# Codebook: War PIT dataset (v.0.2)

*Jakob Frizell, European University Institute, April 2021*

## Introduction

This dataset contains information on the combined top marginal personal income tax rate in 61 conflict-affected low-and middle-income countries between 1960 and 2020. The primary purpose of the data is to provide an indicator of tax redistribution, and particularly the tax burden placed on the rich, in these states over an extended time-period.

Personal income taxes have since their first appearance around two centuries ago (Seelkopf et al., 2019) remained an absolutely central fiscal policy instrument. Not only have they brought in crucial revenue to the state (Andersson, 2017), allowing it to expand its scope, but importantly, to consciously distribute the resultant financial burdens – indeed, to directly decrease economic inequality. Precisely by its potential effect on the top incomes in a society, progressive income taxes have been crucial in supressing and then containing inequality in the last century (Atkinson, Piketty, & Saez, 2011; Piketty, 2014). Moreover, given its fiscal and redistributive potential, the progressive income tax has been a central *political* issue since its conception. From the Communist manifesto – where it a "heavy progressive or graduated income tax" was included as no. 2 of political demands – to the Post-Cold War flat-tax movement, the personal income tax has carried a normative charge beyond any other fiscal instrument. In so being, its character in any specific society and period will also be a reflection of the predominating political and normative forces at play.

Compared to other, more direct measurements of tax progressivity – demanding original data that on many occasions simply do not exist for the concerned countries – the top PIT rate is overwhelmingly efficient in its simplicity, and has indeed been found to serve as an excellent proxy for the overall tax burden placed on the rich (Hope & Limberg, 2020). While PITs in low- and middle-income countries still lags behind those of the West in terms of revenue (Genschel & Seelkopf, 2016), their redistributive potential remains (Chu, Davoodi, & Gupta, 2000; Gemmell & Morrissey, 2005). Equally important, the highest PIT rate constitute a highly visible signalling device of the political elite – whether vis-à-vis domestic constituents or international capital markets.

## Sample coverage

The data includes all 61 low- and middle-income countries which have been actively involved in a major armed conflict[1] in any year between 1960 and 2020.

A country enters the sample in 1960 (or independence). With the last year being 2020, it results in a maximum temporal span of 61 years. In practice, it means that a majority of the countries in the sample are covered over their entire period of independence, or close thereto.

The broad geographical coverages, encompassing multiple countries from all continents, puts it apart from e.g. the much more Europe-focused sample of the otherwise similar Comparative Income Taxation Database (Genovese, Scheve, & Stasavage, 2016).

**Table E1: War PIT dataset country-period coverage**

| Country | From | To | Country | From | To |
|---|---|---|---|---|---|
| Afghanistan | 1960 | 2020 | Morocco | 1960 | 2020 |
| Algeria | 1962 | 2020 | Mozambique | 1975 | 2020 |
| Angola | 1975 | 2020 | Myanmar (Burma) | 1960 | 2020 |
| Azerbaijan | 1991 | 2020 | Nepal | 1960 | 2020 |
| Bangladesh | 1971 | 2020 | Nicaragua | 1960 | 2020 |
| Bosnia-Herzegovina | 1992 | 2020 | Nigeria | 1960 | 2020 |
| Burundi | 1962 | 2020 | Pakistan | 1960 | 2020 |
| Cambodia | 1960 | 2020 | Peru | 1960 | 2020 |
| Cameroon | 1960 | 2020 | Philippines | 1960 | 2020 |
| Chad | 1960 | 2020 | Portugal | 1960 | 2020 |
| Colombia | 1960 | 2020 | Russia | 1991 | 2020 |
| Congo (Brazzaville) | 1960 | 2020 | Rwanda | 1962 | 2020 |
| Congo, (DRC/Zaire) | 1960 | 2020 | Serbia (FYR) | 1990 | 2020 |
| Croatia | 1991 | 2020 | Sierra Leone | 1961 | 2020 |
| Cyprus | 1960 | 2020 | Somalia | 1960 | 2020 |
| Egypt | 1960 | 2020 | South Africa | 1960 | 2020 |
| El Salvador | 1960 | 2020 | South Sudan | 2011 | 2020 |
| Eritrea | 1991 | 2020 | Sri Lanka | 1960 | 2020 |
| Ethiopia | 1960 | 2020 | Sudan | 1960 | 2020 |
| Georgia | 1991 | 2020 | Syria | 1960 | 2020 |
| Honduras | 1960 | 2020 | Tajikistan | 1991 | 2020 |
| India | 1960 | 2020 | Tunisia | 1960 | 2020 |
| Indonesia | 1960 | 2020 | Turkey | 1960 | 2020 |
| Iran | 1960 | 2020 | Uganda | 1962 | 2020 |
| Iraq | 1960 | 2020 | Ukraine | 1991 | 2020 |
| Israel | 1960 | 2020 | Vietnam, (North/Unified) | 1960 | 2020 |
| Jordan | 1960 | 2020 | Vietnam, (South) | 1960 | 1975 |
| Laos | 1960 | 2020 | Yemen (AR/North/Unified) | 1960 | 2020 |
| Lebanon | 1960 | 2020 | Yemen, (PRY/South) | 1967 | 1990 |
| Liberia | 1960 | 2020 | Zimbabwe (Rhodesia) | 1960 | 2020 |
| Libya | 1960 | 2020 | | | |

