NBER WORKING PAPER SERIES

CAPITAL TAXATION, DEVELOPMENT, AND GLOBALIZATION: EVIDENCE FROM A MACRO-HISTORICAL DATABASE

Pierre Bachas
Matthew H. Fisher-Post
Anders Jensen
Gabriel Zucman

Working Paper 29819
http://www.nber.org/papers/w29819

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2022, Revised May 2024

This paper was previously circulated under the title "Globalization and Factor Income Taxation." The database built in this paper is available online at http://globaltaxation.world. Additional material is available in the online supplementary appendix: link here. We acknowledge financial support from the Weatherhead Center for International Affairs at Harvard University and the World Bank Development Economics' Research Support Budget. Zucman acknowledges support from the Stone Foundation. We are indebted to Elie Gerschel, Xabier Moriana, Rafael Proença, Roxanne Rahnama and Anton Reinicke for excellent research assistance. We thank numerous seminar participants and in particular Pierre Boyer, Damien Capelle, Denis Cogneau, Martin Fiszbein, Jason Furman, Simon Galle, Pinelopi Goldberg, Gordon Hanson, Amit Khandelwal, David Lagakos, Etienne Lehmann, Juliana Londoño-Vélez, Benjamin Marx, Sergey Nigai, Mathieu Parenti, Steven Pennings, Nam Pham, Thomas Piketty, Tristan Reed, Ariell Reshef, Bob R kers, Dani Rodrik, Nicholas Ryan, Emmanuel Saez, Nora Strecker, Romain Wacziarg, Daniel Yi Xu and Roman Zarate for insightful comments and discussions. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not represent the views of the World Bank and its affiliated organizations, nor those of the Executive Directors of the World Bank, nor of the governments they represent, nor of the National Bureau of Economic Research.

Capital Taxation, Development, and Globalization: Evidence from a Macro-Historical Database
Pierre Bachas, Matthew H. Fisher-Post, Anders Jensen, and Gabriel Zucman
NBER Working Paper No. 29819
March 2022, Revised May 2024
JEL No. F14,F62,H20,O24

ABSTRACT

This paper builds and analyzes a new global macro-historical database of effective tax rates on capital and labor in 154 countries. We establish a new stylized fact: while effective capital tax rates fell in developed countries between 1965 and 2018, they rose in developing countries since 1990. Multiple research designs at the country, sector and firm-level suggest that trade openness contributed to this rise, by increasing the share of output produced in corporations and larger firms, where effective capital taxation is higher. In contrast to a common view, globalization appears in many countries to have supported governments' ability to tax capital.

Pierre Bachas
The World Bank
1818 H Street N.W.
Washington, DC 20433
pbachas@worldbank.org

Matthew H. Fisher-Post
48 Boulevard Jourdan
PARIS 75014
France
mfp@psemail.eu

Anders Jensen
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
and NBER
anders_jensen@hks.harvard.edu

Gabriel Zucman
Department of Economics
University of California, Berkeley
530 Evans Hall, #3880
Berkeley, CA 94720
and NBER
zucman@berkeley.edu

Appendices are available at http://www.nber.org/data-appendix/w29819

1 Introduction

How has globalization affected the relative taxation of capital and labor? Has it uniformly eroded the amount of taxes paid by capital owners, shifting the burden to workers? Or have some countries managed to increase effective capital tax rates, and if so how? Answering these questions is critical to better understand the macroeconomic effects and social sustainability of globalization in uncertain times (Goldberg & Reed, 2023).

Based on a new long-run global database of effective tax rates on capital and labor, we document that in developing countries, effective capital tax rates have increased in the post-1990 era of hyper-globalization. Consistently across several research designs, we find that a significant share of this rise can be explained by trade openness. By expanding the share of economic activity occurring in the corporate sector, and within the corporate sector in larger firms, our results show that trade improves the effective collection of taxes, particularly corporate income taxes. Globalization has also had widely noted negative effects on capital taxation, due to international tax competition that applies downward pressure on corporate statutory tax rates. We find that the positive tax capacity effect of trade we uncover prevailed in developing economies, causing openness to increase overall government tax revenues (as a % of GDP). The revenue consequences of globalization have not been systematically investigated in developing countries due to limited data, and concerns over potential revenue losses have persisted as a key obstacle to further integration across borders (World Bank, 2020). In contrast to a common view, our findings show that globalization has not uniformly eroded governments' ability to raise revenue, and instead appears to have supported capital taxation in many countries.

To establish these results, this paper makes two contributions. The first is to build and analyze a macro-historical database of effective tax rates on capital (ETR_K) and labor (ETR_L) covering 154 countries, with over half starting in 1965, until 2018. Each ETR divides all taxes collected on the factor by the national income that accrues to it; by relying on actual taxes collected, ETRs capture the net past effect of all tax rules and, importantly for developing countries, tax evasion and avoidance. Complementary to existing ETR series that focus on developed countries, our data provides a global coverage by digitizing and harmonizing thousands of historical and recent public finance records in developing countries. The global database allows us to systematically characterize the evolution of effective tax rates in developing countries and compare trends across development levels.

A novel fact emerges from this database: the evolution of capital taxation has been asymmetric across development levels. In high-income countries, effective capital tax rates declined, from a high of 38-39% in the late 1960s to 32-33% in the late 2010s. By contrast, in developing countries, effective capital tax rates have been on a rising trend since the beginning of the 1990s, albeit starting from a low level. Effective capital tax rates rose from 10% in 1989 to 18% in 2018, with more pronounced increases in larger economies. For example, ETR_K rose from 6% to 24% in China, 5% to 12% in India, and 7% to 27% in Brazil. The positive trend in capital taxation is driven by the corporate sector: the average effective corporate tax rate rose from 12% in 1989 to 20% in 2018.

This rise of capital taxation in low- and middle-income countries had not been noted in the literature before, due to a lack of data on the evolution of taxation globally. The finding appears robust. It holds: when we exclude China and oil-rich countries; with other approaches to computing capital and labor income in unincorporated businesses (where factor shares are not directly observable); and with alternative ways of splitting personal income tax revenue between capital versus labor.

Our second contribution is to formulate and test a hypothesis that sheds light on the rise of capital taxation in developing countries. We hypothesize that openness exerts a positive effect on developing countries' capacity to tax, consistent with trade leading to the expansion of larger firms relative to smaller ones (Mrázová & Neary, 2018) and firm-level effective taxation rising with size, due to better enforcement and higher statutory tax burdens (Almunia & Lopez-Rodriguez, 2018; Best, Shah, & Waseem, 2021).[1] Our hypothesis is motivated by the observation that the rise in ETR_K coincides with trade liberalization. Since the beginning of the 1990s, many developing countries opened their markets and reduced tariffs, leading to a boom in international trade that reshaped the economies of Mexico, India, and China among others (Goldberg, 2023). By disproportionately benefiting larger firms, trade can increase the share of economic activity in corporations and more formal businesses, where effective taxation of capital (and labor) is higher.

To motivate the tax capacity hypothesis, Figure 1 shows that the share of domestic output from the corporate sector (profits and employee compensation) has grown over time in developing countries, at the expense of mixed-income (income of self-employed and unincorporated businesses). While the corporate sector accounted for 53% of domestic output in 1989, prior to the hyper-globalization era, it grew to 62% by 2018; mixed income fell from 32% to 20% over the same period. Thus, developing countries have experienced a relocation of activity from a hard-to-tax sector to a sector with stronger effective taxation.

[1] Higher effective taxation in the corporate sector stems both from stronger enforcement and higher statutory taxes than in the non-corporate sector. Our notion of tax capacity is that these co-determined forces jointly lead to higher ETR_K with firm size (where size is measured as firm output, in our case revenue).

We establish the second contribution in two steps. First, we study the impact of trade on taxation in developing countries, with a focus on ETR_K and corporate taxes. Second, we study mechanisms that link trade to taxation, with a focus on the tax capacity channel.

We implement three research designs to study how trade impacts taxation. First, we estimate the non-parametric association within a country over time between ETR and trade openness. Second, we analyze major trade liberalization events that occurred in seven large developing countries, including China's WTO accession in 2001, and caused sharp reductions in trade barriers (Brandt, Biesebroeck, Wang, & Zhang, 2017; Goldberg & Pavcnik, 2016). We use synthetic control methods and present event-study results. Third, we extend the trade instruments from Egger, Nigai, and Strecker (2019) to our sample.

All three designs show that, in developing countries, trade leads to a large increase in ETR_K, and a smaller increase in ETR_L. The effect is sizable: trade openness can account for 33% of the documented rise in ETR_K since 1989. Although studying macroeconomic outcomes presents identification challenges, the results are consistent across research designs, which differ in their identifying assumptions, and are robust to numerous sensitivity checks. Across the research designs, we also find that trade leads to an increase in total tax revenues (as a % of GDP). Reflecting trade's positive impact on ETR_K, over half of this increase comes from higher corporate income taxes (CIT), and a smaller share from personal income taxes and payroll. Indirect taxes (combining tariff revenues and domestic consumption taxes) slightly rise, but the coefficient is not significant.

We then turn to investigate mechanisms. In the IV and liberalization event-studies, we find that trade increases the share of domestic output produced in the corporate sector, relative to the unincorporated business sector (mixed-income).[2] Thus, output is expanded in the corporate sector where enforcement is stronger and effective taxation is higher (Slemrod & Velayudhan, 2018). Moreover, within the corporate sector we find that trade increases the average effective tax rate on capital, suggesting the expanded corporate output accrues to firms whose ETR_K increases with their output (our proxy for firm size). These two effects of trade are consistent with the tax capacity channel. Simultaneously, we find that trade reduces the statutory corporate tax rate, consistent with a tax competition channel where globalization pushes governments to reduce the statutory tax burden on capital. On net, the positive tax capacity impact outweighs the tax rate reduction in developing countries, causing trade to increase ETR_K at the country-level.

In contrast, we find no tax capacity effect of trade in developed countries, but a stronger decrease in statutory corporate tax rates. These results help reconcile the asymmetric evolution of capital taxation in developing and developed countries.

[2] Trade leads to a sharp rise in corporate profits and an insignificant change in employee compensation.

We sharpen our mechanism analysis by conducting a firm-level investigation of the tax capacity channel. We merge multiple administrative datasets in Rwanda, which allows us to observe each firm's integration into international trade and corporate tax payments. The integration measure accounts for the firm's indirect exposure to trade through its production network (Almunia, Hjort, Knebelmann, & Tian, 2023). Rwanda provides an interesting setting: starting from a relatively low share of domestic output, the corporate sector has grown significantly since the 1990s, in tandem with a rise in trade openness and tax collection. Using the shift-share design of Hummels, Jørgensen, Munch, and Xiang (2014) for identifying variation, we find that trade integration increases both a firm's ETR_K and its size. Though limited to a single country, these firm-level results provide micro-evidence for trade's positive impact on ETR_K, and support the tax capacity mechanism whereby trade's impact is mediated by a positive firm size-ETR_K gradient.

Finally, we study sources of heterogeneity in the pro-tax impact of trade. During our sample period, developing countries have invested in domestic tax enforcement, such as large taxpayer units (C. Basri, Felix, Hanna, & Olken, 2019). We find that trade's impacts on the tax capacity mechanism and on ETR_K hold in the absence of these enforcement policies and, more generally, outside of periods of significant fiscal pressure (Cagé & Gadenne, 2018). Thus, trade's pro-tax impact appears to be a broad feature of the globalization process which does not hinge on governments' initial enforcement and revenue needs. At the same time, we find that openness' pro-tax impact depends on the nature of the trade shock, in ways that are consistent with recent theoretical work on trade and formalization (Dix-Carneiro, Goldberg, Meghir, & Ulyssea, 2021).

Combining multiple empirical strategies, our results at the country, corporate sector, and firm-level consistently suggest that trade openness increases ETR_K and contributed to the newly documented rise of ETR_K in developing countries since the early 1990s. Based on a new global database, our findings show that globalization has supported effective capital taxation and overall revenue collection in many countries around the world.

Section 2 discusses related literature. Section 3 describes the methodology and data. Section 4 presents findings on the long-run evolution of ETR. Section 5 analyzes trade's impact on ETR and Section 6 investigates the mechanisms. Section 7 concludes.

2 Related Literature

Globalization and tax structure Our paper contributes to the macro literature on globalization and tax structure (Alesina & Wacziarg, 1998), reviewed in Adam, Kammas, and Rodriguez (2013). The "race to the bottom" hypothesis posits that governments reduce taxes on factors that become more mobile (e.g., capital) following trade liberalization (Slemrod, 2004). To achieve revenue neutrality, governments raise taxes on less mobile factors (e.g., labor).[3] The "social insurance" hypothesis postulates that governments raise revenue to insure workers displaced by international competition, often via social security and payroll taxes (Rodrik, 1998). These studies mainly focused on high-income countries. By expanding the scope to developing countries, we formulate and test a new mechanism, where trade increases ETR by expanding activity in firms with higher effective tax collection. Our results suggest that globalization has supported the ability of governments to tax capital in many countries.

Our results are based on a new global database of effective tax rates, which complements existing datasets (including Carey & Rabesona, 2004; Kostarakos & Varthalitis, 2020; McDaniel, 2007) by expanding coverage to developing countries (details in Section 3).[4] Our backward-looking ETR measure is complementary to the literature on forward-looking capital tax rates (including Devereux & Griffith, 1999), which models in detail the statutory tax burden a firm would face under different conditions. This literature finds that the statutory tax burden on capital has fallen in developed and developing countries, consistent with the 'race to bottom' mechanism (including Devereux, Griffith, & Klemm, 2002; R. Kumar & James, 2022; Steinmüller, Thunecke, & Wamser, 2019).

