WID INCOME AND WEALTH DISTRIBUTIONAL SERIES UPDATED AND EXTENDED COVERAGE, 1800-2024

MANUEL ARIAS-OSORIO
LUIS BAULUZ
PIERRE BRASSAC
LUCAS CHANCEL
CLARA MARTINEZ-TOLEDANO
ROWAIDA MOSHRIF
THOMAS PIKETTY

TECHNICAL NOTE N°2025/10

DECEMBER 2025

WORLD INEQUALITY LAB

WID Income and Wealth Distributional Series
Updated and Extended Coverage, 1800-2024

Manuel Arias-Osorio[1], Luis Bauluz[2], Pierre Brassac[3], Lucas Chancel[4],
Clara Martinez-Toledano[5], Rowaida Moshrif[1], Thomas Piketty[1]

This version: December 28, 2025

Abstract. This technical note describes the new extended set of distributional series available in the World Inequality Database (WID), including series on pretax income distribution, posttax income distribution and wealth distribution. All benchmark series cover the entire world, which is defined as a sum of a stable set of 216 core countries and jurisdictions over the 1980-2024 period and as the sum of a stable set of 57 core territories (48 main countries + 9 residual regions) over the 1820-2024 period. Our benchmark series are annual over the 1980-2024 period and cover years 1820, 1850, 1880, 1900, 1910, and every ten years until 1980 over the 1820-1980 period. We also provide the list of countries-years for which we offer additional coverage.

[1]WIL (World Inequality Lab), PSE (Paris School of Economics). [2]WIL & CUNEF (Madrid). [3]WIL & UC3M (Madrid). [4]WIL & Sciences Po (CRIS). [5]WIL & Imperial College (London).

1. Introduction

This technical note describes the new extended set of distributional series available in the World Inequality Database (WID), including series on pretax income distribution, posttax income distribution and wealth distribution. All benchmark series cover the entire world, which is defined as a sum of stable set of 216 core countries and jurisdictions over the 1980-2024 period and as the sum of stable set of 57 core territories (48 main countries + 9 residual regions) over the 1820-2024 period. Our benchmark series are annual over the 1980-2024 period and cover years 1820, 1850, 1880, 1900, 1910, and every ten years until 1980 over the 1820-1980 period. We also provide the list of countries-years for which we offer additional coverage.

The rest of this technical note is organized as follows. We describe in section 2 the historical and geographical coverage of our benchmark income and wealth distribution series. We describe in section 3 the set of additional countries-years which are available in our database. Finally, we provide concluding comments in section 4. The online replication package includes all tables, figures and computer codes. The central WID code on distributional series is available on WID GitHub page.

2. WID Benchmark Distributional Series

The structure of WID benchmark distributional series is described on Table 1. These new extended series are used by Andreescu et al (2025) in order to revisit the relation between inequality and development.

The main novelty is that we now provide the same geographical and historical coverage for income and wealth distributional series. Namely, WID benchmark distributional series for pretax income, posttax income and net household wealth now cover all 216 WID core countries and jurisdictions for all years over the 1980-2024 period. Prior to 1980, WID benchmark income and wealth distributional series are restricted to 57 core territories (48 main countries + 9 residual regions) and to a selected number of benchmark years (namely 1820, 1850, 1880, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970) (see Table 2). For now, posttax series are not available in a systematic manner in WID prior to 1980. They will be completed in the near future (see section 4).

The 48 main countries were chosen on the basis of population size, GDP, regional representativity and data availability. Throughout the 1800-2025 period, the 48 main

countries cover about 85-90% of the world population and GDP, while the 9 residual regions cover 10-15% (see Nievas and Piketty (2025)).

WID benchmark series always all cover all income and wealth shares, average income and wealth, income and wealth thresholds for all 127 g-percentiles[1] and all benchmark years countries (variables sptinc, aptinc, tptinc for pretax income, variables sdiinc, adiinc, tdiinc for posttax income, variables shweal, ahweal, thweal for wealth). WID benchmark series also include for all countries-years a number of inequality indicators T10/B50 ratio (r), Gini coefficients (g), inverted Pareto-Lorenz coefficients (b).

WID benchmark series for pretax income, posttax income and personal wealth are always equal-split series (ind=j). They always cover both per capita (ag=999) and per adult (ag=992) income and wealth concepts. By default, the series for normalized distributions (income shares and wealth shares) are the same for per capita and per adult income and wealth.[2]

All WID benchmark series strictly follow the latest edition of DINA Guidelines (see Chancel et al (2025)). Generally speaking, historical pretax income distribution series covering the 1820-1970 period are based upon Chancel and Piketty (2021), with two main novelties. First, the series were extended to the new set of 57 core territories. In the absence of any specific country study, we attribute to the country the normalized distribution of the corresponding residual region. Next, the full historical series for all 127 g-percentiles were homogenized using gpinter (Blanchet et al (2022)).[3]

Regarding posttax income distribution, for now the series that we have made available online in the World Inequality Database only cover the 1980-2024 period. The pre-1980 posttax series used in Andreescu et al (2025) are based upon simplifying assumptions. Namely, we assume that pretax and posttax inequality levels are the same until 1910, and then that the magnitude of redistribution evolves linearly between 1910 and 1980 at the country level.[4] This can be justified by the general evolution of public revenue and expenditure over the past century (from less than 10% of GDP in

[1] Generalized percentiles (or g-percentiles) refer to the 127 quantiles defined by the bottom 99 percentile, the 9 tenth-of-percentile at the top 1%, the 9 hundredth-of-percentile at the bottom of the top 0.1% and the 10 thousandth-of-percentile within the top 0.01%. Lower threshold and average income for each of the 127 g-percentiles provide the basic distributional data that is being stored in WID.world for each country-year. Country-level and sub-regional-level data by g-percentile can be aggregated up to the regional and world levels using the gpinter (generalized Pareto interpolation) facility available online at wid.world/gpinter. See Blanchet, Fournier and Piketty (2022).

[2] For now, per capita and per adult series differ only for a small subset of countries-years.

[3] See online replication package.

[4] See online replication package and computer code associated to Andreescu et al (2025) for all details.

all countries until 1910, up to 40-50% of GDP in recent decades in rich countries) An alternative estimation strategy consists of using detailed country-level series on public expenditure and revenue by categories (see Bharti et al (2025) and explicit assumptions on their distributional incidence (based upon post-1980 observed profiles and other sources; see Fisher-Post and Gethin (2023) and Gethin (2024)). Preliminary estimates suggest that this would make very little difference. We plan to release extended posttax series following such a methodology in the future.

Regarding wealth distribution series, we have put together a large number of historical estimates published by WIL fellows and other researchers covering the entire 1820-2024 period, and we have made imputations for missing countries-years on the basis of their position in terms of income distribution (see section 2.1. below).

2.1. Imputation Method for Missing Wealth Distribution

We have put together a large number of historical estimates published by WIL fellows and other researchers covering the entire 1820-2024 period. For missing countries-years, we use an imputation method that is similar to the method described in Bajard et al (2025), and that is based on the observation that top income shares are a strong predictor of top wealth shares in countries with missing wealth data. The key difference is that we now apply this method not only to the 1995-2024 sub-period but to the entire 1820-2024 period.