[1] As defined by Uppsala Conflict Data Program, whereby the conflict whether inter-state or intra-state, but involving at least one state-actor, caused a minimum of 1000 battle-related deaths within a calendar year.

# Defining the top marginal PIT rate

The primary variable can be summarily defined as the combined marginal statutory tax rate applicable to the highest personal incomes. Below, the different components of this definition is explicated.

**Taxable person:** we are interested in personal rather than corporate income tax; accordingly, the relevant rate is that applied to *physical/natural persons*, rather than legal persons/corporations. If different rates apply, it further concerns *resident nationals*, rather than nationals living abroad or expatriate residents. Tax systems where spouses may file joint tax declarations, are simply treated as if they were based on individual income.

**Type of income:** the relevant rates applies to income derived from *economic activity*, i.e., employment or business activities. On the one hand it excludes income derived from ownership, such as property rent, interest payments, dividends, capital gains, or royalties from copy-rights. By including also business income, the operationalization differs from the relatively contemporary income tax data of e.g. Andrew Young School of Policy Studies (2010) and KPMG (2015), who focus narrowly on wage/salary income. This makes sense if you are interested in the possible marginal rates for the typical income earner, since this most often will be a wage earner. However, when interested in the marginal taxes on those who earn most in a society, the exclusion of income from business activity would be arbitrary. Indeed, with regard to redistribution – and the signalling of redistribution – the taxes levied on independent professionals and businessmen may be of even higher interest. In the last decades most countries have moved towards unifying their income tax schedules (if not making the fully "global"), but this has not always been the case. For most countries under study, different schedules have applied to e.g. wage income, "professional" activities and self-employed persons (sometimes also farmers, and more specific groups such as managers). Including also these schedules is therefore crucial. In this, the present coding aligns with that of the Comparative Income Taxation Database (Genovese et al., 2016).

**Top marginal rate:** the relevant rate applies to *directly assessed income*, as opposed to estimates on the basis of e.g. wealth, property value, or business turnover. This also excludes poll-taxes or any levy set at nominal values; in other words, it must be expressed as a *proportion* of income. Social Security Contributions are not included. The relevant *marginal* rate is further the highest possible proportion of each additional unit of income that a person (as defined above) may be liable to pay in income taxes. This also means that whenever there are several applicable schedules, as per the definition above, the highest top rate in any of them is recorded. It is further the statutory, rather than effective, marginal rate that is of concern. Consequently, deductions and tax credits are ignored.

**Combined rate:** in many instances, there are not only several alternative schedules, but a single income will often be subject to more than one schedule at the same time. Accordingly, what is recorded is the resulting top marginal rate of all applicable income tax schedules combined. This means, for example, that local income taxes are included whenever possible. The rate is then calculated according to the schedule of the largest city in the country. More often, it means the inclusion of one or more income surtax levied at the national level.