Effective taxation and trade in developing countries Our paper contributes to the micro-oriented literature on trade and public finance in developing countries. Many studies focus on *border taxes* and evasion (e.g., Fisman & Wei, 2004; Javorcik & Narciso, 2017; Sequeira, 2016) or cross-border income-shifting (e.g., Bilicka, 2019; Londoño-Vélez & Tortarolo, 2022; Wier, 2020). We focus instead on trade's impacts on the *domestic tax bases* of capital and labor and domestic economic structure.[5] Our results are intuitive when considering that the trade literature finds positive effects of openness on domestic outcomes including market shares (McCaig & Pavcnik, 2018), firm size (Alfaro-Ureña, Manelici, & Vasquez, 2022), and local development (Méndez & Van Patten, 2022), which the public finance literature has separately identified as determinants of effective taxation (Besley & Persson,

[3]Within labor in OECD countries, Egger et al. (2019) find that globalization in the post-1994 era led to a reduction in income taxes for the top 1% of workers and increased income taxes for middle-class workers.

[4]We complement other work in economic history on taxation (including Cogneau, Dupraz, Knebelmann, & Mesplé-Somps, 2021), by providing long-run measures of factor effective tax rates.

[5]The theoretical literature has focused on trade's impact on the optimal indirect tax mix between border and consumption taxes in developing countries (e.g. Emran and Stiglitz, 2005) and mainly abstracted from direct taxes. Benzarti and Tazhitdinova (2021) study the impact of indirect taxes on trade flows.

2014; Best et al., 2021).[6] We contribute by linking these two bodies of work and directly studying trade's impacts on domestic tax bases at the country, sector and firm level.

By incorporating domestic tax bases, we can comprehensively study the total tax revenue impacts of globalization. Previous studies on trade's revenue impact in developing countries have produced mixed findings, possibly due to differences in sample, methods and tax base focus (including Baunsgaard & Keen, 2009; Buettner & Madzharova, 2018; Cagé & Gadenne, 2018). We contribute by implementing multiple identification strategies in the largest sample to date and find that trade's impacts on domestic tax bases are sufficiently large that openness increases total tax revenue (as a % of GDP).

These impacts of trade are mediated by the tax capacity mechanism, which is rooted in two distinct insights from the trade and the public finance literatures. First, a large class of models predicts that trade leads to the expansion of large firms relative to small firms (Mrázová & Neary, 2018); for empirical evidence, see Bernard, Jensen, Redding, and Schott (2007). Second, in developing countries small firms are mainly informal, and effective taxation increases with firm size (measured as firm revenue)[7]; this positive gradient arises because effective tax collection is higher in larger firms and corporations due to their visibility, complex production structures, and employment of many workers (Almunia, Hjort, et al., 2023; Waseem, 2020). The resulting information trails improve enforcement (Naritomi, 2019; Pomeranz, 2015), though with limits (Carillo, Pomeranz, & Singhal, 2017).[8] The positive size-gradient also arises because the tax code in developing countries often leads to higher statutory tax burdens for larger firms and corporations (R. Kumar & James, 2022): Bachas, Brockmeyer, Dom, and Semelet (2023) find a positive size-statutory tax gradient among corporations in 15 countries. Our mechanism is motivated by Abbas and Klemm (2013), who hypothesize that the corporate sector expansion could explain why the reduction in statutory corporate tax burdens in developing countries has not led to a reduction in CIT revenue (% of GDP).[9] The mechanism also relates to studies in high-income countries that link CIT collection to the corporate sector's statutory burden, output-share and profitability (Clausing, 2007; Griffith & Miller, 2014; Sørensen, 2007).

We focus on a mechanism based on firm size, but many links between trade, firm structure, and taxation remain to be explored (Atkin & Khandelwal, 2020; Parenti, 2018).

[6]Our results, which focus on the corporate output-share, are compatible with findings from trade-formalization studies, which instead focus on the share of formal workers or firms (Section 6).

[7]See also Kopczuk and Slemrod (2006), Kleven, Knudsen, Kreiner, Pedersen, and Saez (2011), La Porta and Shleifer (2014), Bachas, Fattal, and Jensen (2019) and Best et al. (2021).

[8]In developed countries including the US, the large corporate sector is considered an important determinant of effective tax collection (Kleven, Kreiner, & Saez, 2016; Slemrod & Velayudhan, 2018).

[9]See also Quinn (1997), M. M. S. Kumar and Quinn (2012) and Abramovsky, Klemm, and Phillips (2014).

3 Construction of Effective Tax Rates

This section presents a new database of effective tax rates (ETR) on labor and capital, which covers 154 countries, starting in 1965 when possible, until 2018. We first outline the conceptual framework to build ETR, then present the data sources, and finally discuss the sample coverage. Further details are in Appendix B.

3.1 Methodology

Effective tax rates We compute macroeconomic effective tax rates following the methodology of Mendoza, Razin, and Tesar (1994). The effective tax rate on labor, denoted ETR_L, is the total amount of taxes effectively collected on labor divided by total labor income in the economy; similarly for capital, denoted ETR_K:

$$ETR_L = \frac{T_L}{Y_L} \qquad \text{and} \qquad ETR_K = \frac{T_K}{Y_K} \tag{1}$$

To construct the numerators, each type of tax revenue is assigned to labor or capital:

$$T_L = \sum_j \lambda_j \cdot \tau_j \quad \text{and} \quad T_K = \sum_j (1 - \lambda_j) \cdot \tau_j \tag{2}$$

where λ_j is the allocation to labor of each type j of tax τ_j. Types of taxes j follow the OECD Revenue classification. We allocate taxes as follows: (1) corporate income taxes, wealth taxes, and property taxes are allocated to capital; (2) payroll taxes and social security payments are allocated to labor; (3) personal income taxes (PIT) are allocated partly to labor and partly to capital, in a country-time specific manner (details below). Indirect taxes are neither assigned to labor nor to capital (but analyzed directly in Section 5.3). Table B2 provides a detailed allocation summary.

To construct the denominators, we decompose net domestic product as follows:

$$Y = Y_L + Y_K = \underbrace{CE + \phi \cdot OS_{PUE}}_{Y_L} + \underbrace{(1 - \phi) \cdot OS_{PUE} + OS_{CORP} + OS_{HH}}_{Y_K} \tag{3}$$

Labor income Y_L equals compensation of employees (CE) plus a share ϕ of mixed income (operating surplus of private unincorporated enterprises, OS_{PUE}). Capital income Y_K equals the remaining share $(1 - \phi)$ of mixed income, plus corporate firms' profits net of depreciation (operating surplus of corporations, OS_{CORP}), plus actual and imputed rental income (operating surplus of households, OS_{HH}).[10]

We also measure the effective tax rate on corporate profits, $\overline{ETR}_C^K$, as the ratio of corporate income taxes to corporate profits. This is an average effective tax rate at the corporate sector level; in Section 6, we analyze the firm-level corporate effective tax rate.

These macroeconomic ETRs rely on several conventions and assumptions (see Carey & Rabesona, 2004). First, as is done in the literature, they do not factor in economic incidence in that the economic cost of taxes is not "shifted" from one factor of production to another: all labor taxes are allocated to labor and all capital taxes are allocated to capital. Second, the tax revenue streams need to be comparable to their macroeconomic tax bases measured in national accounts. This generates two key challenges for our ETRs: (i) in the numerator, what share of personal income tax revenues to allocate to capital versus labor; and (ii) in the denominator, what share of mixed income to allocate to capital versus labor. We outline below our benchmark assumptions (detailed discussion is in Appendix B.2).

Allocation of personal income taxes (PIT) The main empirical difficulty in assigning taxes to labor and capital concerns the allocation of PIT. A naive procedure allocates 70% of the PIT to labor and 30% to capital, roughly matching the labor and capital shares of domestic product. In practice, however, recent work highlights that not all labor and capital income is subject to PIT, since not all individuals are required to file PIT, and exemptions apply to some income types (Jensen, 2022). Exemptions for capital (e.g., imputed housing rents, undistributed profits) are typically larger than for labor (e.g., pension contributions). Further, labor and capital income might not face the same tax rate: dual-income tax systems tax labor income with progressive rates but capital income with flat rates. In the US, Piketty, Saez, and Zucman (2018) use detailed tax and national accounts data to measure that 75% of labor income is subject to PIT, versus 33% of capital income. This suggests allocating 15% of PIT to capital and 85% to labor.[11]

Starting from this baseline where 15% of PIT revenues derive from capital, we perform two country-year adjustments: (i) we raise capital revenues for country-years with a high PIT exemption threshold in the income distribution (Jensen, 2022); (ii) we lower it in country-years where dividends face lower taxes than wages. The resulting capital share of PIT revenue varies between 7% and 32% across country-years. Over time, this share falls

[10]We decompose net domestic product (NDP), which subtracts consumption of fixed capital from gross domestic product (GDP). NDP is lower than GDP, by 10% on average. We exclude capital depreciation since it does not accrue to any factor of production and is usually tax-exempt. Factor incomes also exclude indirect taxes (which are also excluded in the numerator of ETR).

[11]If 75% of labor income is taxable and labor income is 70% of national income (resp. 33% and 30% for capital income), then $75\% \times 70\%/(75\% \times 70\% + 33\% \times 30\%) = 84\%$ of the PIT is labor income.

from a global average of 19% in 1965 to 14% in 2018, due to a reduction in PIT exemption thresholds and increased prevalence of dual tax systems.

In the absence of detailed tax records in every country and year, these adjustments provide an imperfect approximation of the true capital share of PIT. We therefore implement two simple robustness checks where the share allocated to capital is fixed over time at either 0% or 30%, representing low and high-end scenarios.

The labor share of mixed income The labor share of mixed income (unincorporated enterprises) is hard to measure.[12] For our benchmark series we assume $\phi = 75\%$, i.e., 25% of mixed income is considered capital income.[13] In the absence of a consensus over alternatives this assumption has the advantage of being transparent, though factor shares are unlikely in practice to everywhere be time and country-invariant. We therefore implement two robustness checks, which create time and year variation in ϕ. The first method, based on ILO (2019), uses micro-data to estimate the country-specific labor income of self-employed based on the observable characteristics of these workers and their comparison with employees.[14] Second, we assign to ϕ the observed country-year labor share of the corporate sector (as in Gollin, 2002).

The exact ETR formulas which include the above adjustments are in Appendix B.2.

Usefulness and limitations of ETR Since national account statistics are compiled following harmonized guidelines, ETRs are conceptually comparable over time and across countries, though the data limitations described above should be kept in mind. By relying on taxes actually collected, the ETRs incorporate tax avoidance and evasion behavior as well as the net past effects of all tax policies, including rates, exemptions and credits. This is particularly relevant in a development context, where due to widespread evasion, knowledge of statutory tax rules only provides a partial picture of effective tax burdens.

The ETRs are backward-looking measures that comprehensively capture how much capital and labor have effectively paid in taxes. They are helpful for three reasons. First, knowing how much revenues are effectively collected from each factor is important when governments face fiscal pressure (Besley & Persson, 2014): this is characteristic of most

[12] The UN's national accounts framework outlines the combination of multiple methods to overcome challenges of measuring the *level* of mixed income in economies with widespread informality. While information on the methods used is not available on a country-year basis, an inspection of the published frameworks suggests no change in methodologies for mixed income over time.

[13] This is below the 30% used in Distributional National Accounts (DINA) guidelines (Blanchet, Chancel, Flores, & Morgan, 2021), but since the global average of the corporate sector's capital share is 27%, assuming a lower capital share for unincorporated enterprises seems reasonable (see Guerriero, 2019).

[14] Details in Appendix B.2. A challenge with this method is that it can create implausibly large estimates of the level of mixed income compared to their values in national accounts. We implement an adjustment to help with this limitation, but for this reason we choose to use ILO (2019) only for robustness.

developing countries, where potential revenue losses or gains is a key policy determinant. Second, the level of the ETR and its deviation from a statutory rate is frequently an input into policy-making to understand the size of tax gaps (e.g. the recent focus on the firm-level ETR in the global minimum tax agreements). Finally, the tax burden levied on each factor is an important starting point to determine the economic incidence of a tax system.

A limitation of macroeconomic ETRs is that they are impacted by both the tax code and economic changes. Thus, studying ETRs is most helpfully done in combination with analyzing its mechanisms, which we focus on in Section 6. Related, we emphasize that the ETR should not be interpreted as a proxy for the statutory tax burden. An important complementary body of work carefully measures legal tax burdens (Devereux & Griffith, 1999), by constructing forward-looking average tax rates on capital based on the simulated present value of returns and costs of a new investment. Driven by differing objectives, the backward-looking and forward-looking measures are related, yet distinct.[15]

3.2 Data sources

3.2.1 National accounts

To measure factor incomes for 154 countries since 1965 when possible, we create a panel of national accounts using data from the System of National Accounts (SNA) produced by the United Nations. We first use the 2008 SNA online repository that has global coverage for recent decades. In turn, the UN Statistics Division provided us with access to the 1968 SNA offline data which covers historical data from the 1960s and 1970s. To the best of our knowledge, our paper is the first to harmonize and integrate the 2008-SNA and 1968-SNA datasets.[16] Estimating factor incomes requires information on all the components of national income (equation 3). Whenever we have national income for a country-year but information on a component is missing, we attempt to recover it with information from the second SNA dataset, as well as using national accounting identities with non-missing values for the other income components. In the remaining cases, we impute component values following DINA guidelines (Blanchet et al., 2021) (details in Appendix B.1).

3.2.2 Tax revenue

We construct a new tax revenue dataset that disaggregates taxes by type following the OECD Revenue Statistics classification of taxes. Our database includes all taxes—on per-

[15]This is particularly the case for $\overline{ETR}_C^K$: see supplementary appendix for a detailed discussion. Our measure of $\overline{ETR}_C^K$ also relates to the CIT-efficiency measure by IMF (2014). In the supp. apppendix we find that CIT-efficiency measured with our data in the relevant sample matches well the IMF (2014) series.

