx	(+) Portfolio and other income paid to the rest of the world
ptdpx	(+) Debt income paid to the rest of the world
ptepx	(+) Equity income paid to the rest of the world
ptfrp	(+) Reinvested earnings on foreign portfolio investment, paid to the rest of the
ftaxx	(+) Taxes on production paid to the rest of the world

WID Codes	Variables Description and Decomposition
nnfin	*(=) Net foreign income*
flcin	(+) Net foreign labor and capital income
pinnx	(+) Net foreign capital income
fdinx	(+) Net foreign direct investment income
ptfnx	(+) Net portfolio income
ptdnx	(+) Net debt income
ptenx	(+) Net equity income
ptfrn	(+) Net reinvested earnings on foreign portfolio investment
comnx	(+) Net foreign labor income
taxnx	(+) Foreign subsidies less taxes on production and imports

nwnxa	*(=) Net foreign assets*
nwgxa	(+) Gross foreign assets
ptfxa	(+) Portfolio assets
ptdxa	(+) Debt assets
ptexa	(+) Equity assets
ptrxa	(+) Reserves assets
fdixa	(+) Foreign direct investment assets
nwgxd	(-) Gross foreign liabilities
ptfxd	(+) Portfolio liabilities
ptdxd	(+) Debt liabilities
ptexd	(+) Equity liabilities
fdixd	(+) Foreign direct investment liabilities

Public Spending and revenue

WID Codes	**Variables Description and Decomposition**
expgo	*(=) Total public spending (excluding interest payment)*
gpsge	(+) General public services (excluding interest payments)
defge	(+) Defense
polge	(+) Public order and safety
ecoge	(+) Economic affairs
envge	(+) Environmental protection
houge	(+) Housing and community amenities
heage	(+) Health
recge	(+) Recreation, culture and religion
eduge	(+) Education
edpge	(+) Education: Primary
edsge	(+) Education: Secondary
edtge	(+) Education: Tertiary
sopge	(+) Social protection
spige	(+) Social protection: social insurance
sacge	(+) Social protection: social assistance in cash
sakge	(+) Social protection: social assistance in kind
revgo	*(=) Total public revenue*
pitgr	(+) Personal income tax
citgr	(+) Corporate income tax
scogr	(+) Social contributions
pwtgr	(+) Property and wealth taxes
intgr	(+) Indirect taxes
ottgr	(+) Other taxes
ntrgr	(+) Non-tax revenue
retgo	*(=) Total public revenue (excluding non-tax revenue)*
revgo	(+) Total public revenue
ntrgr	(-) Non-tax revenue
psugo	(=) Primary surplus of the government
revgo	(+) Total public revenue
expgo	(-) Total public spending (excluding interest payment)
ssugo	(=) Secondary surplus of the government
psugo	(+) Primary surplus of the government
inpgo	(-) Interest paid by the government

WID Codes	Variables Description and Decomposition
ncanx	*(=) Current Account = pinnx + comnx + tbnnx + taxnx* +scinx
pinnx	(+) Net foreign capital income
pinrx	(+) Property income received from the rest of the world
pinpx	(-) Property income paid from the rest of the world
comnx	(+) Net foreign labor income
comrx	(+) Compensation of employees received from the world
compx	(-) Compensation of employees paid from the rest of the world
tbnnx	(+) Trade balance (exports - imports)
tbxrx	(+) Exports of goods and services
tbmpx	(-) Imports of goods and services
taxnx	(+) Foreign subsidies less taxes on production and imports
fsubx	(+) Subsidies on prod. received from the rest of the world
ftaxx	(-) Taxes on prod. paid to the rest of the world
scinx	(+) Net remittances
scirx	(+) Remittances received from the rest of the world
scipx	(-) Remittances paid to the rest of the world
fkanx	*(=) Capital Account*
fkarx	(+) Capital transfers received from the rest of the world
fkapx	(-) Capital transfers paid to the rest of the world

Indices, Exchange Rates, and Population

WID Codes	Variables Description and Decomposition
inyixx999i	National income price index
xlcusx999i	Market exchange rate, LCU per USD
xlceux999i	Market exchange rate, LCU per EUR
xlcyux999i	Market exchange rate, LCU per CNY
xlcusp999i	PPP conversion factor, LCU per USD
xlceup999i	PPP conversion factor, LCU per EUR
xlcyup999i	PPP conversion factor, LCU per CNY
npopul992i	Population, Adults
npopul999i	Population, All individuals

4.1 GDP

In order to cover the whole period we rely on several assumptions:

- Ex-soviet countries, there is a year of GDP in 1973 (Madison) we interpolate from 1980 up to that year as done in Blanchet, Chancel, and Gethin (2022) . From 1973 to 1970 we use share of URSS GDP to extrapolate backwards.
- Curaçao and Sint Maarten were retropolated as a share of Former Netherlands Antilles for previous years.
- Former Yugoslavan countries are retropolated as shares of Yugoslavia GDP except for Kosovo. Kosovo is retropolated as a share of Serbian's GDP.
- Czech Republic and Slovakia are retroplated as a share of Czechoslovakia's GDP for the period prior to 1980.
- Eritrea is retropolated as a share of Ethiopia's GDP for the period prior to 1993.
- Timor Leste is retropolated as a share of Indonesia's GDP for the period prior to 1990.
- South Sudan is retropolated as a share of Sudan's GDP for the period prior to 2012.
- Zanzibar is retropolated as a share of Tanzania's GDP for the period prior to 1990.
- Isle of Man, Guernsey, Jersey and Gibraltar are retropolated as a share of United Kingdom's GDP.

When the country used in the denominator for calculating shares and retropolating GDP is a core country, then we substract that retropolated GDP from the parent country. Example: Serbia gets substracted Kosovo's GDP. This does not apply to small islands such as Isle of Man, Guernsey, Jersey and Gibraltar as a share of United Kingdom. Since the United Kingdom's GDP does not include those figures, we are not double counting the islands' in the world total. This also applies to Bonaire, Sint Eustatius and Saba, which is a territory that belongs to the Netherlands but its GDP is not reported as part of the Netherland's.

4.2 Price index

- Soviet Union countries are assumed to experience the same inflation rate from 1970 to 1990, years for which we had data for Russia. These countries are Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Tajikistan, Turkmenistan, Ukraine, Uzbekistan.
  - Estonia, Latvia and Lithuania currently use Euro, the inflation rate for 1970-1990 is an average of Russia and the EU.
- Former Yugoslavan countries are assumed to experience same inflation rate as Yugoslavia for years 1970-1990. These countries are Bosnia and Herzegovina, Croatia, Republic of North Macedonia and Serbia.
  - Kosovo, Montenegro and Slovenia currency use Euro, their inflation rate for 1970-1990 is an average of Yugoslavia and the EU.

- Eritrea is assumed to experience the same inflation rate as Ethiopia for period 1970-1990
- Czech Republic and Slovakia are assumed to experience the same inflation rate as Czechoslovakia for period 1970-1993.
- South Sudan is assumed to experience the same inflation rate as Sudan for period 1970-2008.
- Timor Leste is assumed to experience the same inflation rate as Indonesia for period 1970-1990.
- Curaçao, Sint Maarten and Bonaire, Sint Eustatius, and Saba are assumed to experience the same inflation rate as Aruba for period 1970-2005.
- For Isle of Man we use United Kingdom's price index.

4.3 Exchange rates

Market values

- We completed EUR to USD exchange rates before 1990 using UN SNA for available countries.
  - Not available for Estonia, Kosovo, Lithuania, Latvia, Slovenia and Slovakia. Pre 1990 exchange rate for these countries follows the evolution of the average exchange rate of the available EUR countries.
- Soviet countries' exchange rate follows the evolution of the Soviet Union pre 1990.
  - Georgia is an exception. We have estimates of GDP_{LCU} in real terms from interpolating GDP 1990 to GDP 1973 from Madisson. Before 1973 comes by applying shares to USSR. Price index to get nominal GDP_{LCU} was explained above. From UN SNA we have GDP_{USD}, we retropolated it as a share of USSR. We get the exchange rate by dividing $\frac{GDP_{LCU}}{GDP_{USD}}$ both in nominal terms.
- Similar process as for Georgia is followed for Former Yugoslavian countries.
  - We have GDPlcu in real terms from Blanchet et al. (2022)
- Yemen's exchange rate is assumed to follow the evolution of the Former Yemen Arab Republic previous years.
- For Curaçao and Sint Maarten we use exchange rates from Former Netherlands Antilles for previous years.
- For Czech Republic we use exchange rates from Csekoslovakia for previous years.
- For Zimbabwe, years 2017-2021 are replaced by dividing $\frac{GDP_{LCU}}{GDP_{USD}}$ from World Bank data. This is because WB does not report an exchange rate and UN SNA's does not vary, provoking a virtual increase in GDP_{USD}.
- For Gibraltar, we use exchange rates from Guernsey.