More precisely, we describe on Table 4 the set of countries-years with wealth distribution estimates which we use for imputing missing observations. As one can see from Figure 1, the relationship between top 10% income shares and top 10% wealth shares is very strong. It is also striking to see that the relationship holds within each sub-period, namely 1800-1909, 1910-1979 and 1980-2024 (see Figures 2-4). We apply two different bandwidths over the 1820-1979 ($h = 0.11$) and 1980-2024 ($h = 0.27$) periods to predict the distribution of wealth (see Bajard et al (2025)). It ensures the best possible fit with existing recent estimates while producing reasonable historical ones. These two bandwidths are both small, exploiting the strong correlation found between wealth and income inequality. We have tried several alternative methods and this appears to be the most robust method.[5]

[5] See online replication package.

2.2. Imputation Method for MER Historical Distributions

Using the PPP distributions of the residual regions, we construct MER versions for benchmark years prior to 1980. Following, these residual regions can be aggregated with the core countries for estimating MER distributions for the continental regions and the world, exactly as it is done for from 1980 on.

First, we collapsed the distribution of each residual region in 1980 into the percentile structure described above:

i. Standard structure: 0–10, 10–25, 25–50, 50–75, 75–90, 90–99, and 99–100 which will later be homogenized using gpinter tool.

Given data availability in 1980, we observe percentile-level values in both PPP and MER terms, allowing us to compute the differences in shares between the two valuation currencies for each residual region.

We then applied a two-step adjustment procedure. First, we added the calculated differences to the PPP shares for each percentile and for each benchmark year from 1820 to 1970. This produces percentile series that follow the PPP dynamics but are anchored at the observed MER levels in 1980. We then recalculated bracket averages based on the adjusted shares.

While this simultaneous adjustment performs well in most cases, in a few regions—most notably Other South and Southeast Asia and Other North America—it led to non-monotonically increasing bracket averages at percentiles 10, 25, and 50 (sometimes one of them, sometimes all three). The second step of the adjustment involved **dropping** these non-monotonic observations and replacing them using an interpolation method slightly modified from that used in the WID central code to correct the pretax bottom 20 percentiles for certain countries. The interpolation formula used is

$$a_p = a_L + (a_U - a_L) * (\frac{p - p_L}{p_U - p_L})^{1+\alpha}$$

where a_L and a_U are the lower and upper observed bracket averages (i.e., the last ones satisfying the monotonicity condition), p_L and p_U are the corresponding

percentiles, and p is the percentile being interpolated. The parameter α controls the concavity of the interpolated curve and is computed as[6]

$$\alpha = \frac{a_{99}}{mean(a_{75}, a_{90}, a_{75})} - 1$$

To preserve internal consistency, new shares were estimated from the corrected bracket averages, and the entire series was rescaled so that shares sum to one and averages match the average national income.

For thresholds, we estimated the difference between the new MER bracket averages and the original PPP bracket averages at $p = 0$, and added this difference to the PPP threshold value. For the other percentiles, threshold values were set to the midpoint between adjacent bracket averages, in line with standard WID methods for regional aggregation.

These adjusted standardized series were then introduced into the gpinter tool to generate a generalised distribution. The normalized full distributions were subsequently processed by the WID central code, which calibrated averages and thresholds using available macroeconomic data for the residual regions. Finally, these were aggregated with core countries to construct continental and World regional distributions (also calibrated using WID macroeconomic data). The result is a complete set of pretax income and wealth distributions in both MER and PPP terms for benchmark years from 1820 to 1970 for the 57 core territories, as well as full coverage for all countries and regions from 1980 onward.

3. Additional Countries-Years Available in WID

In addition to the set of benchmark countries and years described on Tables 1-2, WID income and wealth distribution series also cover additional years for a number of countries (see Table 3).

We should make clear that all series described on Table 3 cover all income and wealth shares, average income and wealth, income and wealth thresholds for all 127 g-percentiles. All series were homogenized using gpinter.[7] Also, in the same way as the benchmark series described on Tables 1-2, all additional series described on Table 3 are always equal-split series (ind=j). They systematically cover both per capita

[6] We use the bracket averages at percentiles 75, 90, and 99 as reference points because they always satisfy the monotonicity condition.

[7] See online replication package.

(ag=999) and per adult (ag=992) income and wealth concepts. By default, the series for normalized distributions (income shares and wealth shares) are the same for per capita and per adult income and wealth.[8] They also include for all countries-years a number of inequality indicators T10/B50 ratio (r), Gini coefficients (g), inverted Pareto-Lorenz coefficients (b).

Note that full WID distributional series available on wid.world also cover other observation units (individuals, tax units, etc.) and income concepts (factor income, fiscal income, etc.), but only for a relatively small and irregular subset of countries-years-percentiles (not listed here). We do not recommend to use these incomplete series for economic or historical analysis, as they are not representative of the database as a whole and may not be fully comparable across countries-years. In particular, the notion of fiscal income varies with the tax legislation, so that fiscal income series are not fully comparable across countries and years. They have been gradually superseded by DINA pretax and posttax national income concepts, which have been designed in order to allow for more meaningful comparisons.

3.1. Imputation Method for Distributional data for Additional Countries-Years

Following the methodology exposed in Piketty, Saez and Zucman (2018), at the very top of the distributions the fiscal income - where individuals are more likely to be observed and subject to tax collection-, can be a strong predictor of the top of the pretax income. Based on this assumption, part of the additional country–year observations (see Table 3) for which pretax income data were unavailable prior to 1980 were constructed using information on top percentiles from the fiscal income distribution. The fiscal income data (with a fiinc code in WID) was used either to adjust the benchmark trajectory or to refine the trajectory of the residual region to which the country–year observations belong. The implementation of these new data in the WID also required extending the macroeconomic variables necessary for the calibration of the distributions, namely the price index, the exchange rate, and net national income.

After identifying the countries to which this process would be applied (with some availability of top fiscal income shares), we retrieved the historical distributions for the benchmark years for all core territories and extended the distributions of the residual regions to the corresponding countries for years prior to 1980 (for instance, the historical distribution of “Other Eastern Europe” was extended to Bulgaria, Hungary, Croatia, and Poland).

[8] For now, per capita and per adult series differ only for a small subset of countries-years.

As a result, both core and non-core countries have income distributions, but only for benchmark years. Since our objective is to estimate distributions for the intervening years, we interpolated the series between these benchmarks in order to ensure continuity with the historical regional series. The resulting full series (*baseline benchmark-based shares*) were collapsed into the standard percentile structure (presented in section 2.1) and a detailed structure – 0–10, 10–25, 25–50, 50–75, 75–90, 90–99, 99–99.9, and 99.9–100 —which is used as a last resort to complete Top 1% fiscal income data.