[16]Relative to recent work (including Guerriero, 2019; Karabarbounis & Neiman, 2014), our data expands coverage in space and time, mainly to developing countries, and systematically attempts to measure factor incomes for total domestic output (vs. only for the corporate sector).

sonal and corporate income, social security and payroll, property, wealth and inheritance, consumption and international trade—at all levels of government. We ensure a systematic separation of income taxes into personal and corporate income. We collect new archival data and integrate it with existing data sources.

When available, OECD Revenue Statistics data (link) is the preferred source, as it covers all types of tax revenues and goes back to 1965 for OECD countries. It accounts for 2,875 country-year observations (42.3% of the sample). Its drawback is its limited coverage of non-OECD countries, as it covers 93 countries in total and only covers developing countries more recently. We add data from ICTD (link). ICTD includes most developing countries, with coverage that starts in the 1980s. ICTD sometimes combines personal and corporate income taxes, and sometimes lacks social security. ICTD adds 1,246 country-year observations (18.3% of the sample).

To complement these existing sources, we conducted archival data collection to digitize records from government budgets and national statistical yearbooks. This adds 2,011 new country-year observations.[17] We supplemented these archival records with countries' online publications, and offline data from the IMF Government Finance Statistics (1972-1989). In total, this data collection adds 2,678 observations (39.4% of the sample).

Building a dataset based on newly digitized historical sources necessarily requires making a number of decisions. To increase the credibility of our data, we follow four guiding principles. First, we seek to build long historical time-series that overlap in years with existing sources. We aim to only use two data sources per country, but use the overlapping years between multiple sources to corroborate that they are comparable in levels of tax revenue and types of taxes in place.[18] For this reason, a switch in data source rarely leads to a significant change in trend. Second, for the historical periods without overlap with existing data, we corroborate the levels of tax to GDP with academic and policy studies. Third, we draw on historical studies to verify that large changes in revenues collected reflect policy, economic or political changes rather than data artifacts. Fourth, we aim to be conservative and exclude countries in time periods where concerns exist about data quality, due to the economic and political context.

To help assess our approach, the supplementary appendix provides additional material. We provide a table which outlines, in each of the 154 countries, the main considerations and our choices relating to the four guiding principles. The table emphasizes the uncertainty surrounding specific countries and time periods, and flags instances where

[17] The archives were accessed in the Government Section of the Lamont Library (website link).

[18] OECD is the preferred starting point and archival data is initially second in priority since it often disaggregates tax types and goes back far in time, but we revise this based on the source that best matches the OECD data. The supplementary appendix summarizes the data sources used for each country.

the data appears worthy of inclusion but should be interpreted with caution (all our main results are unchanged if we exclude these instances). Moreover, we provide in-depth country case-studies with direct links to the initial archival sources; the case-studies currently cover all countries with more than 15 million inhabitants but will ultimately expand to cover all 154 countries. We invite comments from researchers to help improve the accuracy of the series as we continuously update the data.

3.3 Data coverage of effective tax rates

The final ETR sample contains 6,799 country-year observations in 154 countries (Figure A1). The number of countries starts at 78 in 1965 and grows to 110 by 1975 (due to independence or country creation). The key jump in coverage —from 117 to 148 — corresponds to the entry of ex-communist countries in 1994, including China when it arguably built a modern tax system (Appendix B.1). The data is effectively composed of two quasi-balanced panels. The first covers 1965-1993 and excludes communist regimes, accounting for 85-90% of world GDP. The second covers 1994-2018 and includes former communist countries, accounting for 97-98% of world GDP. Figure A1 shows coverage by development level. We use the World Bank income classification in 2018, classifying low and middle-income countries (LMICs) as developing countries and high-income countries (HICs) as developed countries. We refer interchangeably to LMICs as developing countries and HICs as developed countries. Our sample contains 5,144 observations in LMICs and 1,655 observations in HICs.

Comparison with existing datasets Our database complements previous ETR series by expanding coverage to LMICs. Table B3 summarizes the coverage of existing ETR series, which focus on HICs (Carey & Rabesona, 2004; Kostarakos & Varthalitis, 2020; McDaniel, 2007; Mendoza et al., 1994). Our benchmark ETRs rely on specific choices: Table B3 summarizes the methodological differences with existing ETR series, which relate mainly to allocating capital to both mixed income and PIT.[19] The alternative choices are covered by the robustness checks of Section 3.1, which are implemented in Section 4.2.

4 Stylized Facts on Global Taxation Trends

4.1 Evolution of effective tax rates on capital and labor

Figure 2 documents the evolution of effective tax rates on capital and labor from 1965 to 2018. Aggregates are dollar-weighted, i.e., the global effective tax rate on capital equals worldwide capital tax revenues divided by worldwide capital income. These series can

[19] A comprehensive discussion of the methodological differences is provided in the supp. appendix.

thus be interpreted as the average tax rate on a dollar of capital income derived from owning an asset representative of the world's capital stock. The top panel shows global trends and the bottom panels separate HICs and LMICs.

Globally, effective tax rates on labor and capital converged between 1965 and 2018, due to a rise in labor taxation and a drop in capital taxation. The global ETR_L rose from 16% in the mid-1960s to 25% in the late 2010s, while ETR_K fell from 32% to 26%.

The global trends mask heterogeneity by development levels. While labor taxation rose everywhere, the decline in capital taxation only occurred in HICs: the effective capital tax rate fell from 38-39% to 32-33% between the late 1960s and late 2010s, fueled by a large reduction in effective corporate tax rates, which fell from 27% to 19%. In contrast, starting from a low level, ETR_K increased in LMICs, with the rise happening entirely since the beginning of the 1990s. ETR_K started at 10% in the mid-1960s and was at the same level in 1989; in between, there was a rise and decline in the late 1970s, but this temporary change was fully driven by resource-rich countries (Figure 4). From 10% in 1989, ETR_K saw a sustained increase over the next two decades which reached 18% in 2018. The rise in capital taxation is partly driven by higher effective taxation in the corporate sector: the effective corporate tax rate rose from 12% to 20% between 1989 and 2018 in LMICs.[20]

4.2 The rise of capital taxation in developing countries

The secular decline in ETR_K in HICs has been documented before (Dyreng, Hanlon, Maydew, & Thornock, 2017; Garcia-Bernardo, Janský, & Tørsløv, 2022), but the rise in ETR_K in LMICs starting at the beginning of the 1990s is novel. We therefore need to establish that this result is robust to the assumptions we used to construct the ETR series.

The ETR series depends on four main methodological decisions: (1) how to assign PIT revenue to capital vs labor; (2) how to allocate mixed income to capital vs labor; (3) balanced vs. unbalanced panel; (4) weights to aggregate countries. Our benchmark series: (1) assigns PIT to capital vs. labor for each country-year using data on PIT exemption thresholds and the tax treatment of dividends relative to wages; (2) allocates a fixed 25% of mixed income to capital; (3) consists of two quasi-balanced panels before and after 1994 (when China, Russia and other former command economies enter the sample); and (4) weighs countries by their share of worldwide factor income in each year. We assess how results change when varying one, several, or all of these choices at the same time.

[20] Figure A2 shows that both corporate tax revenues and corporate profits increased since 1989 but the former outpaced the latter, causing $\overline{ETR}_C^K$ to rise. Corporate profits and tax revenue are the largest components that determine changes in ETR_K. Smaller contributions to ETR_K's rise come from the decline in mixed-income, and the steady increase in property and wealth taxes, which outpaced income from rents, albeit starting from a very low level (0.3% of NDP in 1989).

Figure 3 investigates the robustness of the ETR_K trend in LMICs.[21] Panel (a) varies the allocation of personal income tax (PIT) revenue. Our benchmark follows a data-driven country-year assignment; instead we consider two simpler scenarios where the share allocated to capital is fixed, at either 0% or 30% (low and high-end scenarios). Due to high PIT exemption thresholds in LMICs, the benchmark country-specific assignment is closer in levels to the 30% than to the 0% scenario. Though the capital share allocated to PIT slightly changes over time (Section 3.1), the time-invariant robustness series track the trends in the benchmark series closely. This is because the PIT remains limited in LMICs, such that its split into labor versus capital is of minor consequence.

Panel (b) varies the assignment of factor shares in mixed income. We implement two robustness checks, creating mixed income labor shares that vary at the country-level based on the ILO (2019) method, and at the country-year level based on the observed corporate labor share. Both alternative series are very similar to the benchmark.

Panel (c) quantifies the effect of country entry into the panel. In our benchmark, China, Russia, and other former command economies enter in 1994. In this robustness, we balance the panel by imputing missing observations between 1965 and 1993; we use the observed ETR_K value for that country in 1994 and the trends in ETR_K observed for other LMICs in 1965-1993. The imputation raises ETR_K between 1965 and 1993, because Russia had both a high ETR_K and a high weight when entering the sample in 1994.

Panel (d) aggregates countries using net domestic product (NDP), instead of capital income weights. The NDP weights are either time-varying or fixed in 2010. These alternative weighting procedures suggest a slightly higher increase in ETR_K over time.

Finally, panel (e) plots all 54 combinations of the four methodological choices. The rise in ETR_K in LMICs between 1989 and 2018 is clearly apparent in each of the 54 series. How wide is the range of increases and how does our benchmark series compare? Computing the 1989-2018 change in the 54 series, we obtain a fairly tight range of ETR_K increases, between 6.4ppt and 10.3ppt. Our benchmark is at 8.7ppt, which is close to the mean increase of 9.2ppt; there are larger increases than our benchmark in 43 series and smaller increases in 10.[22] Our benchmark series corresponds to an 87% increase in the effective tax rate on capital in LMICs since 1989, reflecting both the strong growth and low baseline.

Comparison with previous studies Pre-existing ETR series mainly cover HICs, which limits the comparison to our sample. In HICs, our benchmark ETR trends are comparable

[21] The robustness for ETR_L in LMICs, and ETR_L and ETR_K in HICs are shown in the supp. appendix.

[22] Setting 1989 as the base year is partly arbitrary, but it allows us to fix a starting level for ETRs immediately before the period of strong trade liberalization in LMICs. If we instead compute the change in ETR_K between 2018 and the lowest point in a given series, the range of changes is 6.8-11.3ppt across the 54 series, with a mean at 9.6ppt and our benchmark at 9.4ppt.

to previous studies, but the levels differ by 16.5% on average (Figure B1). This difference is primarily due to methodological assumptions about the allocation of capital to mixed income and PIT (Table B3). However, the alternative methodologies from the pre-existing series are contained within the range of ETR trends produced by our robustness checks. In HICs, that range of ETR_K trends is indeed wide (due to the quantitative importance of the PIT; ETR series in supp. appendix). However, the range of ETR_K trends in LMICs is comparatively tighter (Figure 3); this is because the rise in ETR_K in LMICs is primarily driven by the corporate sector (Figure A2), which is not strongly affected by the methodological differences between our study and pre-existing studies.

4.3 Where has capital taxation risen?

Figure 4 shows the ETR_K series for subsamples of countries. Panel (a) plots ETR_K series for the most populous LMICs: Brazil, China, India and Indonesia. All display a marked ETR_K rise over time. Starting in 1989 (1994 for China) until 2018, ETR_K rose from 7% to 27% in Brazil, 6% to 24% in China, 5% to 12% in India, and 9% to 15% in Indonesia.

China's weight and fast-rising capital taxation imply that it plays a key role in the aggregate ETR_K trend in LMICs. Panel (b) shows that, when excluding China, the rise in ETR_K is half as large (from 10% to 15%) and a more significant part of the rise occurred earlier in the 1990s. Panel (c) shows that oil-rich countries, measured as deriving more than 7% of GDP in oil in 2018, have been on a completely distinct path. Reflecting the oil commodity shocks, their ETR_K rose in the 1970s but fell in the 1980s, and have stayed flat since. Excluding oil-rich countries yields a more pronounced ETR_K rise, from 9% in 1989 to 21% in 2018, and a flat ETR_K series from 1965 to 1989. If we exclude both China and oil-rich countries, we observe a rise in ETR_K from 9% in 1989 to 17% in 2018, which is similar in magnitude to the benchmark series.

Panel (d) reveals that, among non-oil-rich countries, the ETR_K rise is stronger in large LMICs, defined as the 19 countries with a population above 40 million in 2018. Even when excluding China, the ETR_K of the other 18 most populated countries rose from 9% to 18% between 1989 and 2018; in smaller countries, ETR_K rose from 10 to 14% over the same period. The ETR_K has risen by more than 5 percentage points in 13 of the 19 largest LMICs since 1989, and has only fallen in one country (Russia).[23]

In short, the rise in effective capital taxes is more pronounced in larger countries, including China, but is a general pattern in developing countries, except for oil-rich ones.

[23]The supplementary appendix shows the individual countries' ETR_K and ETR_L time series.

4.4 Suggestive evidence for the role of globalization

We saw that ETR_K fell in HICs but rose in LMICs. Importantly, the rise in LMICs starts in the early 1990s, which coincides with the onset of the "hyper-globalization" period that could a priori have made capital more mobile and harder to tax.[24] Instead, could trade globalization have caused ETR_K to rise in LMICs? Here we take a first pass at investigating this question. We create 5-year growth rates within countries in trade and ETRs. We plot binned scatters of ETR against trade openness (measured as the share of imports and exports in NDP), after residualizing all variables against year fixed effects. Figure 5 depicts these within-country associations, which condition on global time trends. Mirroring the heterogeneity in long-run trends, we observe differences by development level in the association between trade and ETR_K: openness is associated with increases in ETR_K in LMICs, but with decreases in ETR_K in HICs.[25] In sum, from a global and historical perspective, the correlational evidence suggests that trade may have contributed to the newly documented rise in ETR_K in developing countries.