PPP

- For Zanzibar, we rely on the PPP exchange rate of Tanzania.
- For Channel Islands, Isle of Man, Gibraltar, Guernsey, Jersey, Anguilla and Montserrat, we rely on the PPP exchange rate of the United Kingdom.
- For Bonaire, Sint Eustatius and Saba, we rely on the PPP exchange rate of Curaçao.
- For the Faroe Island and Greenland, we rely on the PPP exchange rate of Denmark.
- For the Virgin Islands, we rely on the PPP exchange rate of United States.
- For the Liechtenstein, we rely on the PPP exchange rate of Switzerland.
- For the French Polynesia, New Caledonia and Monaco, we rely on the PPP exchange rate of France.

4.4 Population

We source population figures from the United Nations Department of Economic and Social Affairs (Population Division)[8] and complement it with the United Nations System of National Accounts[9]. Both databases combined ensure a vast coverage, allowing for the complete series for our 216 core countries for the period 1970-2023.

4.5 Balance of Payments and International Investment Position

4.5.1 Foriegn Capital Income

The primary source for data on foreign capital income is the IMF Balance of Payments (BOP)[10], and in situations where IMF data is not accessible, alternative sources like the United Nations or OECD statistics are used (for a detailed coverage see Nievas and Sodano (2024)). Foreign capital income encompasses diverse components, including portfolio and other income received and paid, and reinvested earnings on portfolio investment.

If foreign capital income is not reported for a certain year but an aggregate is reported (e.g.: foreign income), then we use the foreign capital income-to-foreign income ratio of the closest year to fill in the missing value. If foreign capital income received or paid is available but the country does not report its decomposition (FDI or portfolio), then we assume each asset class capital income is proportional to the share of the asset class on aggregate wealth.

For missing values, predictions are made based on asset class stock, GDP in USD, exchange rates, and inflation rates. Return rates predictions are made separately for each asset class since FDI is assumed to be more profitable than portfolio. An Ordinary Least Squares (OLS) regression model is

[8] https://population.un.org/wpp/Download/Standard/MostUsed/
[9] https://unstats.un.org/unsd/snaama/Index
[10] https://data.imf.org/?sk=7a51304b-6426-40c0-83dd-ca473ca1fd52

used, including country-specific fixed effects to account for time-invariant characteristics of each economy, as well as region-year fixed effects to capture unobserved shocks affecting the region uniformly. Specifically:

$$i^{B}_{\rho,ct} = \beta_0 + \beta_1 \frac{wealth^{B}_{\rho,ct}}{GDP_{ct}} + \beta_2 e_{ct} + \beta_3 \pi_{ct} + \alpha_c + \gamma_{rt} + \epsilon_{ct} \tag{A1}$$

Where i refers to the return rate, B to asset or liability, ρ to the asset class (FDI or portfolio), c to the country, t to the year, e to the nominal exchange rate with respect to US dollars, π to the inflation rate and α, γ and ϵ to the country fixed effects, region-year fixed effects and error term, respectively. Whenever data is still missing, we impute the value based on the regional average.

4.5.2 Foreign Wealth

The data on foreign wealth is sourced from "The External Wealth of Nations" (Lane and Milesi-Ferretti, 2018), which provides a standard breakdown of external assets and liabilities based on the Balance of Payments (BOP) Statistics Manual 6. External financial assets and liabilities encompass various components, such as foreign direct investment, portfolio equity, portfolio debt, other investment, and financial derivatives. Notably, foreign exchange reserves are included as financial assets, while gold holdings are excluded. In cases where data coverage is incomplete, countries are assumed to follow the regional trend. Only six countries have been completely imputed using a regional average.[11]

4.5.3 Corrections

The adjustments made ensured that the net foreign capital income and net foreign wealth collectively sum up to zero globally, contingent upon the presence of all 216 economies, by simply distributing proportionally any excess. For a full discussion of the advantages of this adjustment procedure with respect to other alternatives the reader should refer to Nievas and Piketty (2024).

Retained earnings on portfolio investment: The income that a company retains after having paid its suppliers, its employees, its shareholders, and its corporate income tax bill is what we call "undistributed profits" or "retained earnings." This flow is part of national income.

However, imagine that a company in country A has some undistributed profits, but is actually owned by residents of country B. If the ownership takes the form of portfolio investment, meaning that the residents of country B do not have a direct control over the company's decisions, then the SNA currently considers that the entire flow of undistributed profits belongs to the national income of country A, not country B.

We correct SNA following Blanchet et al. (2021), by redistributing the corresponding share to country B. The correction estimates both the flow of foreign retained earnings that accrue to residents and the flow of domestic retained earnings that accrue to foreigners. The difference between these two items leads to our adjustment. We completed the procedure for all 216 countries and made sure that aggregates add up to 0. Tax Havens do not play a role here.