Parallelly, we extracted the top values of the fiscal income distributions. Owing to the scarcity of these data, the available brackets are heterogeneous across countries, differing in percentile cutoffs as well as population and age groups. In some cases, the data are directly available for the Top 1%, as desired, for population–age groups such as 992i, 992j, or 992t. In other cases, we completed the data (Top 1% share) using two methods.

First, we used bracket averages, taking advantage of their correspondence with income shares and re-estimated shares. Second, we calculated ratios between Top 10% and Top 1%, Top 1% and Top 0.1%, and Top 0.1% and Top 0.01% in the pretax income distributions for each country-year. These ratios were then interpolated between benchmark years—most often relying on ratios observed in non-core countries, typically from 1980 onward—to proportionally estimate the Top 10%, Top 1%, and Top 0.1% for all relevant countries and years.

For Bulgaria, Cameroon, Ghana, Tanzania, and Uganda, fiscal income observations are located too far from 1980 (around 30 years before), making extrapolation not trustable. For these countries, we instead collapsed into the detailed percentile structure described above, and we applied the ratios calculated from the corresponding residual regions. As a result, we obtain full coverage of the p99p100 sfiinc for all intended countries, as shown below.

- Sfiinc992i AR, AU, CA, ES, FI, GB, GH, GR, HU, IT, JP, KE, KR, MW, NO, NZ, SC, SG, TZ, UG, ZA, ZM, ZW
- Sfiinc992j BG, CL, CN, FR, HR, MY, NO, PL, RU, US
- Sfiinc992t CH, CM, DE, DK, DZ, FI, FR, GB, HU, ID, IE, IN, MU, NL, NZ, PT, SE, TN, TW, US, VN, ZA

Using these two elements—the fiscal income Top 1% share and the interpolated historical pretax series— we compute adjustment ratios at the benchmark years and these ratios interpolate to obtain a continuous measure of the fiscal–pretax income adjustment for the p99p100 percentile. Multiplying the fiscal income shares by these ratios yields a new series that exactly matches pretax income at the benchmark years (thus remaining consistent with the regional pretax estimations available for those years) and, for the intervening years, follows pretax income levels while varying in line with fiscal income dynamics (*directly adjusted shares*).

While this outcome is consistent with our objective, we observe that, for country–year observations where both series are available (e.g. the United States, Canada, India, and France), fiscal income series often display pronounced year-to-year fluctuations that are not present in pretax income series, even over the same years. As a result, this direct adjustment may induce exaggerated variability in the estimated pretax Top 1% pretax income shares. To mitigate this issue, we apply a robust shrinkage correction to the raw difference between the *baseline benchmark-based shares* and the *directly adjusted shares*. This correction is implemented by adding a penalization term to the estimation.

The penalization term is computed as described in the equation below

$$penalisation = \frac{raw\ difference}{1 + \left|\frac{raw\ difference}{\lambda * StandardDeviation_{coutry}(raw\ difference)}\right|}$$

As can be seen, the degree of shrinkage applied to the series depends on the parameter λ. The optimal value of λ was determined numerically; that is, by pooling all country–year observations for which both pretax income shares and fiscal income shares are available, computing penalized adjustments on the *directly adjusted shares* over a grid of possible λ values, and comparing the mean squared error of the difference between each penalized series and the observed pretax distributions. This procedure yielded a hypothetical optimal λ of 4.6 for all country–year observations.

Using this penalization, we produce, for each country, a regularized (*shrinkage-adjusted*) series, which is retained as the final Top 1% value. Given this series, the next step consists of rebalancing the entire distribution within the standard percentile structure described above. This involves computing a proportional bracket threshold for the new Top 1% that is consistent with the revised share, and then redistributing the difference between the original and adjusted Top 1% values across the remaining

percentiles, in proportion to their population weights within each relevant tranche and adding up 100%.

Finally, the resulting series are homogenized using gpinter. The detailed final availability of full series can be seen in the Table 3.

3.2. Imputation Method for Additional Countries-Years Macroeconomic data

The calibration of the normalized bracket averages and thresholds of the income distributions using macroeconomic data relies on the availability of net national income series in the WID. Although this variable is fully available for the main countries covered through the implementation of Nievas & Piketty's (2025) World Historical Balance of Payments Database, this is not the case for non-core countries. To construct this variable for those countries, we generate series of price indices, MER USD exchange rates (and their counterparties PPP USD exchange), and gross national product.

Price index: This variable is available for the majority of countries in the WID from 1950 onward. To extend the series backward, we relied on several additional data sources, including the Consumer Price Index from the Central Statistics Office of Ireland (available from 1922), Reinhart and Rogoff (2011) data for all covered countries and years, the consumer price indices reported in International Historical Statistics by Mitchell (2013), and inflation data from Franses and Janssens (2018) for African countries.

Following the same methodology as in the main WID code, we computed the logarithms of the raw price level series and calculated their annual differences. These differences were then chained to cover as many years as possible for each country and accumulated backward from 2024 to the earliest available year. Taking the exponential of the resulting cumulative sums yields the price index.

Also, we applied the following exceptions: For the Central Statistics Office data, we assumed that inflation in 1921 was equal to that observed in January 1922. In Reinhart and Rogoff (2011) (i) Poland's inflation for 1940–1945 was completed using Germany's inflation rates; and (ii) Hungary's inflation in 1946 was interpolated. Finally, Bulgaria's missing observations prior to 1925 were completed using the regression estimated by Ljungberg (2025) on Hungary's and Austria's consumer price indices, setting the same parameters found by the author.

Exchange rates: This variable was manually compiled for countries with available fiscal income data, using information from World Bank documents, central banks, academic journals, other public reports and other primary sources. In a more systematic approach, exchange rate series for the remaining countries were retrieved from Harald Müller's tabulations available at liganda.ch.

We also ensured data availability wherever possible by imputing common exchange rates across countries. For example, we used the exchange rates of former colonizing countries (typically core countries) for their former colonies prior to independence, and extended exchange rate series from core countries to other countries sharing the same currency.

GDP: Some non-core countries with available fiscal income data already have gross domestic product series in the WID prior to 1970. To extend this coverage backward, we relied on data from the Maddison Project (2023) for the countries included, which are already used in the WID but over a shorter time span. We interpolated missing years in order to construct complete series back to 1900.

The Maddison database reports output aggregates in 2010 international dollars, a currency unit that is not compatible with WID series. For matching the series, following the methodology implemented in the WID core code, we transformed both the Maddison and WID series into logarithms, computed annual growth rates, and reconstructed GDP backward from the last observed retaining always the WID available observations. This procedure yields unified GDP series expressed in constant 2024 prices. However, while the trend underlined is useful, because the WID does not provide exchange rates to "international dollars," this approach may overlook changes in inflation and exchange rates, potentially leading to over- or undervaluation of GDP levels when expressed in constant prices.