Naturally, LMICs have undergone significant development since the 1960s and this growth is likely to also have contributed to the long-run rise in ETR_K. In the supplementary appendix, we find that the associations in Figure 5 hold in LMICs when controlling for GDP per capita growth. This correlational evidence, combined with the observation that while globalization is a major process in LMICs, its revenue impacts are still not established (Section 1), motivate us to investigate trade as a determinant of ETR and study its mechanisms.

5 Trade Globalization and Capital Taxation

In this section, we implement two distinct research designs to investigate the impact of trade openness on capital and labor taxation in developing countries.

5.1 Event-studies for trade liberalization

5.1.1 Empirical design

In the first design, we implement event studies of trade liberalization policy events in key developing countries. To discern sharp breaks from trends in our outcomes, our selection criteria was to select events that caused large trade barrier reductions and which

[24]Individual trends in the four largest LMICs (Figure 4) also suggest an association between liberalization episodes and an uptick in ETR_K (Brazil in 1988; China in 2001; India in 1991; Indonesia in the mid-1980s).

[25]The supplementary appendix further shows that early globalized LMICs saw trade and ETR_K rise in tandem prior to the 1990s and stagnate thereafter. By contrast, LMICs which participated in the second wave of globalization post-1990 saw a rise in trade and ETR_K in the 1990-2018 period.

have been studied in the literature. This led us to select the six events from the review papers by Goldberg and Pavcnik (2007, 2016) (Colombia in 1985, Mexico in 1985, Brazil in 1988, Argentina in 1989, India in 1991, Vietnam in 2001), and add the well-known event of China's accession to WTO in 2001 (Brandt et al., 2017). These liberalization events led to large reductions in tariffs: from 59% to 15% in Brazil; 80% to 39% in India; and, 48% to 20% in China. We can rely on pre-existing narrative analyses to discuss threats to identification and interpretation of results.[26] Appendix C.1 provides more details on our selection criteria and the liberalization events.

For each of the seven treated countries and outcomes, we construct a synthetic control country, as a weighted average over the donor pool of never-treated countries (Abadie, Diamond & Hainmueller, 2010).[27] We match on the level of each outcome in the 10 years prior to the event, while minimizing the mean squared prediction error between the event-country and the synthetic control.[28] We plot the average levels of the outcome for treated and synthetic control countries by relative time to the event. Moreover, we estimate the event-study model in 10 years both before and after the events:

$$y_{ct} = \sum_{e=-10, e\neq -1}^{10} \beta_e \cdot \mathbb{1}(e=t)_t \cdot D_c + \theta_t + \kappa_c + \pi_{Year(t)} + \epsilon_{ct} \tag{4}$$

where we include fixed effects for event-time, θ_t, country κ_c, and calendar year, $\pi_{Year(t)}$ (the latter control for shocks that correlate with events clustered in calendar time). D_c is a dummy equal to one if country c is treated. The coefficient β_e captures the difference between treated and synthetic control countries in event time e, relative to the pre-reform year $e = -1$ (omitted period). Since inference based on small samples is challenging, we plot 95% confidence bounds using the wild bootstrap, clustered at the country event level. In Table A1 we estimate the simple difference-in-differences, which captures the average treatment effect in the 10 years post-liberalization, and the imputed treatment effect based on Borusyak, Jaravel, and Spiess (2021), which addresses challenges from two-way fixed effects and heterogeneous event-times.

5.1.2 Event-study results

Figure 6 displays the event studies in levels (left-hand panels) and the dynamic regression coefficients (right-hand panels). The top panels show that, as expected, trade openness

[26]The reductions in trade barriers are sometimes implemented over several years. To be conservative, we focus on the earliest start year for each event as defined in published studies.

[27]For each country-event, we can include eventually-treated countries in the donor-pool (excluding those with treatment within 5 years of the event); the results, available upon request, are similar.

[28]The supp. appendix lists the countries included in the synthetic control for each event and each outcome.

rises in the years post-event by 10 percentage points, and its trend changes in post-reform years compared to pre-liberalization years.[29] The middle panels show that ETR_K followed stable pre-trends and sharply rises post-liberalization, by 4 to 5 percentage points. The bottom panels show ETR_L also rose, but only by 2 percentage points. Despite the small sample size, the dynamic post-treatment coefficients for each period are often significant at the 5% level. The p-values for the joint significance of all post-reform dummies are well below 5%. Table A1, Panel A, reports the DiD results, which are marginally more significant when estimated from imputed treatment effects. Panel B shows that results are comparable when we jointly match on all outcomes for each country-event.

The identifying assumption is that there are no changes in confounding determinants of ETR that coincide with the liberalization events. The breaks from stable pre-trends imply that confounding changes would have to sharply coincide with the event onset. Narrative analyses of the timing for each event (Appendix C.1) do not suggest obvious concurrent changes. Moreover, using data from Wacziarg and Wallack (2004), we verify that other cross-border reforms (e.g. capital liberalization) or domestic reforms (e.g. privatization) do not occur in the same year as the trade events.[30] However, reforms sometimes occur a few years after: for example, Mexico joined NAFTA and removed capital inflow restrictions, Argentina and Brazil joined MERCOSUR, and India liberalized its FDI rules (Appendix C.1). These reforms occurred several years post-trade liberalization, but ETRs sharply rise in the immediate post-event years. This discussion highlights that the causal interpretation of trade-centered macroeconomic reforms requires caution. A plausible interpretation is that the short-run rise in ETRs with sharp breaks from stable pre-trends reflects the impact of trade reforms, but that the medium-run coefficients also reflect the impacts of additional, mainly cross-border, reforms.

Our results are based on a (small) sample of liberalization events that satisfied specific criteria. In Appendix C.3, we study the robustness to using very different selection criteria for trade liberalization. Specifically, we re-estimate the event-study using the 68 liberalization events in LMICs from Wacziarg and Welch (2008). We find very similar impacts of trade on ETR using this alternative and expanded set of liberalization events.

We further probe the identification and robustness of our results. First, given the limited number of liberalization events, we investigate if the average effects are influenced by one particular event. Removing one treated country at a time, we find the dynamic treatment effects for all subsets of events are similar to the full sample (supp. appendix).

[29]The absence of a pre-reform dip limits concerns about inter-temporal substitution, although some of the liberalization events may have been predictable, including China's WTO accession.

[30]Only Mexico had a concurrent domestic reform, and results hold without it: see supp. appendix.

Second, Table A1 addresses concerns related to the control group. We find similar results when we remove from the donor pool each liberalizing country's 5 major export and import trading partners (measured in the immediate pre-event years), alleviating concerns of spillovers to countries in the control group.[31] Results are also comparable when the donor pool excludes countries that have already liberalized (based on Wacziarg & Welch, 2008), to guard against the concern that the trends in the synthetic control group reflect the longer-run effects of the treatment (liberalization). Finally, to lessen the concern that treated and control countries experience different unobservable shocks, we find similar results when the donor pool for each treated country is restricted to the same region (or to LMICs only).

5.2 Regressions with instrumental variables for trade

5.2.1 Empirical design

Our second design employs instrumental variables for trade. One attractive feature is that the IV provides causal estimates under different identifying assumptions than the event-study. We estimate the following model in developing countries:

$$y_{ct} = \mu \cdot trade_{ct} + \Theta \cdot X_{ct} + \pi_c + \pi_t + \epsilon_{ct} \tag{5}$$

where y_{ct} is the ETR (or another outcome) in country c in year t, $trade_{ct}$ is the share of imports and exports in NDP, and π_c and π_t are country and year fixed effects.[32] We cluster ϵ_{ct} at the country level. X_{ct} contains confounding determinants of ETR: the exchange rate, gross capital formation, log of population, and capital openness (Chinn & Ito, 2006 Rodrik, 1998). ETR time series are sometimes volatile (Figure 4), so we winsorize ETR at the 5%-95% level by year separately for LMICs and HICs.

OLS estimation of equation (5) may be biased due to reverse causality and unobservable confounding factors that correlate with trade. To try to address these issues, we use the two instruments for trade from Egger et al. (2019). The first instrument, denoted $Z^{gravity}$, relies on the structure of general equilibrium models of trade. Under the standard gravity model assumptions, it uses the average bilateral trade frictions between exporting and importing countries as variation (aggregated to the country-year level). This instrument

[31] We also verify that none of the main countries in the synthetic control (supp. appendix) had external or domestic reforms in the event-year or in the post-event periods (using the data in Wacziarg & Wallack, 2004). Consistent with this, the levels of the outcomes in the synthetic control are relatively stable throughout the event periods. Finally, note that if the spillovers correspond to coordination of policies, this would likely bias our estimation towards finding null effects.

[32] We include fixed effects for imputed and interpolated values, as well as for each tax and national account data source (Section 3.2), to ensure results are not driven by changes to data quality. Results also hold without imputed values and within each data source (Table A3).

is valid if the distribution (not the level) of trade costs among individual country-trading pairs is not influenced by ETRs in the import or export country. The second instrument, denoted $Z^{oil-distance}$, interacts time-series variation in global oil prices with a country-specific measure of access to international markets. Access is captured by the variance of distance to the closest maritime port for the three most populated cities. This time-invariant measure captures the internal geography of a country and impacts transportation costs: following a global shock to oil prices, transportation costs will be higher in countries with less concentrated access to ports, leading to a larger drop in imports and exports. This instrument is valid if the interaction between global oil prices and country-specific measures of spatial concentration is uncorrelated with changes in ETRs. Conceptually, both instruments capture variation in trade costs driven by economic forces that are plausibly exogenous to ETRs and their determinants (details in supp. appendix).

In LMICs, the 1^{st}-stage is stronger in the 2000s and at higher income levels for Z^{oil}, and in earlier periods and at lower income levels for $Z^{gravity}$ (supp. appendix). Restricting the analysis to sub-samples where one of the instruments has a strong first-stage introduces bias (Mogstad, Torgovistky, & Walters, 2021). Instead, we combine the two instruments to estimate a local average treatment effect that is representative of LMICs across income levels and time periods. Table A2 shows the 1^{st}-stage.[33]

5.2.2 Instrumental variable results

Table 1 presents the results in LMICs for ETR_K in Panel A, and ETR_L in Panel B.[34] In column (1), the OLS uncovers positive, significant associations between trade and both ETRs. In column (2), we employ the two instruments. The 1^{st}-stage Kleibergen-Paap F-statistic is 24.59. The IV shows that trade causes an increase in both effective tax rates, and the coefficient for ETR_K (0.151) is three times larger than for ETR_L (0.047). These magnitudes are economically meaningful: moving from the 25^{th} to the 75^{th} percentile of trade openness in LMICs would cause a 8.9 percentage points increase in ETR_K.

The remaining columns of Table 1 present three sets of robustness checks. In the first set (Columns 3 to 7), we modify the specification. The most notable difference is that the coefficient on ETR_K increases (to 0.211) when we weigh the regression using NDP (Column 4), putting thus more weight on the variation in larger developing countries. The results hardly change when we: use non-winsorized ETRs (Column 3); include

[33]Table D1 shows the instruments impact imports and exports, and trade in intermediate goods-services (G-S) and final G-S. Thus, our IV-estimates comprehensively reflect the impacts of trade through rises and falls in final and intermediate goods and services that flow both in and out of the country.

[34]Relative to ETR coverage, the sample size drops by 4.5% due to data-availability of instruments.

controls (Column 5);[35] allow oil-rich countries to be on a separate non-parametric time path (Column 6), which addresses the concern that the identifying variation for $Z^{oil-dist}$ is correlated with trends in ETRs specific to oil-rich countries (Figure 4); winsorize trade openness (Column 7).

In the second set of robustness checks, we implement the alternative capital and labor assignments to PIT and mixed-income, described in Section 4.2. In our benchmark, the capital share of mixed income is time-invariant, yet trade may cause factor shares to change. In columns (8)-(9), we allow factor shares to respond to trade by implementing the two methods which create country-year varying capital-shares of mixed income: the results are comparable. They also remain similar when we fix the capital share of PIT at 0% (column 10) or at 30% (column 11) in all countries over time. In the third robustness set (columns 12-13), we estimate IVs using each instrument separately. The 1^{st}-stage F-statistic is 45.13 for $Z^{gravity}$ and 10.75 for Z^{oil}. The IV estimates are comparable, though larger when based on Z^{oil}. Leveraging the opposite sign effects of the two instruments on trade, the reduced form results (Table A2) suggest that openness effects are symmetric: increased trade increases both ETR_L and ETR_K, while reduced trade decreases both ETRs.

Finally, our results are based on an unbalanced panel combining several data sources (Section 3.2- 3.3). Table A3 shows that the results are qualitatively similar within each data-source for taxes (newly digitized government records; OECD; ICTD) and national accounts (SNA1968; SNA2008), as well as in both quasi-balanced panels (pre and post-1994) and in a strongly balanced panel (1965-2018).[36]

Quantifying the role of trade How much of the ETR_K rise in LMICs since 1989 can be accounted for by rising trade? Between 1989 and 2018, the weighted ETR_K in LMICs rose by 8.7ppt (Section 4.2) and trade openness by 13.6ppt. The NDP-weighted IV for trade's impact (col.4 of Table 1) is arguably the most comparable, since the ETR_K trends in Section 4 are also weighted. Using this estimate would imply that trade openness accounts for 33% of the rise in ETR_K ($0.211 * 0.136/0.087 = 0.329$). Considering all estimates in Table 1 generates a range of 21-42%.[37]

5.3 Impacts of trade openness on total tax revenues

We find positive effects of trade on capital and labor taxation, but how does trade impact *overall* tax revenues, including indirect taxes? This is a relevant question, as trade-induced tax losses from liberalization remain an important concern for policymakers (Hallaert,

[35]Results also hold when controlling for GDP per capita (not shown).
[36]Variation between coefficients may reflect data quality or 1^{st}-stage and treatment heterogeneity.
[37]For reasons discussed in 5.1.2, we do not use the event-study estimates for this exercise.