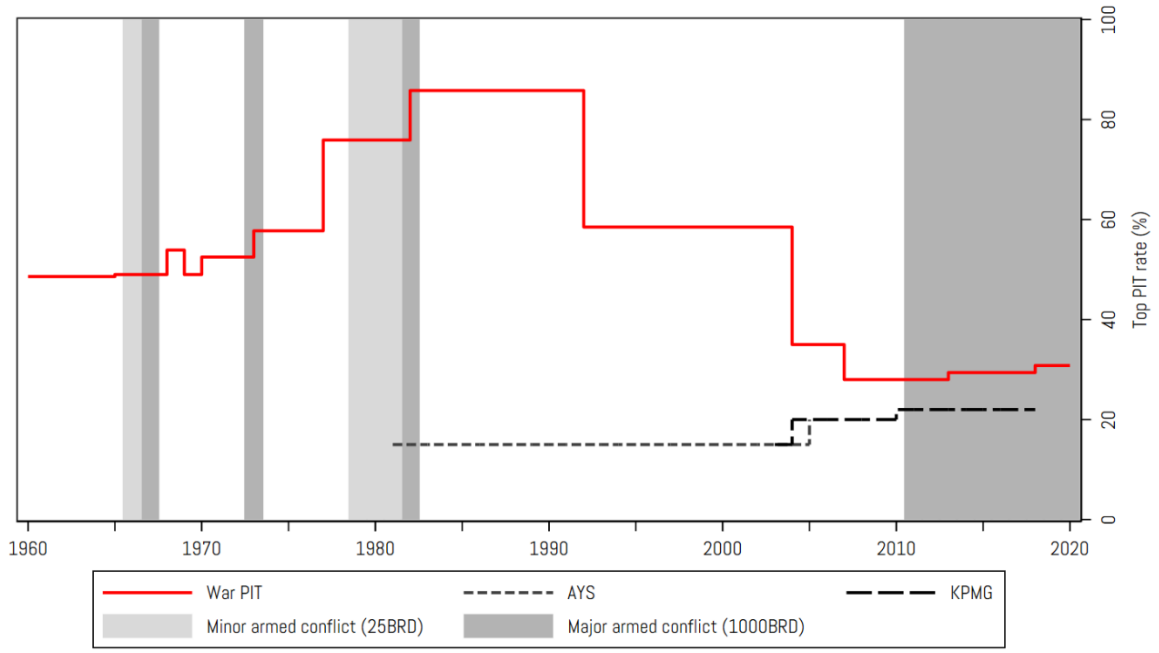
Important as these are, they often lead to some complexity, and occasionally uncertainty, in how to calculate the rate structure. This is because their relationship to the basic schedule(s) are not always clearly stated in secondary literature (and indeed opaquely expressed in legal documents). Suffice it here to state that there are three basic types of income surtaxes: (1) *additive*, applied to the same taxable income as the primary schedule, (2) *derivative*, applied to the tax liability of the primary schedule, and (3) *residual*, applied to the net ("residual") income, after deducing the primary tax liability.[2]

Inevitably, however, the precise coding will in many instances be a matter or sound judgement. In cases where the above criteria are not sufficient to arrive at an unequivocal figure, the priority is to achieve coherence, and thus comparability *within countries, over time.* This has two motivations. First, with regard to the general purpose of the dataset, is intended to be used to analyse how factors or events in a country affects the evolution of its income tax. While between-country comparisons may still be helpful, the causal questions under analysis – as is generally the case – fundamentally concerns inter-temporal effects. Second, and concerning the validity of the measurement, a rate change within a single country (rather than a difference between countries) will more accurately capture the relative redistributive effect of the tax. Indeed, taken by itself, the top rate does not necessarily reflect the absolute redistributive effect of the income tax. To make at least a fair theoretical estimate, one would also need information on e.g., the number of brackets, their relative progressivity, their cut-off points in terms of income, the size and nature of deductions, and, indeed the underlying distribution of incomes in the population. To that, we have to add essentially unmeasurable factors such as prevalence and distribution of tax evasion. All of these factors will vary considerably between different countries. On the other hand, fortunately, they show much less variation within the same country, over time. Indeed, the basic components of an income tax system is only rarely changed. Underlying factors such as degree of tax evasion is likewise likely to change only slowly. Hence, while a difference in top rate between two countries may not accurately reflect the relative redistributive effect of these countries' income tax systems, a change in the top rate within the same country, is very likely to reflect a change in the redistributive effect of that country's income tax system.