To address this issue, we ensured that the extended country-year GDP series remained coherent with amounts of their corresponding residual regions, as estimated in Nievas and Piketty (2025). For this we partially replicated the way the regional aggregations in the WID main code are constructed adding values in current USD values. As so, we first converted all country–year GDP observations with available

exchange rate and price index data into current USD. Then, we rescaled the sum of GDP across countries within each residual region (QM, OC, OE, and OI) to match the regional totals reported in Nievas and Piketty (2025) for all years from 1970 to 1900.

As we move backward in time and country-level observations progressively disappear, missing country values are imputed sequentially by fixing each country's GDP share within its residual region at its last observed value (year n + 1). The regional aggregate for year n is then rescaled accordingly, along with its components—the country aggregates. This procedure ensures that countries with available data transmit their observed dynamics to their share of the regional aggregate, while countries without data remain static and adjust residually in the opposite direction, so that the overall regional total remains consistent with the target value reported in Nievas and Piketty (2025).

The outcome of this procedure is a complete set of theoretical GDP series in current USD from 1900 onward for all countries. Only countries with available exchange rate and price index data prior to 1970 can subsequently be converted back into constant local currency units (LCU) in 2024 prices. Moreover, only the GDP extensions corresponding to targeted countries, and starting from the first year with available fiscal income data, are ultimately imported into the WID.

Net National income: Finally, we applied a similar approach to net national income (NNINC). We first extended historical NNINC series by extrapolating the last observed NNINC-to-GDP ratio and generating proportional historical NNINC estimates. Second, these series were then adjusted with the same method described to the GDP, to match the residual-region NNINC aggregates reported in Nievas and Piketty (2025). As before, only the data corresponding to the targeted country–year observations were retained.

4. Concluding Comments

In this technical note, we have presented the new extended set of WID income and wealth distributional series. There are still many limitations which will need to be addressed in the future. In particular, our historical coverage for income and wealth

distribution series has been substantially improved, but it could be further extended in future years. The posttax series could and will also be extended backward. For now, posttax series are not available in a systematic manner in WID prior to 1980. They will be completed in the near future.[9] This will be explained in future WIL technical notes.

References

M. Andreescu, M. Arias-Osorio, L. Bauluz, N. Bharti, P. Bothe, P. Brassac, L. Chancel, J. Dietrich, D. El Hariri, M. Fisher-Post, I. Flores, V. Gabrielli, A. Gethin, R. Gomez-Carrera, S. Hong, T. Jenmana, R. Loubes, C. Martinez-Toledano, Z. Mo, C. Mohren, M. Morgan, R. Moshrif, S. Mutti, T. Neef, G. Nievas, M. Odersky, T. Piketty, A.S. Robillard, E. Saez, A. Sodano, A. Somanchi, L. Yang, G. Zucman, A. Zuniga-Cordero, "Equality and Development A Comparative & Historical Perspective 1800-2025", WIL Working Paper 2025/25

F. Bajard, L. Bauluz, P. Brassac, L. Chancel, C. Martinez-Toledano, T. Piketty, A. Sodano, "Global Wealth Inequality on WID.world Estimates and Imputations", WIL Technical Note 2025/01

N. Bharti, A. Gethin, T. Jenmana, Z. Mo, T. Piketty L. Yang. "Human Capital, Unequal Opportunities and Productivity Convergence A Global Historical Perspective 1800–2100." WIL Working Paper 2025/15

T. Blanchet, J. Fournier, T.Piketty, "Generalized Pareto Curves Theory and Applications", Review of Income and Wealth, 2022, vol.68(1), p.263-288 (WIL WP 2017/03)

L. Chancel, T. Piketty, "Global Income Inequality, 1820-2020 The Persistence and Mutation of Extreme Inequality", *Journal of the European Economic Association*, 2021 (WIL Working Paper 2021/19)

[9] Andreescu et al (2025) assume that pre-tax and post-tax inequality levels are the same until 1910, and then that the magnitude of redistribution evolves linearly between 1910 and 1980 at the country level. This simplifying assumption can be justified by the fact that redistribution was relatively small in all countries prior to 1910 and rose significantly between 1910 and 1980. An alternative estimation strategy consists of using newly-constructed country-level 1800-2025 series on public expenditure and revenue by categories (see Bharti et al (2025)) and explicit assumptions on their distributional incidence (based upon post-1980 observed profiles and other sources). Preliminary estimates suggest that this would make relatively little difference (as far as the broad evolutions are concerned). More research on this important issue will be released in the future.

L. Chancel, I. Flores, R. Moshrif, G. Nievas, T. Piketty (editors), *Distributional National Accounts Guidelines Methods and Concepts used in the World Inequality Database*, WIL, 2025

M. Fisher-Post, A. Gethin, "Government Redistribution and Development Global Estimates of Tax and Transfer Progressivity 1980-2019", WIL WP 2023/17

A. Gethin, "A New Database of General Government Revenue and Expenditure by Function, 1980-2023", WIL WP 2024/01

G. Nievas, T. Piketty, "WID National Accounts Series Updated and Extended Coverage 1800-2023", WIL Technical Note 2025/02

T. Piketty, E. Saez, G. Zucman, "Distributional National Accounts: Methods and Estimates for the United States", *Quarterly Journal of Economics*, 2018, vol.133(2), p 553-609

Sources for Macroeconomic variables extensions

Borg, M., *L'industrie du papier en Finlande*, Doctoral thesis, University of Neuchâtel, 1932.

Central Statistics Office (CSO), Consumer Price Index (1922-1967), Ireland, Table CPM07, data retrieved from data.cso.ie, accessed 2025.

P. Franses, E. Janssens, "Inflation in Africa, 1960–2015", Journal of International Markets, 2018, Institutions and Money, vol.57, pages 261-292.

J. George, "*The Malayan Currency Board (1938–1967)*", *Studies in Applied Economics*, SAE/No. 53, March ,2016, Johns Hopkins Institute for Applied Economics, Global Health, and Study of Business Enterprise.

E. Lafaye de Micheaux, "British Malayan Capitalism, 1874–1957 An Economic-Historical Perspective," *Economic History Malaysia*, online article, accessed 2025, https//www.ehm.my/publications/articles/british-malayan-capitalism-1874–1957-an-economic-historical-perspective

J. Ljungberg, “European Consumer Price Indices since 1870”, Cliometrica, 2024, vol.19, pages 29-80

B. Mitchell, “International Historical Statistics”, 2013, Palgrave Macmillan, doi https//doi.org/10.1057/978-1-137-30568-8

H. Müller (n.d.), “*End-year Forex Rates” Tabulations*, Liganda.ch, online database, accessed 2025.

C. Reinhart, K. Rogoff, “From Financial Crash to Debt Crisis,” *American Economic Review, 2011,* vol.11(5)

K. Schuler, *Historical Financial Statistics*. Center for Financial Stability, 2015. Available at Center for Financial Stability, accessed 2025.

Obserwator Finansowy, “Building the Strength of the Polish Zloty,” online article, *ObserwatorFinansowy.pl*, 2018, accessed 2025.

Magyar Nemezeti Bank. Exchange rates data retrieved from https//www.mnb.hu/, accessed 2025.