2010; World Bank, 2020). Table 2 presents the OLS and the IV estimation of the effect of trade on total taxes (% of NDP), in LMICs, as well as on individual tax revenue streams. Total taxes include direct taxes on capital and labor, as well as indirect taxes (sum of taxes on trade and domestic consumption).[38]

The trade coefficient for total tax collection is positive and significant in both the OLS and the IV. The IV coefficient (0.101) is economically large: moving from the 25^{th} to the 75^{th} percentile of openness in LMICs would cause a 5.9ppt increase in total taxes (the unweighted average tax/NDP ratio in LMICs is 17.6%). This result is mainly driven by higher corporate income taxes, which account for just over half of the increase in total taxes, and to a lesser extent by social security and personal income taxes.[39] Trade has a positive, but statistically insignificant, effect on indirect taxes.

Trade's impact on total taxes is robust to using NDP-weights; including controls; winsorizing trade; and using each instrument separately (Table A4).

We can also study the impact of the trade liberalization events from Section 5.1 on total tax revenue. Using the event-study methodology, Figure A3 shows that the trade events led to an increase in overall tax collection, with a break from the stable pre-trend.

One limitation is that we do not separately study trade's impact on tariff revenues versus domestic consumption taxes, as our data does not contain a systematic breakdown between these two indirect taxes. This reflects our initial focus on direct capital and labor taxes, but additional data work would permit a separation of these indirect taxes.[40]

Both the event study and the IV indicate that trade leads to higher overall taxation in LMICs. This finding relates to the literature on the net impact of openness on tax revenues, which finds mixed results due to differences in sample, empirical strategy and definition of openness (Section 2);[41] moreover, some of these studies focus on the net impact of trade on indirect taxes and abstract from direct domestic taxes. We contribute by comprehensively studying the total tax impact of openness, based on implementing several identification strategies in the largest sample of developing countries to date.

6 Mechanisms

This section investigates mechanisms for trade's impact on taxes, especially ETR_K.

[38] Long-run trends in taxation by type and development level are in the supplementary appendix.

[39] CIT grew significantly, as a share of NDP, between 1989 and 2018: see Figure A2.

[40] While the sign of openness' impact on tariff revenue could in principle differ depending on whether the reduction in trade costs is initially due to economic forces (as in the IV) or policy changes (as in the event study), we find positive impacts in both cases on domestic capital and labor taxes, and on total taxes.

[41] An important study in this literature, Baunsgaard and Keen (2009) writes in the conclusion: "it is possible that indirect effects operating through higher levels of openness and income consequent upon trade reform have more than offset the direct loss of revenue identified here."

6.1 Outlining the tax capacity mechanism

The *tax capacity* mechanism combines two distinct insights from the trade and public finance literature (Section 2): first, trade expands activity in corporate structures and larger firms relative to smaller businesses and self-employment; second, effective taxation increases with firm size. To fix ideas, consider the following decomposition of ETR^K:

$$ETR^K = \int_{i \in C} ETR_i^K f(i)\, di + \int_{i \in NC} ETR_i^K f(i)\, di \tag{6}$$

$$= \mu_C^K \cdot \overline{ETR}_C^K + (1 - \mu_C^K) \cdot \overline{ETR}_{NC}^K \tag{7}$$

This decomposition shows that the effective tax rate on capital ETR^K is composed of two parts.[42] The first part captures capital taxation within the corporate sector. It is the product of the corporate sector's share of NDP, μ_C^K, and the average effective tax rate on capital in the corporate sector, $\overline{ETR}_C^K$. The former is directly measured in national accounts (employee compensation plus corporate profits net of depreciation), while the latter is computed as the ratio of corporate income tax revenue to corporate profits. In the second part, $\overline{ETR}_{NC}^K$ measures the effective tax rate on capital in the non-corporate sector; it is multiplied by the non-corporate sector's income share, $1 - \mu_C^K$, which includes mixed income of unincorporated enterprises and household surplus (rents and imputed rents).[43]

In LMICs, $\overline{ETR}_C^K$ is 50% larger than the overall ETR^K (19.9% versus 13.3%). This stems from both stronger enforcement and higher statutory tax burdens in the corporate sector.[44] Hence, the expansion of the corporate sector relative to the non-corporate sector (i.e. an increase in μ_C) could increase ETR^K.

The conjecture that trade exerts a tax capacity effect is rooted in the literature on trade and firm size (described in Dix-Carneiro et al., 2021). First, trade can lead to increased market opportunities that disproportionately benefit large exporters (Melitz, 2003). Second, trade can expand the availability of intermediate goods and lower their prices, which could disproportionately benefit initially larger firms (for example due to fixed costs as in Kugler & Verhoogen, 2009). Through these two channels, trade could expand the corporate sector's share of national income (μ_C), as larger firms are more likely to be incorporated. Moreover, by benefiting initially larger firms or leading to firm

[42] In this section, capital taxation is denoted with a K-superscript to accommodate additional notation.

[43] $\overline{ETR}_{NC}^K$ is measured as the ratio of tax revenue from property and wealth, self-employment, and the PIT assigned to capital, over capital mixed-income and the surplus of the household sector. It is thus composed of a mix of variables, which are arguably not as well measured as those from the corporate sector.

[44] The ability to levy higher tax rates is endogenous to enforcement (Bergeron, Tourek, & Weigel, 2024). Our notion of tax capacity is that these co-determined forces jointly contribute to effective taxation.

size growth within the corporate sector, trade could also increase the average effective corporate tax rate, $\overline{ETR}_C^K$. This effect would be driven by a positive firm size-ETR_i^K gradient, where size is measured as firm revenue. The positive gradient arises because compliance and enforcement increase with size.[45] It also arises because the tax code in LMICs often leads to higher statutory tax burdens for larger firms (R. Kumar & James, 2022): using administrative tax data, Bachas et al. (2023) find a positive association between firm size and the statutory effective tax rate for corporate firms in 15 LMICs.[46]

6.2 Results on mechanisms: Tax capacity and race to the bottom

We investigate mechanisms relating trade to ETR, focusing on the tax capacity and 'race to the bottom' channels. In the race to bottom, international tax competition leads governments to reduce statutory corporate tax rates, which would reduce $\overline{ETR}_C^K$ (Section 2). We study both mechanisms in LMICs with the empirical strategies of Section 5.

Table 3 shows the OLS (Panel A) and the IV (Panel B) from equation 5. Consistent with race-to-bottom, column (1) shows that trade causes a decrease in the statutory CIT rate (significant at 10%).[47] The CIT rate is an imperfect proxy of firms' tax incentives as it ignores the tax base (Abbas & Klemm, 2013), but it can be measured in our full sample.

In line with the tax capacity mechanism, trade raises the corporate share of domestic output (μ_C), and reduces mixed income by an equivalent magnitude.[48] This is consistent with the conjecture that trade disproportionately benefits larger firms, which are more likely to be incorporated. Trade also raises $\overline{ETR}_C^K$ (column 6), consistent with the trade-induced corporate output accruing to firms whose ETR^K-size gradient is positive.

How is the additional income of the corporate sector allocated between capital and labor? Columns (4)-(5) show that the corporate sector rise is entirely driven by higher corporate profits, while the change in employee compensation growth is small and sta-

[45]See studies cited in Section 2. For example, Best et al. (2021) uncover a negative size-evasion gradient using randomized audit data on firms in Pakistan, finding also that firm-size is the most significant predictor of evasion. Models of tax compliance provide micro-foundations for the negative size-evasion gradient (including Kleven et al., 2016; Kopczuk & Slemrod, 2006).

[46]The gradient is positive everywhere except at the very top of the size-distribution, where it becomes negative. The gradient is driven by preferential tax treatments that increase with firm size and with characteristics that correlate with size such as total profits. The gradient can also reflect avoidance behavior, if larger firms are on average less able to take actions that reduce their legal tax liability.

[47]The outcome is the first-differenced tax rate (Romer & Romer, 2010). Table A4 shows results with the level of the CIT rate. We combine data from Végh and Vuletin (2015), Egger et al. (2019), Tax Foundation (link) and country-specific sources. A next step could be to study trade's impact on the more detailed statutory measures (Section 2). The downward trend in CIT rates in LMICs (supp. appendix) is related to, but does not fully capture, changes over time in the detailed statutory measures.

[48]The quality of data-sources used by national statistics offices can affect the measurement of mixed income in LMICs, but we find no impact of trade on countries' statistical capacity (World Bank link).

tistically insignificant.[49] This, in turn, causes trade to expand the capital share, both of national income and of the corporate sector (columns 7-8).[50]

The mechanism IV-results are robust to several checks (Table A4): using NDP weights; including controls; winsorizing the trade variable; and, estimating IVs separately based on each instrument. The CIT rate result remains less robust than the tax capacity results.

Figure A3 studies the same mechanism-outcomes but using the event-study design (Section 5.1). The trade liberalization events led to a decrease in the CIT rate and raised both corporate income (μ_C) and the effective corporate tax rate ($\overline{ETR}_C^K$). Some individual event-time coefficients are less precisely estimated, but the post-event dummies are jointly statistically significant for all outcomes. Although they are based on different identifying variation in trade, the event-study and IV results are therefore both consistent with the existence of the tax-capacity and race-to-bottom mechanisms in developing countries.

6.3 Firm-level investigation of tax capacity mechanism

The tax capacity mechanism is based on a firm level channel, combining a positive impact of trade on firm size with a positive firm size-ETR^K gradient. While the macro-results on μ_C and $\overline{ETR}_C^K$ in the previous subsection are consistent with it, in this subsection we directly investigate the tax capacity mechanism at the firm level.

We conduct the analysis in Rwanda between 2015 and 2017, where we leverage multiple administrative datasets to observe each formal Rwandan firm's exposure to trade and domestic tax payments. To our knowledge, there is limited firm level evidence in LMICs on how trade impacts a firm's domestic effective tax rate. Rwanda is an interesting setting as the corporate sector, starting from a comparatively low output share, has grown significantly since the 1990s, in tandem with a rise in trade openness and tax revenues.

We use corporate income tax returns to measure each firm's effective tax rate ETR_i^K as the ratio of corporate taxes paid divided by reported net profit. Net profit is revenue minus material, labor, operational, depreciation and financial costs. In Rwanda, this firm-level ETR_i^K varies due to firm characteristics (including revenue, our proxy for size), reduced rates and exemptions (Mascagni, Monkam and Nell, 2016). This ETR_i^K can also vary due to tax avoidance but, since the denominator is based on tax returns, it will not capture

[49] There is also a null effect of trade on households' operating surplus OS_{HH} (result not shown).

[50] This could occur due to an increase in markups. De Loecker and Eeckhout (2021) find that markups have risen in most regions over the past 40 years. De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) and Goldberg (2023) study the impact of trade on markups. Gupta (2023) and Atkin et al. (2015) find that markups increase with firm size, respectively in India and Pakistan. The rise in corporate profits and limited change in employee compensation may also arise if trade raises firms' labor market power (Felix, 2022). Finally, it may arise if trade benefits more capital-intensive production in developing countries, including through a reduction in CIT rates (Kaymak and Schott, 2023).

outright evasion.[51] The corporate ETR_i^K in Rwanda is everywhere positively associated with size (proxied by firm revenue), apart from in the top percentile (Bachas et al., 2023). Outside of the very top, an increase in firm i's size may cause ETR_i^K to rise.

We merge the CIT returns with customs data to record firms' direct exposure to trade. Following recent work (reviewed in Atkin and Khandelwal, 2020, Bernard & Moxnes, 2018), we measure a firm's total exposure to trade by also accounting for the firm's indirect exposure via its linkages to domestic suppliers that use traded goods in their production.[52] We merge administrative data that record transaction linkages between formal firms (details on data and sample in Appendix D.1). To measure a firm's total trade exposure in a network setting, we follow the methodology in Dhyne, Kikkawa, Mogstad, and Tintelnot (2021) that uses similar datasets to measure Belgian firms' exposure to trade. Specifically, we define firm i's total foreign input share as the share of inputs that it directly imports (s_{Fi}), plus the share of inputs that it buys from its domestic suppliers l (s_{li}), multiplied by the total import shares of those firms:

$$s_i^{Total} = s_{Fi} + \sum_{l \in V_i} s_{li} \cdot [s_{Fl} + \sum_{r \in V_l} s_{rl} \cdot (s_{Fr} + ...)] \tag{8}$$

where V_i is the set of domestic suppliers of firm i, and V_l is the set of domestic suppliers of firm l. The denominator of the input shares is the sum of imports and purchases from other firms. We limit the recursive calculation in (8) to inputs from a firm's immediate suppliers l and the suppliers to their suppliers r (adding more levels only marginally raises s_i^{Total}).[53] Inspecting s_i^{Total} and s_{Fi} reveals that while just under 30% of Rwandan formal firms import directly, 93% rely on trade directly or indirectly through suppliers which use foreign inputs in their production. Most firms are therefore dependent on foreign trade, but only a limited number show that dependence through the direct foreign inputs observed in customs data. The median total foreign input share is 48%.

We estimate regressions in the sample of corporate firms of the form:

$$ETR_{it}^K = \mu \cdot s_{it}^{Total} + \Theta \cdot X_{it} + \pi_t + \pi_i + \epsilon_{it} \tag{9}$$

[51]For this reason, $\overline{ETR}_C^K$ measured in national accounts differs from the (appropriately weighted) corporate ETR_i^K measured in tax returns. They also differ because of conceptual differences in the measurement of profits: see the supplementary appendix for a detailed discussion.