As an illustration of how the data produced by the project at hand compare to other available sources for non-OECD countries, Figure E1 below plots the evolution of top income taxes for Syria from 1960 to 2020. Not only does the collected data extend further in time, but the values differ completely due to the specific operationalisations applied. Notably, whereas the War PIT data codes the combined highest top rate for personal income tax – similar to CITD – AYS (2010) and KPMG (2015) codes only national-level, standard top rates (excl. surcharges) for the sole category of wage/salary income.

[2] Calculations are made according to the following formulas, where $a$ is primary tax rate, $b$ is the surtax rate, and $c$ is the final combined tax rate, all expressed as fractions:
Additive: $a+b=c$
Derivative: $a+a*b=c$
Residual: $a+b*(1-a)=c$

**Figure E1: Top PIT rates and conflict episodes in Syria (War PIT, AYS & KPMG)**




*Source: War PIT dataset (v.0.2); Andrew Young School of Policy Studies (2010); KPMG (2015)*

## Sources

The collection of data has to a significant extent relied on documentation from the IMF Archives. Foremost of these are the reoccurring, country-specific "Recent Economic Developments"-reports. For a majority of countries in the sample, these reports have been issued annually or bi-annually, albeit with considerable gaps, from the early 1960's (then often titled "Background notes") to the late 1990's. Secondly, most of the relevant "Staff Reports", including Statistical Appendices, as well as mission reports from the Fiscal Affairs Department and so-called "Selected Issues"-reports have also been consulted.

These IMF documents are particularly valuable, for two reasons: first, rather than simply reporting and analysing legal revisions, they are primarily concerned with the practical implementation of various taxes. This is important since enactment of tax laws does not necessarily imply its immediate or complete implementation. Second, the fact that they are published in fairly regular intervals, often makes it possible to follow a tax reform at several points in time, thus allowing for a more accurate evaluation of its effective implementation. On the other hand, it is obvious that the interest payed to tax issues have varied considerably between different regional departments as well as over time. Hence, for some countries, tax issues will only be mentioned in relation to substantial reforms, making it difficult to verify periods of stability. More importantly, it is only from the establishment of the Fiscal Affairs Department in 1964, that high-quality analyses (as opposed to brief references to top-rates) start appearing in the country reports. Furthermore, after around 2010, IMF essentially stop publishing information on actual taxes in their reports (as opposed to presenting elaborate formal models of fiscal sustainability etc.). Hence, in addition to IMF documentation, the following types of sources have also been used:

- **Tax briefs and expert publications.** For earlier periods (ca. 1960-2000), a variety of specialized tax publications, such as "Bulletin for international fiscal documentation" and "Tax News Service" (both IBFD), and "Income taxes outside the United Kingdom" (Board of Inland Revenue), held at the IBFD library in Amsterdam, have been consulted. For the last 15-20 years of the sample period, annual "tax guides" from major tax consultancy firms (such as PwC, KPMG and Deloitte) have been extensively used.
- **Legal documents** have, to the extent possible, been consulted in case of any uncertainty (secondary sources giving vague, conflicting, or no information).
- A variety of **second-hand sources** such as books, journal articles, and news articles have been used. These range from in-depth studies of tax systems (e.g. by International Bureau for Fiscal Documentation) to online news articles reporting on new tax legislation. These sources seldom provide more than snapshot pictures, but can, in the first instance, give a detailed overview of the tax system, and in the second give valuable information on the precise timing of rate change.
- Inquiries with **legal/tax experts** have, finally, provided assistance with interpretation for particularly complex country-periods.