Annex 1. Sources of Exchange rate by country

- Bulgaria 1913–1947 Schuler (2015). 1948–1970 Müller (n.d.).
- Switzerland 1936–1970 Müller (n.d.).
- Cameroon Exchange rate extended from France, adjusted by a factor of 1.7.
- Finland 1914–1924 Borg (1932). 1925–1938 Müller (n.d.). Indexed to the United Kingdom exchange rate. 1939–1970 Müller (n.d.).
- Ghana 1943–1957 Indexed to the United Kingdom. 1958–1970 Indexed to Kenya.
- Greece 1913–1943 Schuler (2015). 1944 Set equal to the 1943 value. 1945–1970 Müller (n.d.).
- Croatia 1965–1970 Projected backward from the 1970 observation.

- Hungary 1900–1913 Schuler (2015), using data for Austria. 1914–1919 Müller (n.d.), using data for Austria. 1919–1948 Müller (n.d.). 1949–1970 Magyar Nemzeti Bank.
- Ireland 1927–1970 Indexed to the United Kingdom.
- Mauritius and Seychelles 1914–1935 Indexed to India. 1936–1970 Indexed to the United Kingdom and adjusted by a factor of 40/3, following Schuler (2015).
- Malawi, Tanzania, Uganda, and Zambia Indexed to the United Kingdom.
- Malaysia and Singapore 1903–1952 Indexed to the United Kingdom, adjusted following George (2016). 1953–1970 Indexed to the United Kingdom, adjusted following Lafaye de Micheaux (n.d.).
- Poland 1918–1924 Obserwator Finansowy (2018). 1924–1970 Müller (n.d.).
- Tunisia Indexed to France and adjusted following Müller (n.d.).
- Zimbabwe Treated as in the WID, with values expressed directly in USD.

Table 1. WID Benchmark Distributional Series: Geographical and Historical Coverage

Pretax income (sptinc, aptinc, tptinc) **Posttax income** (sdiinc, adiinc, tdiinc) **Net household wealth** (shweal, ahweal, thweal) (equal-split, per capita and per adult)	All 216 core countries	1980-2024 (annual series)	All 127 g-percentiles
	All 57 core territories (48 main countries + 9 residual regions)	1820, 1850, 1880, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980-2024 (annual series)	All 127 g-percentiles

WID benchmark distributional series for pretax income, posttax income and net household wealth cover all 216 core countries and jurisdictions for all years over the 1980-2024 period, and are restricted to 57 core territories (48 main countries + 9 residual regions) and to a selected number of benchmark years over the 1800-1980 period (1820, 1850, 1880, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970). WID benchmark series also include for all countries-years a number of inequality indicators: T10/B50 ratio (r), Gini coefficients (g), inverted Pareto-Lorenz coefficients (b). See wid.world/code-dictionary for variable names and the list of core countries and territories. WID series also cover additional years for a number of countries (see Tables 3). WID benchmark series for pretax income, posttax income and personal wealth are always equal-split series (ind=j). They always cover both per capita (ag=999) and per adult (ag=992) income and wealth concepts. By default, the series are the same for per capita and per adult income and wealth. They differ only for a subset of countries-years. WID distributional series also cover other observation units (individualistic, tax units, etc.) and income concepts (factor income, fiscal income, etc.), but only for a relatively small and irregular subset of countries-years-gpercentiles. **Note**. For now, posttax series are not available in a systematic manner in WID prior to 1980. They will be completed in the near future.

Table 2. Core Territories Used in WID Benchmark Historical Series (57 core territories = 48 main countries + 9 residual regions)	
East Asia (5)	China, Japan, South Korea, Taïwan Other EASA
Europe (11)	Britain, Denmark, France, Germany, Italy, Netherlands, Norway, Spain, Sweden, Other W.EUR, Other E.EUR
Latin America (6)	Argentina, Brasil, Chile, Colombia Mexico, Other LATAM
Middle East/ North Africa (8)	Algeria, Egypt, Iran, Morocco, Saudi Arabia, Turkey, UAE, Other MENA
North America/ Oceania (5)	USA, Canana, Australia, New Zealand Other NAOC
Russia/ Central Asia (2)	Russia Other RUCA
South/South-East Asia (9)	Bengladesh, India, Indonesia, Myanmar, Pakistan, Philipinnes, Thailand, Vietnam, Other SSEA
Sub-Saharan Africa (11)	DR Congo, Ethiopa, Kenya, Ivory Coast, Mali, Niger, Nigeria, Rwanda, Sudan, South Africa, Other SSAF
For recent decades (1980-2024), WID series cover all 216 WID core countries and jurisdictions for all years. Regarding long-run historical series (1800-1980), WID series generally cover all 57 core territories (48 main countries + 9 residual regions) for all years (national accounts) or for a selected set of benchmark years (1820, 1850, 1880, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970). The 48 main countries were chosen on the basis of population size, GDP, regional representativity and data availability. Throughout the 1800-2025 period, the 48 main countries cover about 85-90% of the world population and GDP, while the 9 residual regions cover 10-15%. See Nievas and Piketty (2025).	

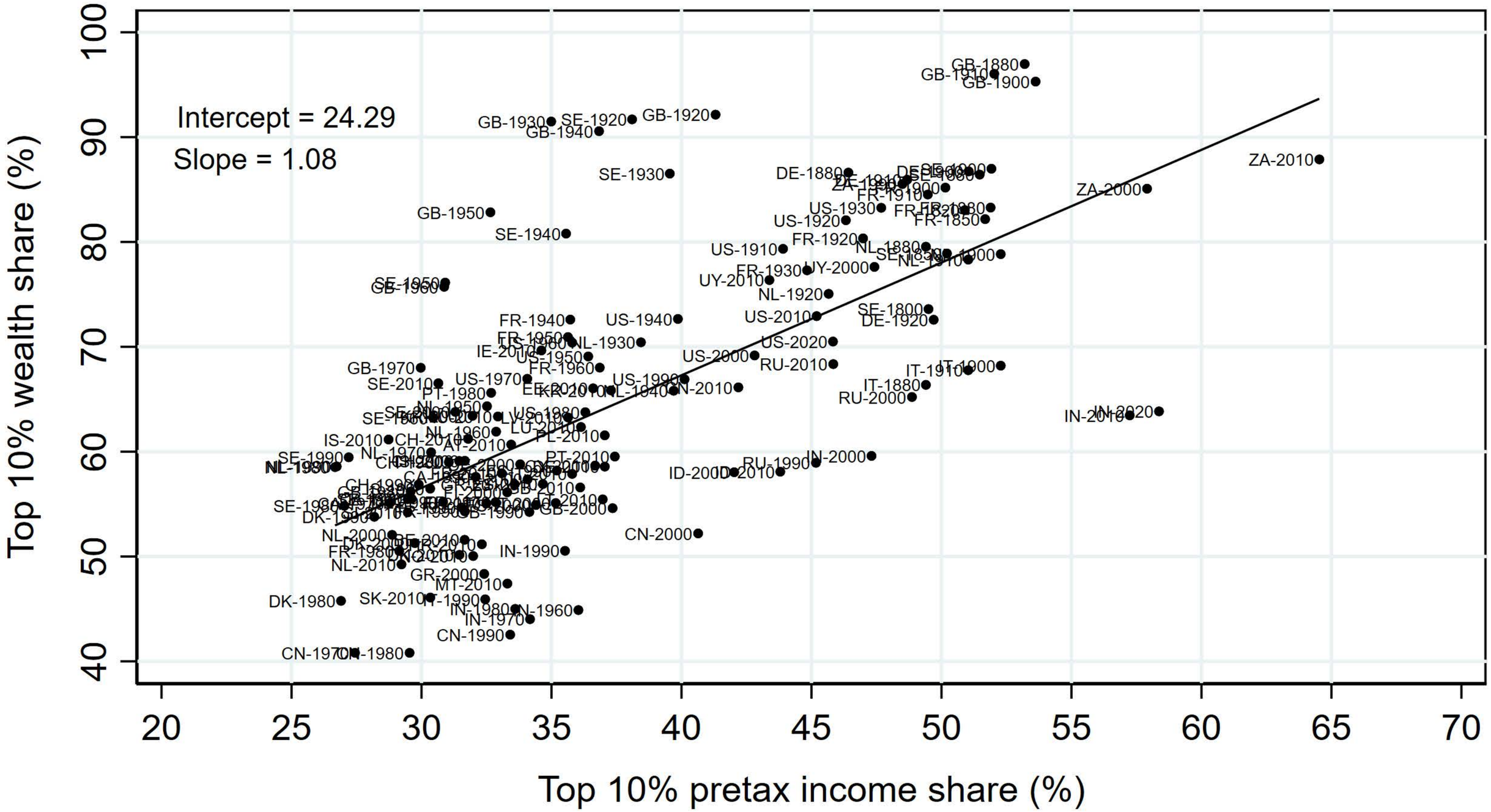
Table 3. WID Distributional Series: Additional Coverage