[52]Recent papers study domestic linkages in LMICs and their role in propagating trade shocks (including Almunia, Hjort, et al., 2023; Fieler, Eslava, & Xu, 2018; Javorcik, 2004).

[53]We focus on firms' exposure to imports through their supplier network; we find qualitatively similar results when we study firms' exposure to exports through their client network (results available).

where ETR_{it}^{K} and s_{it}^{Total} are the corporate effective tax rate and total trade exposure of firm i in year t, and π_t and π_i are year and firm fixed effects. X_{it} includes number of employees and number of clients and suppliers, and ϵ_{it} is clustered at the firm level.

In Table 4, the OLS estimation of (9) shows that a within-firm increase in trade exposure is associated with a higher corporate effective tax rate. This result holds with only year fixed effects π_t (column 1); with industry-geography fixed effects (column 2); with firm controls X_{it} (column 3); with firm fixed effects π_i (column 4).

In Table 4, column (5), we implement an IV that generates firm-level variation in trade exposure using the shift-share design from Hummels et al. (2014). The identifying variation is trade shocks from changes in the world export supply of specific country-product combinations in which a Rwandan firm had a previous import relationship. Specifically, the direct import trade shock for firm i in year t is:

$$logM_{it}^{D} = log \sum_{a,c} s_{ic,t-1}^{a,M} \cdot WES_{a,c,t} \tag{10}$$

where $s_{ic,t-1}^{a,M}$ is the share of imports of firm i in year $t-1$ that falls on product a from country c, and $WES_{a,c,t}$ is the world export supply (excluding sales to Rwanda) of country c for product a. Product a is measured at the detailed six-digit HS level. Rwandan firms import over 3,510 distinct products from 174 different countries of origin.

The shocks to Rwandan firms' trading environment are time-varying and specific to each partner-country $\times$ product being traded. They capture transportation costs and worldwide shocks to export supply for the relevant country $\times$ product, and contain granular variation across products and countries. The identification strategy rests on the joint hypotheses that these shocks are plausibly exogenous to Rwandan firms' trading environment and that they create varied impacts across firms because Rwandan importers have few imported inputs in common. Indeed, the customs data shows that the median number of unique importing firms in a given HS6 product $\times$ country and time period is 1; the 95^{th} percentile is 3. Hence, if only one Rwandan firm imports metal cored wires from Turkey, an idiosyncratic shock to Turkey's global export supply of those wires will affect just one firm in Rwanda. Note also that, to construct the trade shocks, we rely on prior information about importers' sourcing patterns, which removes concerns over contemporaneous shocks affecting both the choice of imported goods and firm outcomes.

We build the trade shocks for all firms. In turn, the 1^{st}-stage instruments are the firm's own trade shocks, as well as the trade shocks to its suppliers and to the suppliers of its suppliers. Specifically, the 1^{st}-stage regression is:

$$s_{it}^{Total} = \beta_1 \cdot logM_{it}^{D} + \beta_2 \cdot logM_{it}^{S} + \beta_3 \cdot logM_{it}^{SS} + \kappa_t + \kappa_i + \epsilon_{it} \tag{11}$$

where $logM_{it}^{D}$, $logM_{it}^{S}$, and $logM_{it}^{SS}$ are the trade shocks to firm i, to firm i's suppliers, and to the suppliers of firm i's suppliers. We construct weighted averages of trade shocks in the supplier network using the recursive formulation in (8) (details in Appendix D.1).

We find that both direct trade shocks to a firm's own imports and indirect shocks to a firm's network of suppliers cause significant changes to the firm's total exposure s_{it}^{Total}, generating a strong 1^{st}-stage (Kleibergen-Paap F-statistic of 18.17).

The IV specification shows that trade causes an increase in the individual firm's effective tax rate on capital (column 5). In Panel B, the IV reveals that trade causes an increase in firm size (proxied by revenue). Panel C shows a positive OLS association between firm size and ETR_{i}^{K} (we cannot use the IV in this panel due to the exclusion restriction).

In Appendix D.1, we find that the main results are robust to controlling for trade shocks to firm i's potential suppliers (firms that operate in the same industry and geographical area as i's current suppliers but are not currently supplying to i) and horizontal suppliers (firms that are suppliers to firm i's current clients). These results provide additional support for the exogeneity assumption.[54]

Though the analysis in Rwanda is based within a single country over a limited time range, it supplements the macro-level results in two ways. First, it provides firm-level identified evidence that trade exerts a positive impact on effective corporate taxation in a developing country, which complements the country-level results in LMICs. Second, by showing that trade increases firm size and that size is positively associated with ETR^{K}, it supports the tax capacity mechanism interpretation that trade's impact on ETR^{K} is mediated by a positive size-ETR^{K} gradient.

Discussion: Links to trade-formality literature At the firm, sector and country level, we find positive effects of trade on outcomes related to formalization. Recent studies focused on the number of formal versus informal firms or formal versus informal workers, and found mixed evidence that trade increases formality by these measures (reviews in Engel & Kokas, 2021; Ulyssea, 2020).[55] One way to reconcile our results with these studies is to note that our focus is on the share of output produced in larger and formal firms: output expansion in these firms may occur without changes to the number of formal or informal firms, and does not imply an increase in the number of formal workers, since informal workers may work in formal firms and contribute to their output (Ulyssea, 2018). In 6.4, we also show that openness' impact on our formal-outcomes depends on the nature of the trade shock, consistent with recent theoretical work in trade (Dix-Carneiro et al., 2021).

[54]In an extension, we find that increased *output* exposure to imports through the client network has positive effects on ETR^{K}, though this average effect could mask heterogeneity across firms.

[55]Goldberg and Pavcnik (2003), Bosch et al. (2012), Cruces et al. (2018), Dix-Carneiro and Kovak (2019).

6.4 Sources of heterogeneity in trade's pro-tax impacts

We return to the country-level IV (equation 5) to study sources of heterogeneity in trade's pro-tax impacts on the tax capacity mechanism and ETR.

Heterogeneity: Domestic enforcement reforms Over our sample period, LMICs have implemented tax enforcement policies. A challenge for the mechanism interpretation is that trade, potentially due to revenue concerns, may have prompted governments to implement these policies that increase ETR^K. To investigate this, we measure the year of adoption (if any) in LMICs of four policies that increase domestic tax enforcement: (i) large taxpayer unit; (ii) organizational integration of customs and domestic tax authorities; (iii) VAT; (iv) international accounting standards (IAS).[56] We estimate heterogeneous IV effects by including an interaction term between trade and the policy adoption variable in (5).[57] Table A5 shows a positive effect of trade on ETR^K without these policies, though the effect is larger following their adoption. Trade has a similar impact on the corporate income-share (μ_C) with and without the enforcement policies, but trade's positive impact on $\overline{ETR}_C^K$ is significantly amplified when enforcement policies are in place.[58] That is, the trade-induced expansion of the corporate sector seems to occur regardless of enforcement policies, but the extent to which the additional corporate output translates into higher effective corporate taxation is reinforced when such policies have been enacted.[59]

Governments in LMICs may have sought to raise domestic revenue, possibly in response to openness, through other channels apart from these specific enforcement policies. We investigate this in Table A6, finding that trade's positive impact on the tax capacity mechanism and ETR^K hold outside of periods of significant revenue loss, when defined in various ways including the episodes of trade revenue loss in Cagé and Gadenne (2018). Thus, trade's pro-tax impacts appear to be broadly present in the globalization process in LMICs, and do not hinge on government's revenue need or enforcement investments.

Heterogeneity: Nature of trade shock Trade theories highlight that the impacts of trade on formality-related outcomes depend on the nature of the trade shock. In Appendix D.2, we use both instruments and equation (5) in LMICs to investigate if the ETR and

[56]The enforcement focus on large firms increases collection (Almunia & Lopez-Rodriguez, 2018; M. C. Basri, Felix, Hanna, & Olken, 2021). The customs-tax unification improves domestic audit capacity (IMF, 2022). The VAT creates information trails (Almunia, Henning, Knebelmann, Nakyambadde, & Tian, 2023; Waseem, 2020). IAS deepen accounting requirements for tax reporting (Barth et al., 2008).

[57]The timing of adoption for each reform is endogenous; however, our focus is on the trade coefficients with and without these reforms in place, which are identified (Bun and Harrison, 2019).

[58]Only the VAT was adopted in all liberalizing countries by the time of the events studied in Section 5.1.

[59]Intuitively, the enforcement policies all disproportionately raise enforcement on larger firms, thereby further increasing the slope of the ETR^K-size gradient inside the corporate sector. Whether these enforcement policies are themselves driven by globalization is a topic for future research.

mechanism impacts differ along two dimensions (Dix-Carneiro et al., 2021): imports versus exports; and, trade in intermediate versus final goods and services (G-S). We find that *exports* increase ETR^K and the corporate income-share (μ_C), while *imports* decrease both outcomes. These results are consistent with 'Melitz-type' demand effects, whereby increased exports represent a pure positive demand shock for export-oriented firms, while increased imports may constitute a negative demand shock for domestic firms, disproportionately affecting larger ones. In additional IV regressions, trade in *intermediate G-S* increases ETR^K and μ_C, while trade in *final G-S* decreases both outcomes.[60] Results are similar for $\overline{ETR}_C^K$. These results are consistent with the increased availability of intermediate goods benefiting larger firms; by contrast, the increased availability of final goods may constitute a negative domestic demand shock, particularly for larger firms. These results suggest trade's pro-tax impacts depend on the nature of the trade shock.

Heterogeneity: Developing vs developed countries We investigate if trade's impacts on mechanisms and ETR differ across development levels, by expanding our sample to include HICs. We expect that the tax capacity mechanism is less likely to operate in HICs, where enforcement constraints on effective taxation are less binding and the corporate sector's size has been stable since the 1970s (Figure 1). On the other hand, the race-to-bottom is likely to be active in HICs, given previous research (Section 2). Table A7 reports heterogeneous IV effects by augmenting (5) with an interaction between trade and a dummy for high-income countries.[61] Trade only raises ETR^K in LMICs, but raises ETR^L everywhere. The negative race-to-bottom effect on the CIT rate is much stronger in HICs than in LMICs. The positive impact of trade on tax capacity outcomes (μ_C, $\overline{ETR}_C^K$) is limited to LMICs, with null effects in HICs.[62] These results suggest countervailing mechanisms that differ by development level, through which trade may have contributed to the diverging trends in ETR^K between HICs and LMICs documented in Figure 2.

We study additional country characteristics in the supplementary appendix. We find that trade's negative impact on the CIT rate is larger in countries that are smaller and with fewer capital restrictions – two settings where capital flight concerns are more pronounced (Hines, 2006). Mirroring this result, trade's positive impact on ETR^K occurs in larger countries and with more capital restrictions. The tax capacity and race-to-bottom mechanisms therefore appear to occur simultaneously: countries with larger markets and lower capital mobility reap more of the tax-capacity benefits of trade.

[60] Which suggests that imports of intermediate (final) G-S increases (decreases) ETR_K and μ_C.

[61] We note these results should be interpreted with caution, given the econometric challenges of estimating IV effects with multiple endogenous regressors (Andrews, Stock, & Sun, 2019).

[62] The IV-coefficients for developing countries differ qualitatively between Table A7 and Tables 1-3. This is because the two instruments' strength changes in the 1^{st}-stage regression (Table A2).

Extension: Capital openness We focused on trade openness but another relevant dimension of globalization is capital openness (Ilzetzki, Reinhart, & Rogoff, 2019; Van Patten, 2022). Due to differences in reporting requirements, data on capital openness is not as available and comparable as trade data, and finding credible exogenous variation for capital openness is challenging. Notwithstanding, we try to investigate the impacts of capital openness in Appendix E. We rely on capital inflow liberalization events for 25 developing countries from Chari, Henry, and Sasson (2012), which capture the first time that foreign investment in the domestic stock market is allowed. Employing the event-study design of Section 5.1, we find that capital liberalization events raise capital openness and positively impact ETR^K and the tax capacity mechanism (μ_C, $\overline{ETR}_C^K$). The pro-tax impacts of globalization in LMICs may be robust to using capital instead of trade openness.

7 Conclusion

This paper provides evidence on long-run trends in capital taxation and causal effects of globalization. Based on a new macro-historical database, we document that effective capital tax rates have increased in developing countries in the post-1990 era of hyper-globalization. By expanding the share of economic activity in incorporated and larger firms, we find that trade improves the effective collection of taxes, particularly corporate income taxes. We provide evidence on this tax capacity effect across multiple research designs and at the country, corporate sector and firm-level. Despite a simultaneous negative effect on corporate statutory tax rates induced by international tax competition, the positive tax capacity effect is sufficiently large that trade increases the effective tax rate on capital and overall government revenues (% of GDP) in developing countries.

Due to limited data, the revenue consequences of globalization in developing countries had not been systematically investigated and policy concerns over revenue losses have persisted in a context of uncertainty surrounding the future of globalization (Goldberg & Reed, 2023). We find that globalization has pro-tax impacts that have supported the effective taxation of capital and overall revenue collection in many countries.

Our results show that openness increased the share of market income going to corporations, profits, and capital. Simultaneously, trade's pro-tax impacts mean that developing countries raised more taxes from capital. As a result, openness is likely to have widened pre-tax income inequality, but its effect on post-tax income inequality is more nuanced. We adopted a macro focus, but a next step could be to combine the ETRs with individual-level estimates of the progressivity of capital (and labor) taxes. This would allow a comparison of the distributional effects of globalization on pre versus post-tax income, raising empirical questions for future research (Goldberg, 2023; Pavcnik, 2017).

References

Abbas, A., & Klemm, A. (2013). A partial race to the bottom: Corporate tax developments in developing economies. *International Tax and Public Finance*, *20*.