- **Existing databases**. For the years 1981-2005, AYS' *World Tax Indicators*; for the years 2003-2018 KPMG's data on top corporate and income taxes. Note that the operationalizations for said databases are not identical to those of the present dataset. However, in those instances where they do indeed coincide, they have been used to triangulate other sources and fill in gaps in the time series.

For each country, all relevant data, including sources, have been compiled into "Background Notes" forming the basis of the coding of the annual dataset. The Background Notes are available upon request.

# Data reliability / Missing data

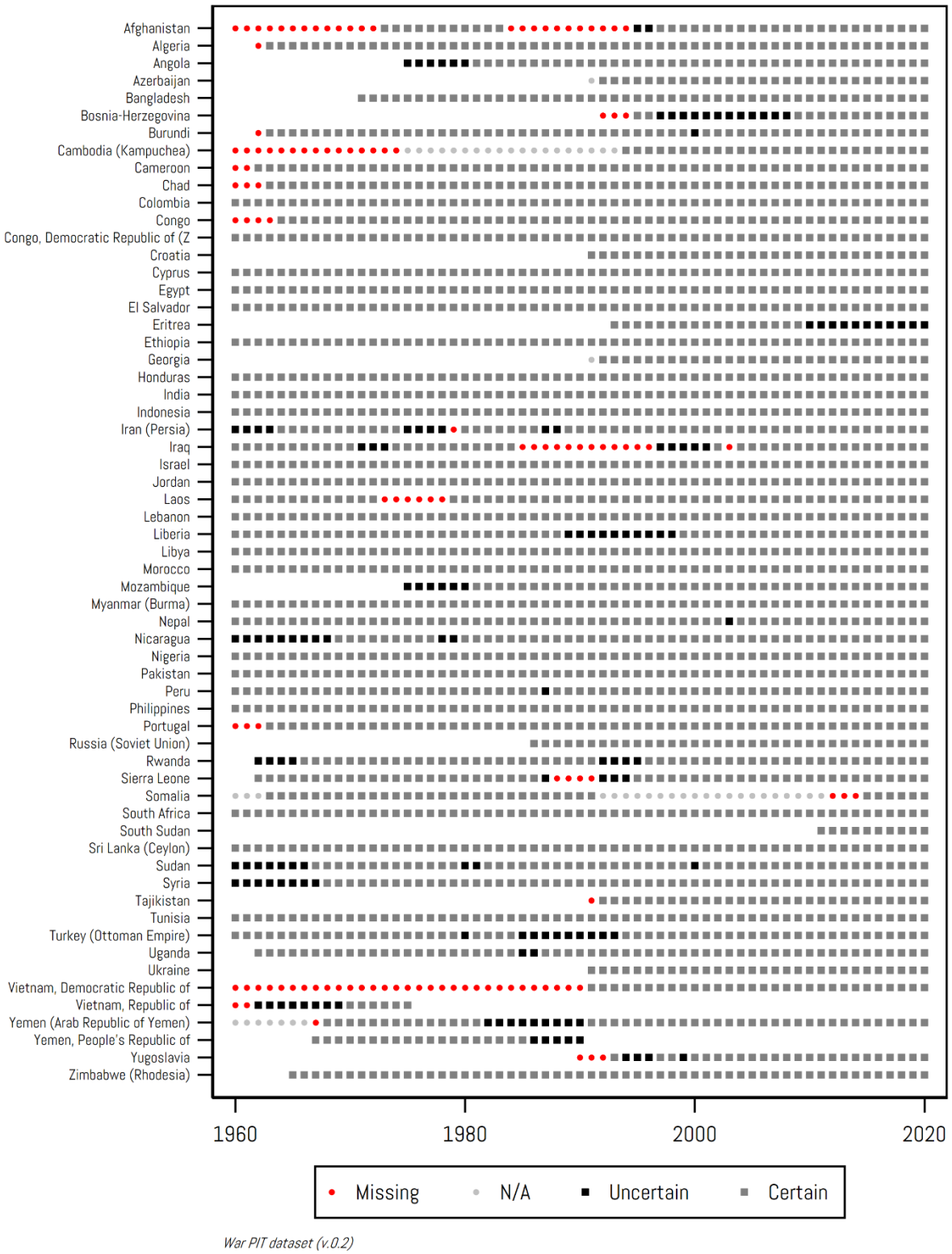
To establish reliable and unbroken time-series is of highest priority for the data-project at hand. While this is always to some extent the case, depending on the purpose of the data gathering, and against the backdrop of finite time and resources, different decisions will be reached as to how to deal with "problematic" observations. E.g., if the insecurity of the sources is too high, the observation may simply be left as "missing". In the case of taxes – which usually remain fairly stable of over time, only taking on step-wise changes – one may simply leave the value unchanged in t+1 despite indications of a temporary change. This could indeed be justified if such unrecorded change can be understood as unrelated "noise" for the question at hand. However, it is precisely those periods for which information may be inaccessible, unreliable or difficult to interpret – such as during major armed conflicts – that are of particular importance here. Hence, disproportionate effort has been spent on accurately coding the "difficult cases", including the precise effect of temporary surtaxes. With that said, whenever a particular observation is deemed less than fully reliable – most often because it has been imputed rather than directly confirmed – this is flagged as uncertain. Out of a total of 3,197 country-year observations, 121 observations are missing (3.8%); a further 146 are flagged as uncertain (4.6%).