Pretax national income (variables sptinc, aptinc, tptinc) (equal-split, per capita & per adult)	AR 1932-1961; AU 1911-1979; BG 1924-1945; CA 1921-1979; CH 1933, 1934, 1936, 1939, 1941, 1943, 1945, 1947, 1949, 1951, 1953, 1955, 1957, 1959, 1961, 1963, 1965, 1967, 1969, 1971, 1973, 1975, 1977, 1979; CL 1964-1979; CM 1945; CN 1978-1979; DE 1871-1938, 1940, 1949, 1954, 1957, 1961, 1965, 1968, 1971, 1974, 1977; DK 1903,1908,1910,1915, 1917-1979; DZ 1932-1941, 1946-1955, 1957; ES 1933-1935, 1941-1955, 1957-1961, 1971; FI 1920-1979; FR 1915-1979; GB 1908-1979; GH 1943, 1951-1959; GR 1967-1979; HR 1967,1968, 1973,1974,1977-1979; HU 1914,1915,1927,1930-1940, 1951, 1955-1966, 1968, 1970, 1972, 1974, 1976, 1978; ID 1921-1939; IE 1922-1927, 1929-1953, 1964-1973, 1975-1979; IN 1922-1949, 1953-1968, 1971-1979; IL 1979; IN 1951-1979; IT 1974-1979; JP 1886-1979; KE 1936,1943,1948-1969; KR 1933-1979; MU 1933-1979; MW 1938, 1945, 1953-1958,1964-1979; MY 1948-1975; NG 1952-1959; NL 1914-1941, 1946, 1952, 1953, 1957-1959, 1962, 1964, 1966,1967,1973,1975,1977; NO 1875,1888,1892-1903, 1906, 1913, 1929,1938, 1948-1955, 1957-1979; NZ 1921-1979; PL 1924-1931, 1935,1936,1947,1955-1965, 1967,1970, 1972, 1976, 1978; PT 1936-1979; RU 1905-1976; SC 1955, 1961-1971; SE 1903, 1907, 1910-1912, 1916, 1919, 1934, 1935, 1941, 1943-1979; SG 1947-1954, 1956-1979; TN 1946-1948, 1952-1956; TW 1977-1979; TZ 1948-1970; UG 1948-1970; US 1913-1979; VN 1921, 1926-1935, 1937-1939, 1942; ZA 1913-1949, 1954-1965, 1967, 1969-1971, 1974, 1975, 1978, 1979; ZM 1929-1937, 1943-1958, 1963, 1968, 1970; ZW 1917-1939, 1945-1978	All 127 g-percentiles
Net household wealth (variables shweal, ahweal, thweal) (equal-split, per capita & per adult)	CA 1945-1953, 1961-1968, 1970; CN 1978-1979; DE 1895-1897, 1899, 1902, 1905, 1908, 1911, 1914, 1924, 1927, 1930, 1934, 1953, 1957, 1960, 1963, 1966, 1969, 1972, 1974, 1977; FR 1807, 1817, 1827, 1837, 1847, 1857, 1867, 1877, 1887, 1902-1905, 1907, 1909-1979; GB 1895-1914, 1919-1941, 1946-1979; IN 1961, 1971; IT 1891, 1894, 1901-1915; NL 1894-1938, 1947, 1951, 1953-1956, 1958-1961, 1963-1967, 1968-1971, 1973-1974, 1976-1979; SE 1800, 1850, 1873-1877, 1900, 1906-1908, 1920, 1930, 1935, 1937, 1945-1951, 1966, 1970, 1975, 1978; US 1913-1979	All 127 g-percentiles

In addition to the benchmark set of countries/territories/years described on Table 1, WID distributional series (equal-split, per capita & per adult) are also available for a number of countries-years described on this table. In addition, WID distributional series also cover other observation units (individualistic, tax units, etc.) and income concepts (factor income, fiscal income, etc.), but only for a relatively small and irregular subset of countries-years-gpercentiles (not listed here).