Abramovsky, L., Klemm, A., & Phillips, D. (2014). Corporate tax in developing countries: Current trends and design issues. *Fiscal Studies*, *35*(4), 559–588.

Adam, A., Kammas, P., & Rodriguez, A. L. (2013). The effect of globalization on capital taxation. *Journal of Macroeconomics*, *35*.

Alesina, A., & Wacziarg, R. (1998). Openness, country size and government. *Journal of Public Economics*, *69*.

Alfaro-Ureña, A., Manelici, I., & Vasquez, J. (2022). The effects of joining multinational supply chains. *Quarterly Journal of Economics*, *137*(3).

Almunia, M., Henning, D., Knebelmann, J., Nakyambadde, D., & Tian, L. (2023). Leveraging trading networks to improve tax compliance. *Working Paper*.

Almunia, M., Hjort, J., Knebelmann, J., & Tian, L. (2023). Strategic or confused firms? Evidence from missing transactions in Uganda. *Review of Economics and Statistics*.

Almunia, M., & Lopez-Rodriguez, D. (2018). Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy*, *10*(1).

Andrews, I., Stock, J., & Sun, L. (2019). Weak instruments in IV regression: Theory and evidence. *Annual Review of Economics*, *11*.

Atkin, D., & Khandelwal, A. (2020). How distortions alter the impacts of international trade in developing countries. *Annual Review of Economics*, *12*.

Bachas, P., Fattal, R., & Jensen, A. (2019). Size-dependent tax enforcement and compliance. *Journal of Development Economics*, *140*, 203–222.

Bachas, P., Brockmeyer, A., Dom, R., & Semelet, C. (2023). Effective tax rates and firm size. *Working Paper*.

Basri, C., Felix, M., Hanna, R., & Olken, B. (2019). Tax administration vs. tax rates. *NBER Working Paper*, (26150).

Basri, M. C., Felix, M., Hanna, R., & Olken, B. (2021). Tax administration versus tax rates. *American Economic Review*, *111*(12).

Baunsgaard, T., & Keen, M. (2009). Tax revenue and (or?) trade liberalization. *Journal of Public Economics*, *94*(9).

Benzarti, Y., & Tazhitdinova, A. (2021). Do value-added taxes affect international trade flows? *American Economic Journal: Economic Policy*, *13*(4), 469–89.

Bergeron, A., Tourek, G., & Weigel, J. (2024). The state capacity ceiling on tax rates. *Econometrica*, *Forthcoming*.

Bernard, A., Jensen, B., Redding, S., & Schott, P. (2007). Firms in international trade. *Journal of Economic Perspectives*, *21*(3).

Besley, T., & Persson, T. (2014). Why do developing countries tax so little? *Journal of Economic Perspectives*, *28*(4).

Best, M., Shah, J., & Waseem, M. (2021). Detection without deterrence. *Working Paper*.

Bilicka, K. A. (2019). Comparing UK tax returns of foreign multinationals to matched domestic firms. *American Economic Review*, *109*(8).

Blanchet, T., Chancel, L., Flores, I., & Morgan, M. (2021). *Distributional national accounts*.

Borusyak, K., Jaravel, X., & Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *Working Paper*.

Brandt, L., Biesebroeck, J. V., Wang, L., & Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review*, *107*(9).

Buettner, T., & Madzharova, B. (2018). WTO membership and the shift to consumption taxes. *World Development*, *108*(1).

Cagé, J., & Gadenne, L. (2018). Tax revenues and the fiscal cost of trade liberalization, 1792-2006. *Explorations in Economic History*, *70*.

Carey, D., & Rabesona, J. (2004). Tax ratios on labor and capital income and on consumption. In P. B. Sørensen (Ed.), *Measuring the tax burden on capital and labor*.

Carillo, P., Pomeranz, D., & Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics*, *9*(2).

Chari, A., Henry, P., & Sasson, D. (2012). Capital market integration and wages. *American Economic Journal: Macroeconomics*, *4*(2).

Clausing, K. (2007). Corporate tax revenues in OECD countries. *International Tax and Public Finance*, *14*.

Cogneau, D., Dupraz, Y., Knebelmann, J., & Mesplé-Somps, S. (2021). Taxation in Africa from colonial times to present. *Working Paper*.

De Loecker, J., & Eeckhout, J. (2021). Global market power. *Working Paper*.

De Loecker, J., Goldberg, P., Khandelwal, A., & Pavcnik, N. (2016). Prices, markups and trade reform. *Econometrica*, *84*(2).

Devereux, M., & Griffith, R. (1999). The taxation of discrete investment choices. *Institute for Fiscal Studies Working Paper Series*, *No. W98/16*.

Devereux, M., Griffith, R., & Klemm, A. (2002). Corporate income tax: Reforms and tax competition. *Economic Policy*.

Dhyne, E., Kikkawa, A., Mogstad, M., & Tintelnot, F. (2021). Trade and domestic production networks. *Review of Economic Studies*, *88*(2).

Dix-Carneiro, R., Goldberg, P., Meghir, C., & Ulyssea, G. (2021). Trade and informality in the presence of labor market frictions and regulations. *NBER Working Paper*.

Dyreng, S. D., Hanlon, M., Maydew, E. L., & Thornock, J. R. (2017). Changes in corporate effective tax rates over 25 years. *Journal of Financial Economics*, *124*(3).

Egger, P., Nigai, S., & Strecker, N. (2019). The taxing deed of globalization. *American Economic Review*, *109*(2).

Engel, J., & Kokas, D. (2021). The distributional impacts of trade. *Working Paper*.

Felix, M. (2022). Trade, labor market concentration, and wages. *Working Paper*.

Fieler, A., Eslava, M., & Xu, D. (2018). Trade, quality upgrading and input linkages: Theory and evidence from Colombia. *American Economic Review*, *108*(1).

Fisman, R., & Wei, S.-J. (2004). Tax rates and tax evasion: Evidence from "missing imports" in China. *Journal of Political Economy*, *112*(2).

Fisunoglu, A., Kang, K., Arbetman-Rabinowitz, M., & Kugler, J. (2011). Relative political capacity dataset. Version 2.4. Harvard Dataverse.

Garcia-Bernardo, J., Janský, P., & Tørsløv, T. (2022). Decomposing multinational corporations declining effective tax rates. *IMF Economic Review*, *70*(2).

Goldberg, P. (2023). *The unequal effects of globalization*. MIT Press.

Goldberg, P., & Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature*, *45*(1).

Goldberg, P., & Pavcnik, N. (2016). The effects of trade policy. *NBER Working Paper*.

Goldberg, P., & Reed, T. (2023). Is the global economy deglobalizing? *Brookings Paper on Economic Activity*.

Gollin, D. (2002). Getting income shares right. *Journal of Political Economy*, *110*(2).

Griffith, R., & Miller, H. (2014). Taxable corporate profits. *Fiscal Studies*, *35*.

Guerriero, M. (2019). Labor share of income around the world. *Working Paper*.

Gupta, A. (2023). Demand for quality, variable markups and misallocation: Evidence from India. *Working Paper*.

Hines, J. (2006). Corporate taxation and international competition. *Working Paper*.

Hummels, D., Jørgensen, R., Munch, J., & Xiang, C. (2014). The wage effects of offshoring. *American Economic Review*, *104*(6).

ILO. (2019). *The global labour income share and distribution.* Intl. Labor Organization.

Ilzetzki, E., Reinhart, C., & Rogoff, K. (2019). Exchange arrangements entering the twenty-first century. *Quarterly Journal of Economics*, *134*(2).

IMF. (2014). Spillovers in international corporate taxation. *IMF Staff Report*.

Javorcik, B. (2004). Does foreign direct investment increase the productivity of domestic firms? *American Economic Review*, *94*(3).

Javorcik, B., & Narciso, G. (2017). WTO accession and tariff evasion. *Journal of Development Economics*, *125*, 59–71.

Jensen, A. (2022). Employment structure and the rise of the modern tax system. *American Economic Review*, *112*(1).

Karabarbounis, L., & Neiman, B. (2014). The global decline of the labor share. *Quarterly Journal of Economics*, *129*(1).

Kleven, H., Kreiner, C., & Saez, E. (2016). Why can modern governments tax so much? *Economica*, *83*(330).

Kleven, H., Knudsen, M., Kreiner, C., Pedersen, S., & Saez, E. (2011). Unable or unwilling to cheat? Evidence from a tax audit experiment. *Econometrica*, *79*(3).

Kopczuk, W., & Slemrod, J. (2006). Putting firms into optimal tax theory. *American Economic Review*, *96*(2).

Kostarakos, I., & Varthalitis, P. (2020). Effective tax rates in the EU. *Working Paper*.

Kugler, M., & Verhoogen, E. (2009). Plants and imported inputs: New facts and an interpretation. *American Economic Review*, *99*(12).

Kumar, M. M. S., & Quinn, M. D. P. (2012). *Globalization and corporate taxation*. International Monetary Fund.

Kumar, R., & James, S. (2022). World Bank global marginal effective tax rate database report. *World Bank Technical Report*.

Londoño-Vélez, J., & Tortarolo, D. (2022). Revealing 21% of GDP in hidden assets: Evidence from Argentina's tax amnesties. *UNU WIDER Working Paper*, (103).

McCaig, B., & Pavcnik, N. (2018). Export markets and labor allocation in a low-income country. *American Economic Review*, *108*(7).

McDaniel, C. (2007). Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003. *Working Paper*.

McDaniel, C. (2020). Average tax rates on consumption, investment, labor and capital in the OECD 1995-2018. *Working Paper*.

Méndez, E., & Van Patten, D. (2022). Multinationals, monopsony, and local development. *Econometrica*, *90*(6).

Mendoza, E. G., Razin, A., & Tesar, L. L. (1994). Effective tax rates in macroeconomics. *Journal of Monetary Economics*, *34*(3).

Mrázová, M., & Neary, P. (2018). Selection effects with heterogeneous firms. *Journal of European Economic Association*, *17*.

Naritomi, J. (2019). Consumers as tax auditors. *American Economic Review*, *109*(9).

Parenti, M. (2018). Large and small firms in a global market: David vs. Goliath. *Journal of International Economics*, *110*.

Pavcnik, N. (2017). The impact of trade on inequality in developing countries. *Proceedings of Jackson Hole Economic Symposium*.

Piketty, T., Saez, E., & Zucman, G. (2018). Distributional national accounts. *Quarterly Journal of Economics*, *133*(2).

Piketty, T., Yang, L., & Zucman, G. (2019). Capital accumulation, private property, and rising inequality in China. *American Economic Review*, *109*(7).

Pomeranz, D. (2015). No taxation without information. *American Economic Review*.

Quinn, D. (1997). The correlates of change in international financial regulation. *American Political Science Review*, *91*(3), 531–551.

Rodrik, D. (1998). Why do more open economies have bigger governments? *Journal of Political Economy*, *106*(5).

Sequeira, S. (2016). Corruption, trade costs, and gains from tariff liberalization: Evidence from Southern Africa. *American Economic Review*, *106*.

Slemrod, J. (2004). Are corporate tax rates, or countries, converging? *Journal of Public Economics*, *88*(6), 1169–1186.

Slemrod, J., & Velayudhan, T. (2018). Do firms remit at least 85% of tax everywhere? New evidence from India. *Journal of Tax Administration*, *4*(1).

Sørensen, P. (2007). Can capital income taxes survive? *CESifo Economic Studies*, *53*.

Steinmüller, E., Thunecke, G. U., & Wamser, G. (2019). Corporate income taxes around the world. *International Tax and Public Finance*, *26*, 418–456.

Ulyssea, G. (2018). Firms, informality, and development. *American Economic Review*.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics*, *12*.

Van Patten, D. (2022). The dismantling of capital controls after Bretton Woods and Latin American productivity. *AEA Papers and Proceedings*, *112*.

Wacziarg, R., & Wallack, J. (2004). Trade liberalization and intersectoral labor movements. *Journal of International Economics*, *64*.

Wacziarg, R., & Welch, K. H. (2008). Trade liberalization and growth: New evidence. *World Bank Economic Review*, *22*(2), 187–231.

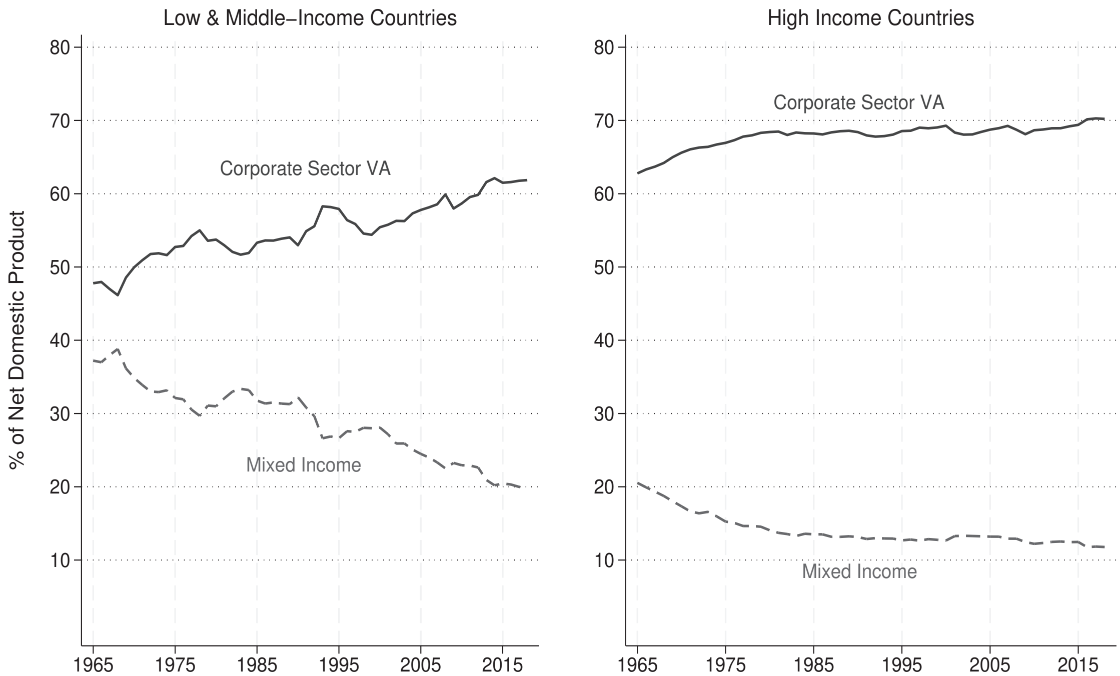
Waseem, M. (2020). The role of withholding in the self-enforcement of a value-added tax. *Review of Economics and Statistics*, *104*(2).