**Table E2: Top rate data, missing/uncertain observations (War PIT dataset v.0.2)**

| | **Missing** | **Not applicable** | **Uncertain** | **Certain** | Total | (column %) |
|---|---|---|---|---|---|---|
| No Conflict | 71 | 39 | 111 | 2,457 | 2,678 | 83.77% |
| (row %) | 2.65% | 1.46% | 4.14% | 91.75% | | |
| Conflict-year | 50 | 12 | 35 | 422 | 519 | 16.23% |
| (row %) | 9.63% | 2.31% | 6.74% | 81.31% | | |
| Total (full period) | 121 | 51 | 146 | 2,879 | 3,197 | |
| (row %) | 3.78% | 1.60% | 4.57% | 90.05% | | (100%) |

*Note: Conflict-years are those reaching the threshold of 1,000 BRD; Not applicable denotes years where country is included in sample but where no PIT is effectively in place.*
*Source: War PIT dataset v.0.2; UCDP-ACD v.20.1 (2020)*

**Figure B2: Top PIT rate data, coverage and missing obs. (War PIT v.0.2)**




*War PIT dataset (v.0.2)*

# Variables

Main variable

**toprate** Continuous variable. For detailed definition, see below.

Secondary variables

**year** Calendar year.

**country** Country name.

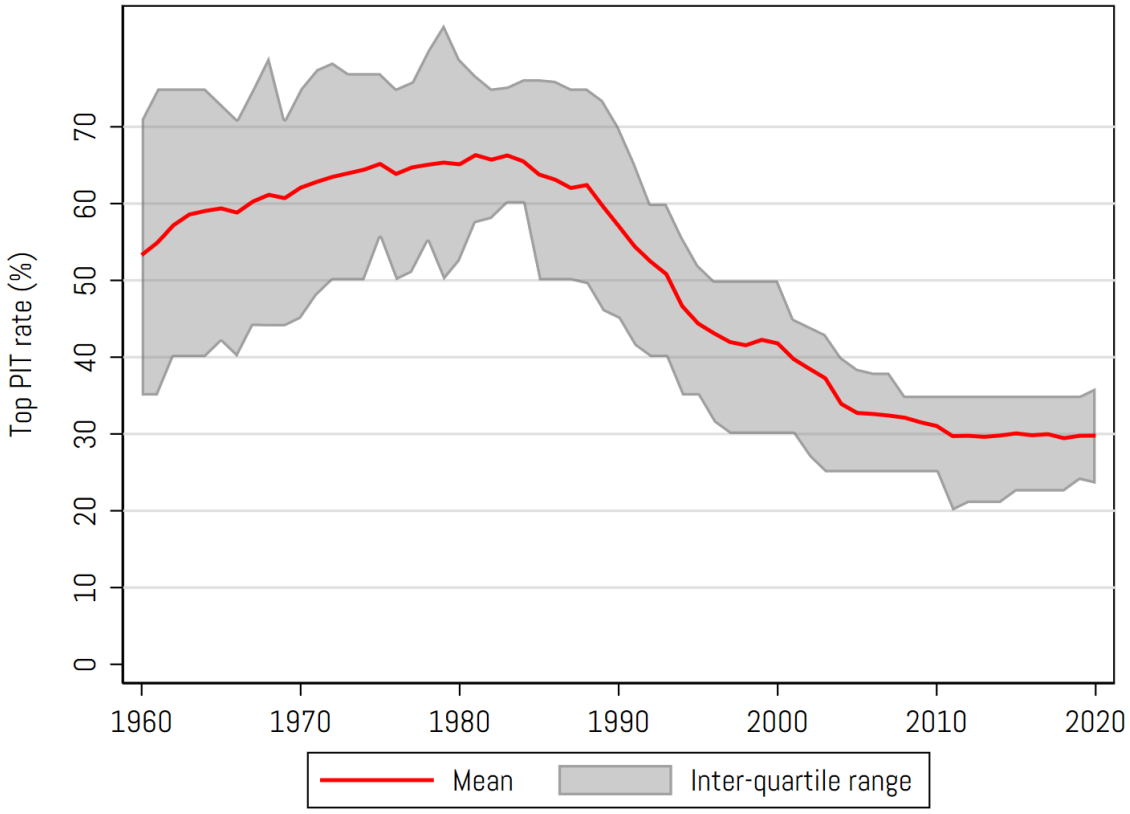
**ccode** Country-code according to Gledditch/Ward.[3]

**comment** Clarifying comments on reforms or legal sources.

**uncertain** Flag for uncertainty: 1=uncertain, 0=certain. [missing value = no PIT levied]

[3] One cases have been adjusted so as to account for substantial continuity:
Serbia 2006-2018 changed to 345 (successor to Yugoslavia). Otherwise ambiguous cases: unified Yemen keeps ccode (678) of North Yemen; unified Vietnam keeps ccode (816) of North Vietnam.

**Figure E3: Top PIT rates over time (War PIT v.0.2)**



## References


Andersson, P. (2017). *Essays on the Politics of Taxation.* Lund University,