[11] Bonaire, Cuba, Kosovo, Monaco, North Korea, Puerto Rico

Current and capital account: We ensure trade in goods global aggregates to be consistent by exploiting the well recorded bilateral statistics. We apply what is common practice in the trade literature, computing exports by mirroring imports. This is, in other words, assuming that the recorded imports to country A from country B are equal to the exports from B to A. This allows for global imports of goods to equal global exports of goods in each year. We then estimate trade in services for countries by subtracting trade in goods from total trade as reported in the BoP. As bilateral trade in goods and BoP trade data come from different sources, some times figures might be conflicting. In particular, there are cases where exports (imports) in goods is higher than total exports (imports) from BoP. For these cases we assume total exports (imports) from BoP is correct and use the ratio of goods exports (imports) to total exports (imports) from the previous year to correct for the figure of trade in goods. With the other components of the current account (trade in services, compensation to employees, other primary income, secondary income) and the capital account, the solution to get a consistent global estimate is not so clear, so we opt for decreasing credit (debit) proportionally whenever the net global is different than zero and we report that results hold without such a correction.

5 Core distributional variables

Based on the concept of core countries and core variables, we extend this approach to distributional series, with a particular focus on pre-tax income. Table 5 provides distributional data for 216 core countries from 1980 onwards, while Table 4 presents data for core territories in the years 1820, 1850, and 1880. Subsequent data is shown at 10-year intervals from 1900 to 1970. The data includes bracket averages, bracket shares, thresholds, and Gini coefficients.

Compared to prior annual updates from the World Inequality Lab (WIL) on income inequality, this technical note adds 40 new countries to the series. For these countries, we could not provide income inequality estimates based on surveys or tax data. Instead, we used simplified assumptions to generate distribution series.

The series for these 40 countries is based on the distribution data from the core territories they belong to: Other Latin America, Other North America & Oceania, and Other Western Europe, as outlined in Chancel and Piketty (2021):

- **Other Latin America**: Antigua and Barbuda; Anguilla; Aruba, Barbados; Bonaire, Sint Eustatius and Saba; Curacao; Dominica; Grenada; Saint Kitts and Nevis; the Cayman Islands; Saint Lucia; Montserrat; Puerto Rico; Sint Maarten (Dutch part); the Turks and Caicos Islands; Saint Vincent and the Grenadines; the British Virgin Islands.
- **Other North America & Oceania**: Bermuda; Fiji; Micronesia; Greenland; Kiribati; the Marshall Islands; New Caledonia; Nauru; the French Polynesia; Palau; Solomon Islands; Tonga; Tuvalu; Vanuatu; Samoa.
- **Other Western Europe**: Andorra; Guernsey; Gibraltar; the Isle of Man; Jersey; Liechtenstein; Monaco; San Marino.

Table 4

The World as the Sum of 33 Core Territories: core distributional variables (1820-2022)

Regions	Core Territories	Core Variables	Years
East Asia (3)	China, Japan, Other East Asia	**Share of Pre-tax national income** (sptinc) **Average of Pre-tax national income** (aptinc) **Threshold of Pre-tax national income** (tptinc) **Gini of Pre-tax national income** (gptinc)	**1820**, **1850**, **1880**, **1900-1980 (10 years interval)**, **1980-2023 (annual)**
Europe (8)	Britain, France, Germany, Italy, Spain, Sweden, Other Western Europe, Eastern Europe		
Latin America (6)	Argentina, Brasil, Chile, Colombia, Mexico, Other Latin America		
Middle East/North Africa (4)	Algeria, Egypt, Turkey, Other Middle East/North Africa		
North America/Oceania (5)	USA, Canada, Australia, New Zealand Other North America/Oceania		
Russia/Central Asia (2)	Russia Other Russia/Central Asia		
South/South-East Asia (3)	India, Indonesia, Other South/Southeast Asia		
Sub Saharan Africa (2)	South Africa, Other Sub-Saharan Africa		

Interpretation. For the purpose of long-run series, WID divides the world into 8 regions and 33 core territories (24 countries + 9 subregions). The core distributional variables include the full distribution of pretax income for all g-percentiles (shares, thresholds, average income) as well as Gini coefficients. All core variables are available for all core territories for years 1820, 1850, 1880, at 10 years interval from 1900 to 1980 and for all years from 1980 to 2022 (annual series).