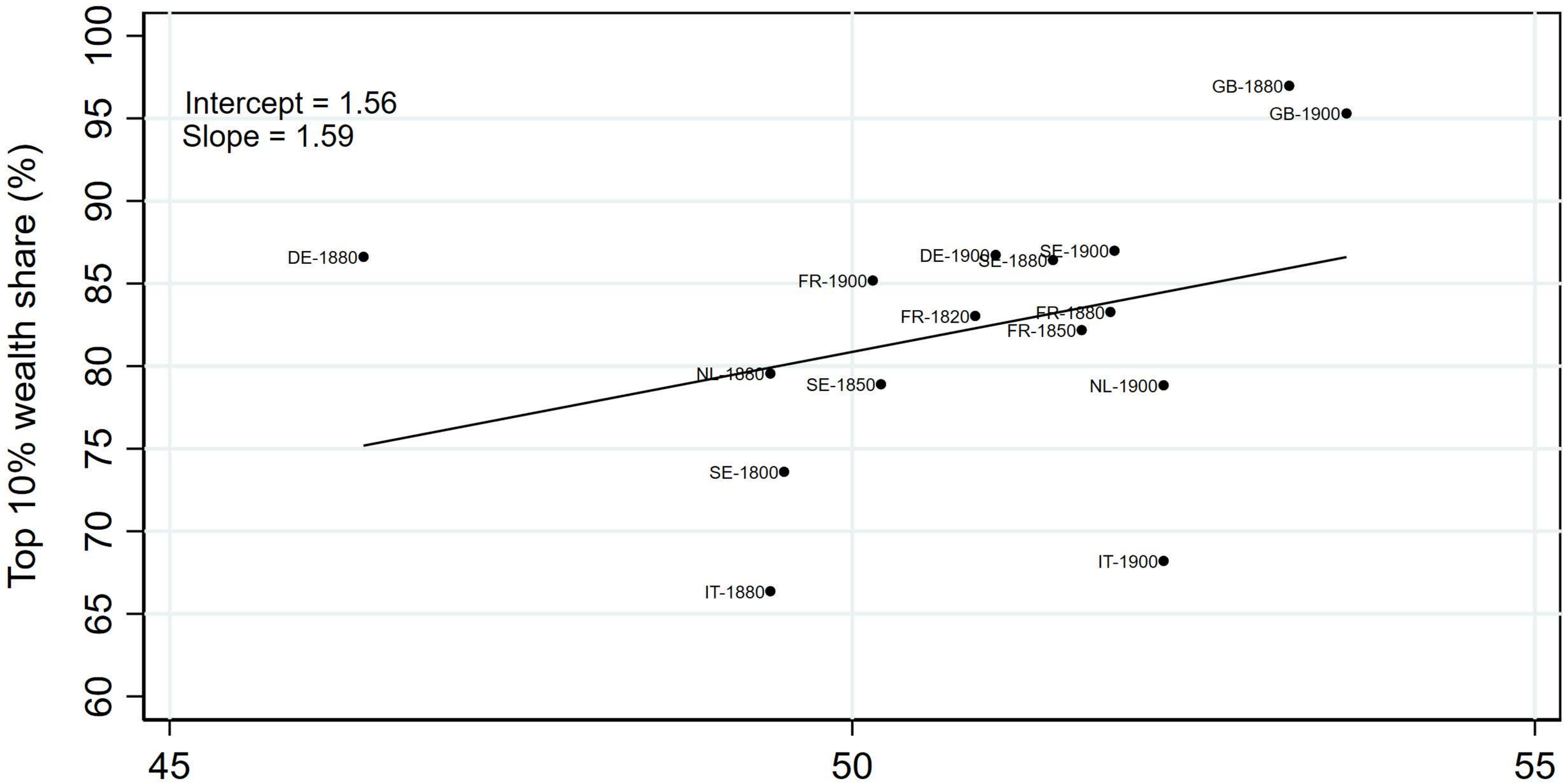
Table 4. Countries-Years with Wealth Distribution Estimates used in Imputations for Missing Observations	
Period 1800-1909	CA 1892, 1902; DE 1895-1897, 1899, 1902, 1905, 1908; FR 1807, 1817, 1827, 1837, 1847, 1857, 1867, 1877, 1887, 1902-1905, 1907, 1909; GB 1895-1909; IT 1891, 1894, 1901-1909; NL 1894-1909; SE 1800, 1850, 1873-1877, 1900, 1906-1908
Period 1910-1979	CA 1945-1953, 1961-1968, 1970; CN 1978-1979; DE 1911, 1914, 1924, 1927, 1930, 1934, 1953, 1957, 1960, 1963, 1966, 1969, 1972, 1974, 1977; FR 1910-1979; GB 1910-1914, 1919-1941, 1946-1979; IN 1961, 1971; IT 1910-1915; NL 1910-1938, 1947, 1951, 1953-1956, 1958-1961, 1963-1967, 1968-1971, 1973-1974, 1976-1979; SE 1920, 1930, 1935, 1937, 1945-1951, 1966, 1970, 1975, 1978; US 1913-1979
Period 1980-2024	AT 2010-2017; BE 2010-2017; CA 1984, 1999, 2012, 2016; CH 1981, 1991, 1997-2017; CN 1980-2015; CY 2010-2017; DE 1993, 1998, 2003, 2008, 2013, 2018; DK 1980-2012; EE 2013-2017; ES 1995-2015; FI 1995-2017; FR 1980-2014; GB 1980-2018; GR 2009-2018; HR 2017; HU 2014-2017; ID 2000-2014; IE 2013-2018; IN 1981, 1991, 2002-2022; IS 1997-2019; IT 1995, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016; KR 2000-2007, 2010-2013; LT 2018; LU 2010-2018; LV 2014-2017; MT 2010-2017; NL 1980, 1982, 1984, 1986, 1988-2019; NO 2010-2019; PL 2014-2016; PT 1980-1982, 2010-2017; RU 1995-2015; SE 1983, 1985, 1988, 1990, 1992, 1997, 1999-2012; SI 2010-2017; SK 2010-2017; US 1980-2022; UY 2009-2016; ZA 1993-2017
We describe in this table the set of countries-years with wealth distribution estimates which we use in imputations for missing observations (see Figures 1-4). Other countries-years reported on Tables 1-3 were estimated using the imputation equation.	