Andrew Young School of Policy Studies. (2010). *Andrew Young School World Tax Indicators*. Retrieved from: https://aysps.wufoo.com/forms/s7x0x5/

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature, 49*(1), 3-71

Chu, K.-y. C., Davoodi, H., & Gupta, S. (2000). *Income Distribution and Tax, and Government Social Spending Policies in Developing Countries*. UNU-WIDER Working Papers. UNU-WIDER. Helsinki.

Gemmell, N., & Morrissey, O. (2005). Distribution and Poverty Impacts of Tax Structure Reform in Developing Countries: How Little We Know. *Development Policy Review, 23*(2), 131-144.

Genovese, F., Scheve, K., & Stasavage, D. (2016). *Comparative Income Taxation Database*.

Genschel, P., & Seelkopf, L. (2016). Did they learn to tax? Taxation trends outside the OECD. *Review of International Political Economy, 23*(2), 316-344. doi:10.1080/09692290.2016.1174723

Hope, D., & Limberg, J. (2020). *The Knowledge Economy and Taxes on the Rich*. Working paper prepared for special issue of the Journal of European Public Policy (JEPP) on 'The Politics of Taxing the Rich: Declining Tax Rates in Times of Rising Inequality'. Working Paper.

KPMG. (2015). *2015 Global Tax Rate Survey*. Retrieved from www.kpmg.com/taxrates

Piketty, T. (2014). *Capital in the twenty-first century : the dynamics of inequality, wealth, and growth*. Cambridge MA: The Belknap Press of Harvard University Press.

Seelkopf, L., Bubek, M., Eihmanis, E., Ganderson, J., Limberg, J., Mnaili, Y., . . . Genschel, P. (2019). The rise of modern taxation: A new comprehensive dataset of tax introductions worldwide. *The Review of International Organizations*. doi:10.1007/s11558-019-09359-9