Table 5

The World as the Sum of 216 Core Countries: core distributional variables (1980-2022)

Regions	Core Territories	Core Variables	Years
East Asia (8)	China, Hong Kong, Japan, Macao, Mongolia, North Korea, South Korea, Taiwan	**Share of Pre-tax national income** (sptinc) **Average of Pre-tax national income** (aptinc) **Threshold of Pre-tax national income** (tptinc) **Gini of Pre-tax national income** (gptinc) **Share of Post-tax national income** (sdiinc) **Average of Post-tax national income** (adiinc) **Threshold of Post-tax national income** (tdiinc) **Gini of Post-tax national income** (gdiinc)	**1980-2022 (annual)**
Europe (47)	Albania, Andorra, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Gibraltar, Greece, Guernsey, Hungary, Iceland, Ireland, Isle of Man, Italy, Jersey, Kosovo, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Moldova, Monaco, Montenegro, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, San Marino, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom		
Latin America (43)	Anguilla, Antigua and Barbuda, Argentina, Aruba, Bahamas, Barbados, Belize, Bonaire, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Curaçao, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Montserrat, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Sint Maarten, Suriname, Trinidad and Tobago, Turks and Caicos Islands, Uruguay, Venezuela		
Middle East/North Africa (20)	Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestinian Territories, Qatar, Saudi Arabia, Syria, Tunisia, Turkey, United Arab Emirates, Yemen		
North America/Oceania (19)	Australia, Bermuda, Canada, Fiji, French Polynesia, Greenland, Kiribati, Marshall Islands, Micronesia, Nauru, New Caledonia, New Zealand, Palau, Samoa, Solomon Islands, Tonga, Tuvalu, United States, Vanuatu		
Russia/Central Asia (11)	Armenia, Azerbaijan, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan.		
South/South-East Asia (19)	Afghanistan, Bangladesh, Bhutan, Brunei, Cambodia, India, Indonesia, Laos, Malaysia, Maldives, Myanmar, Nepal, Pakistan, Papua New Guinea, Philippines, Singapore, Sri Lanka,Thailand, Timor-Leste, Vietnam		
Sub Saharan Africa (54)	Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Djibouti, Equatorial Guinea, Eritrea, Eswatini, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Republic of the Congo, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, South Sudan, Sudan, Tanzania, Togo, Uganda, Zambia, Zimbabwe		

Interpretation. For the purpose of long-run series, WID divides the world into 8 regions and 216 core countries. The core distributional variables include the full distribution of pretax income for all g-percentiles (shares, thresholds, average income) as well as Gini coefficients. All core variables are available for all core territories for years 1820, 1850, 1880, at 10 years interval from 1900 to 1980 and for all years from 1980 to 2022 (annual series).

References

Blanchet, T., & Chancel, L. (2016). National accounts series methodology. *WID.world WORKING PAPER SERIES N°(2016/1)*.

Blanchet, T., Chancel, L., Flores, I., Morgan, M., Alvaredo, F., Atkinson, A. B., ... others (2021). Distributional national accounts guidelines. *Methods and Concepts Used in the World Inequality Database. World Inequality Lab*.

Blanchet, T., Chancel, L., & Gethin, A. (2022). Why is europe more equal than the united states? *American Economic Journal: Applied Economics*, *14*(4), 480–518.

Chancel, L., Moshrif, R., Piketty, T., & Xuereb, S. (2023). Historical inequality series on wid.world - update. *World Inequality Lab – Technical Note N° 2023/01*.

Chancel, L., & Piketty, T. (2021). Global income inequality, 1820–2020: the persistence and mutation of extreme inequality. *Journal of the European Economic Association*, *19*(6), 3025–3062.

Lane, P. R., & Milesi-Ferretti, G. M. (2018). The external wealth of nations revisited: international financial integration in the aftermath of the global financial crisis. *IMF Economic Review*, *66*, 189–222.

Nievas, G., & Piketty, T. (2024). *Allocating missing foreign wealth and missing foreign income: Sources and methods used in wid.world* (Tech. Rep. No. Technical Note 2024/03). World Inequality Lab. Retrieved from https://prod.wid.world/www-site/uploads/2024/07/WorldInequalityLab_TechnicalNote_2024_03_Allocating-Missing-Foreign-Wealth-And-Missing-Foreign-Income_Final.pdf

Nievas, G., & Sodano, A. (2024). Has the us exorbitant privilege become a rich world privilege? rates of return and foreign assets from a global perspective, 1970-2022. *World Inequality Lab – Working Paper N° 2021/14*.

Piketty, T., & Zucman, G. (2014). Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly journal of economics*, *129*(3), 1255–1310.