Period 1800-2023



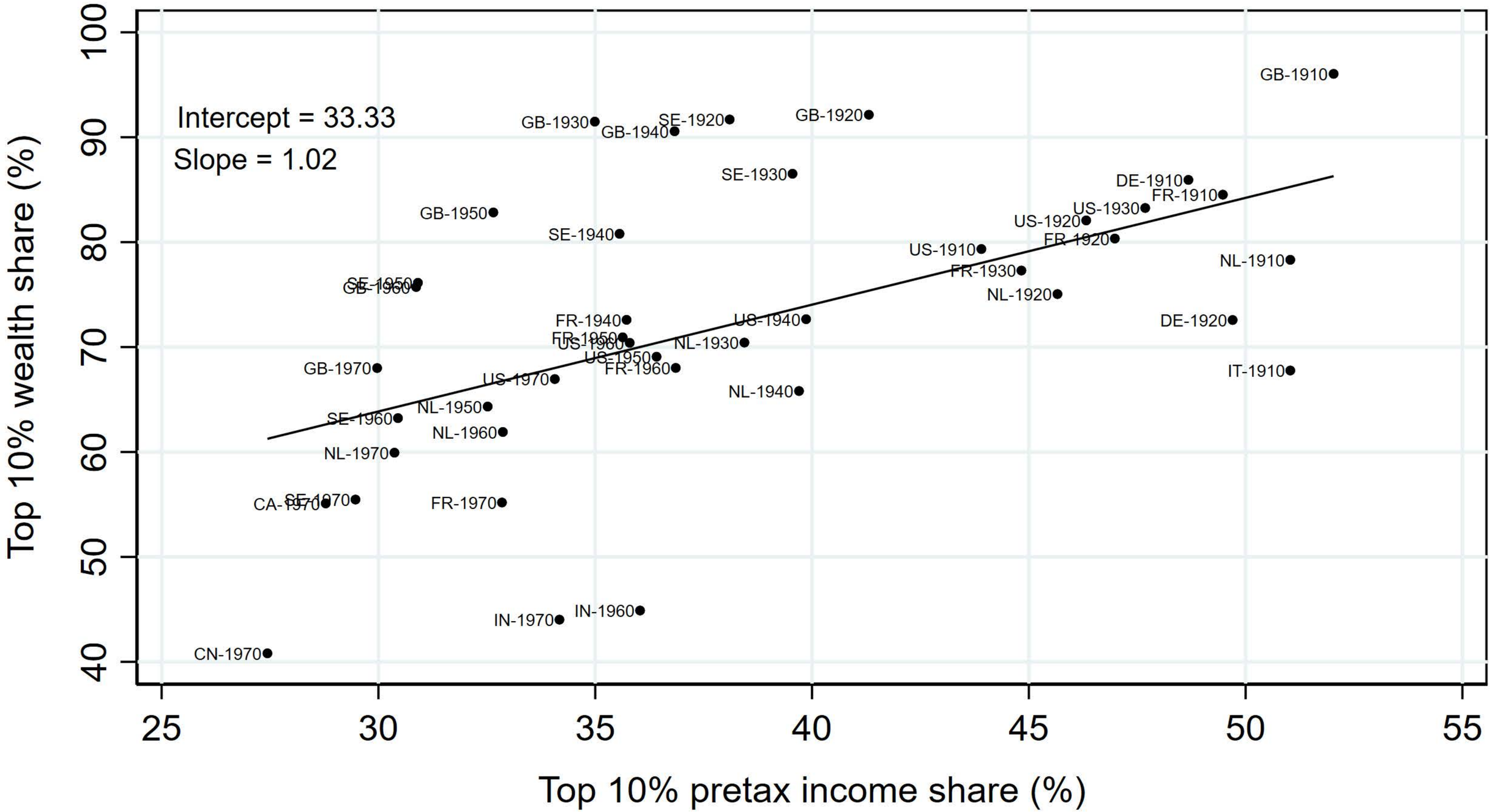
N.B.: The figure plots average top 10% income and wealth shares over subperiods 1800–1819, 1820–1849, 1850–1879, 1880–1899 and each decade from 1900–onwards for all countries for which we observe top income and wealth shares at least once during a given subperiod.

Period 1800-1909



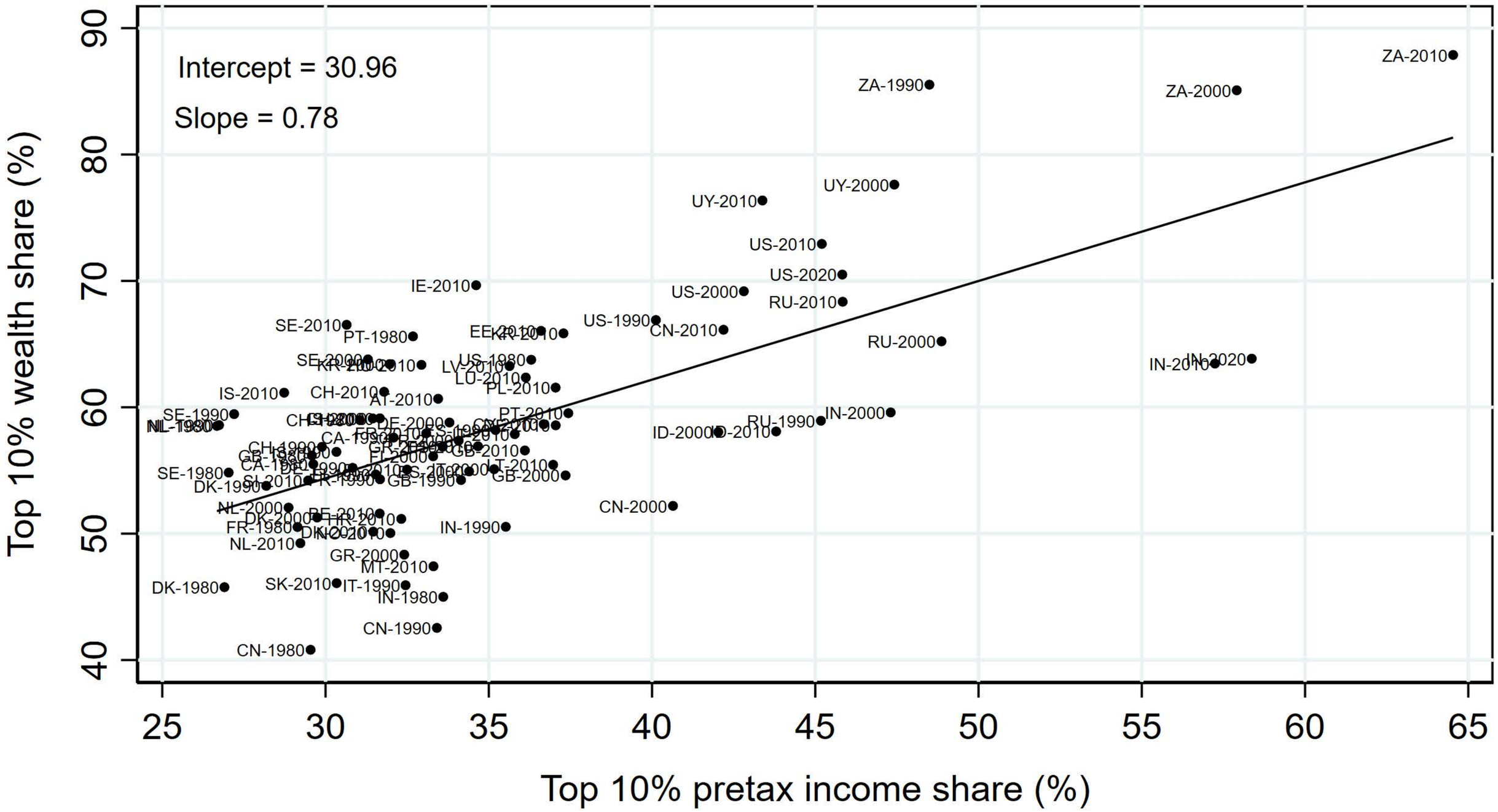
N.B.: The figure plots average top 10% income and wealth shares over subperiods 1800–1819, 1820–1849, 1850–1879, 1880–1899 and 1900–1909 for all countries for which we observe top income and wealth shares at least once during a given subperiod.

Period 1910-1979



N.B.: The figure plots average top 10% income and wealth shares over subperiods 1900–1909, 1910-1919, 1920-1929, 1930-1939, 1940-1949, 1950-1959, 1960-1969 and 1970–1979 for all countries for which we observe top income and wealth shares at least once during a given subperiod.

Period 1980-2023



N.B.: The figure plots average top 10% income and wealth shares over subperiods 1980–1989, 1990-1999, 2000-2009, 2010-2019 and 2020-2023 for all countries for which we observe top income and wealth shares at least once during a given subperiod.