Wier, L. (2020). Tax-motivated transfer mispricing. *Journal of Public Economics*.

World Bank. (2008). *Public finance in China* (J. Lou & S. Wang, Eds.). WB Group.

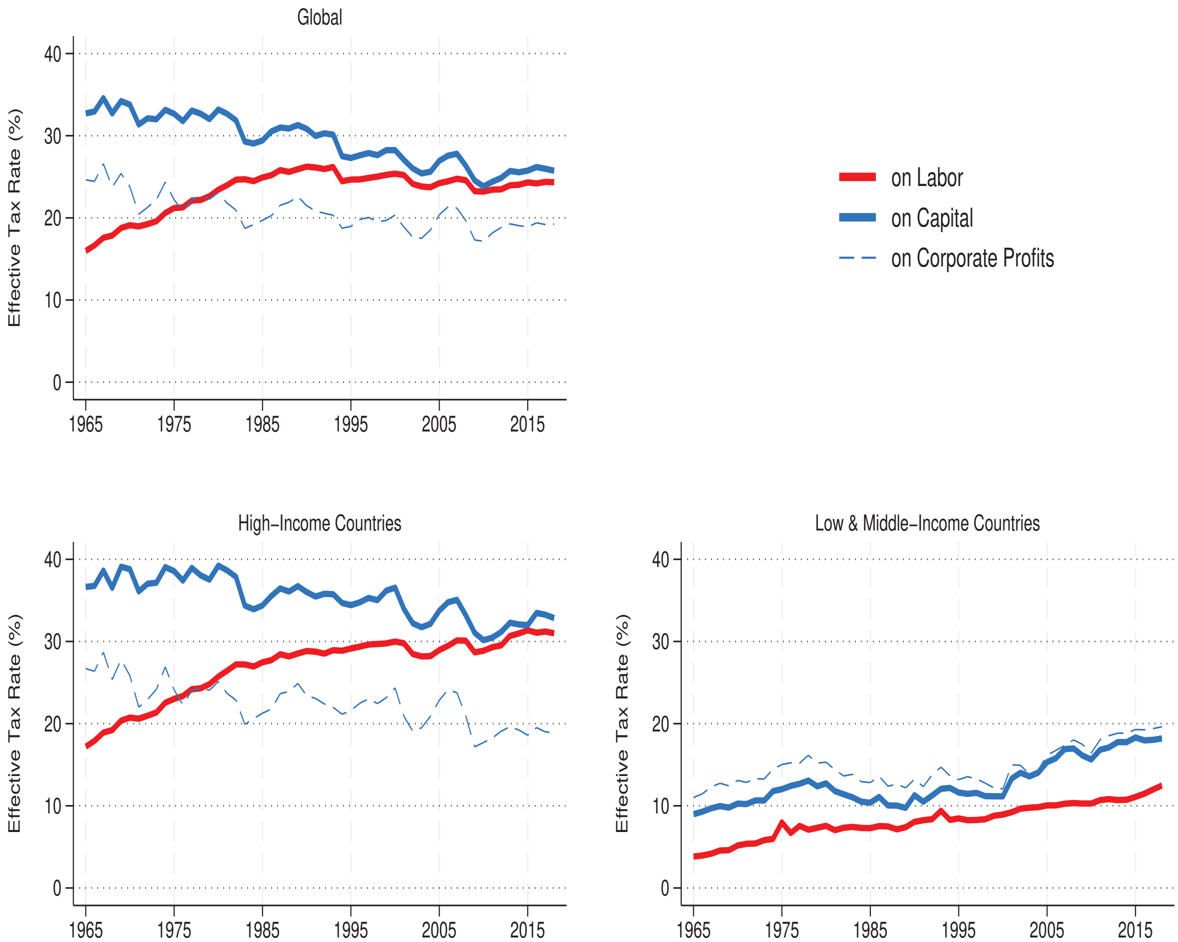
World Bank. (2020). *The African continental free trade area: Economic effects*. WB Group.

Figure 1: Corporate Sector Income and Mixed Income (1965-2018)



Notes: These panels plot the time series of corporate sector income and of mixed income between 1965 and 2018 by level of development, from national accounts statistics. Both outcomes are expressed as a percent of net domestic product and weighted by countries' net domestic product in constant 2010 USD. Corporate income is the sum of corporate profits and corporate employee compensation. Mixed income accounts for income from self-employed and unincorporated businesses. The left panel show the results for low and middle-income countries (N=117), and the right panel show the results for high income countries (N=37), based on the World Bank income classification in 2018.

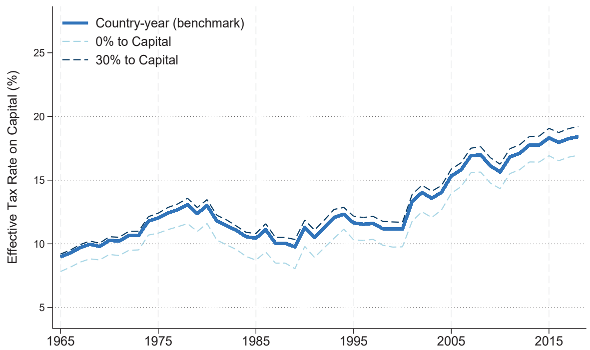
Figure 2: Effective Taxation of Capital and Labor (1965-2018)



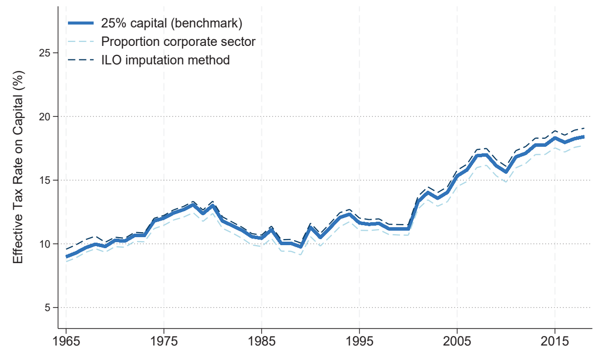
Notes: This figure plots the time series of average effective tax rates on labor (red) and capital (blue), as well as the average effective tax rate on corporate profits (blue dashed line). The top-left panel corresponds to the global average, weighting country-year observations by their share in that year's total factor income, in constant 2019 USD (N=154). The bottom-left panel shows the results for high-income countries (N=37), and the bottom-right panel for low- and middle-income countries (N=117). Income classification is based on the World Bank income groups in 2018. The dataset is composed of two quasi-balanced panels. The first covers the years 1965-1993 and excludes communist regimes. It accounts for 85-90% of world GDP during those years. The second covers 1994-2018 and integrates former communist countries, in particular China and Russia, and accounts for 97-98% of world GDP. This figure is discussed in Section 4.1.

Figure 3: Robustness of Effective Capital Taxation in Developing Countries

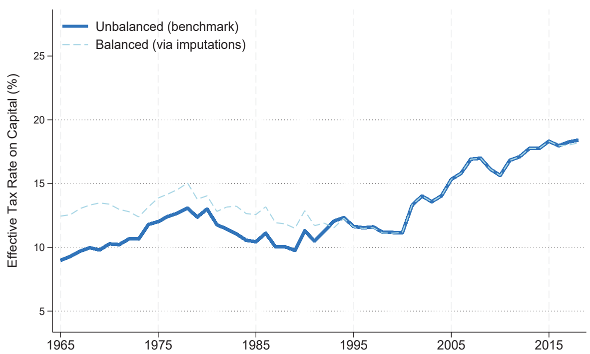
(a) PIT revenue allocation



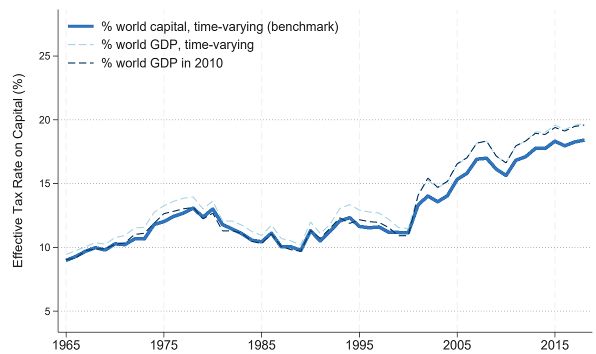
(b) Mixed-income allocation



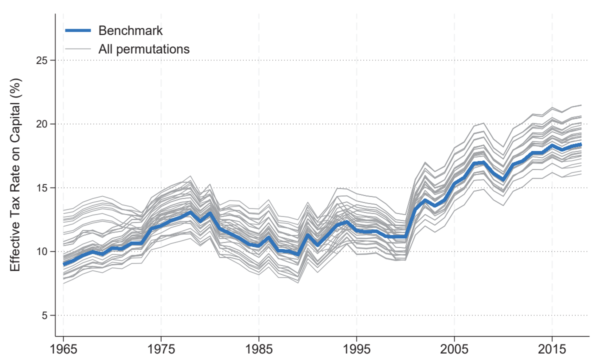
(c) Panel balancedness



(d) Weights for aggregation



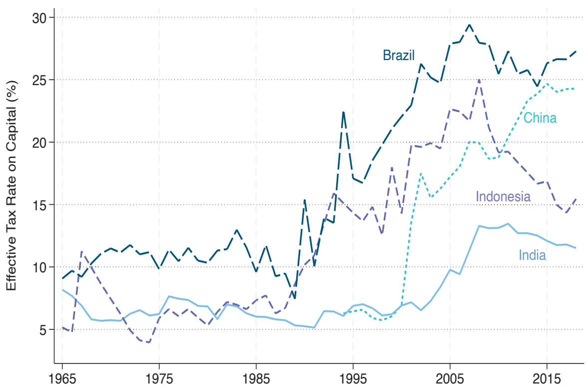
(e) All 54 combinations



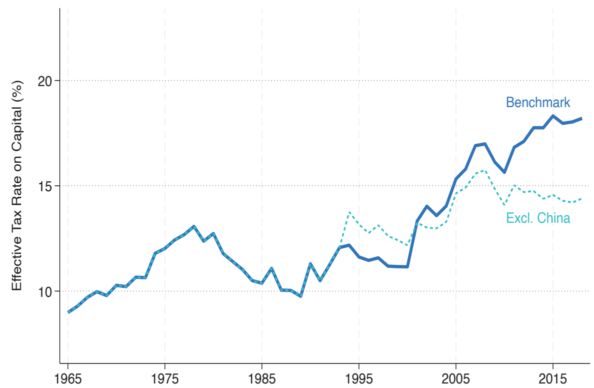
Notes: These panels show trends in the effective tax rate on capital in the 117 developing countries in our sample. The panels vary our four key methodological choices: the allocation of personal income tax revenue to capital vs labor (panel a); the allocation of mixed income to capital vs labor (panel b); presenting results for an unbalanced panel of countries vs a balanced panel via imputations (panel c); and, the use of weights to aggregate individual countries' time-series (panel d). Panel (e) shows all 54 possible combinations that can be constructed by combining these choices. In all panels, the blue line corresponds to our benchmark series. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. This figure is discussed in Section 4.2.

Figure 4: Heterogeneity of Effective Capital Taxation in Developing Countries

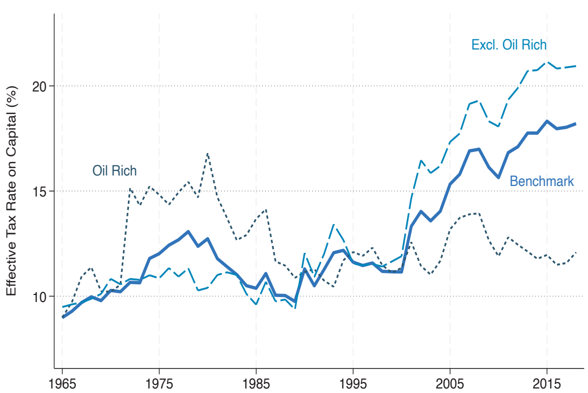
(a) Four largest countries



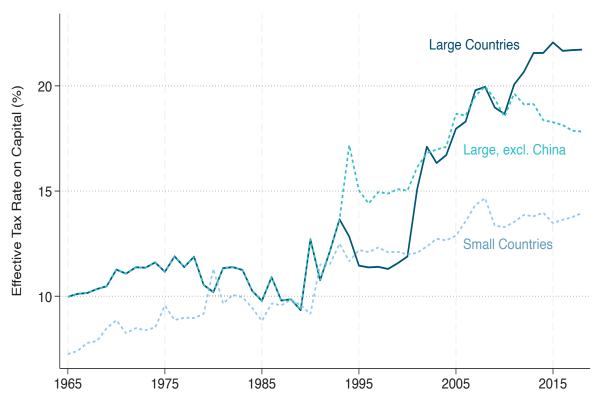
(b) Excluding China



(c) Oil Rich Countries



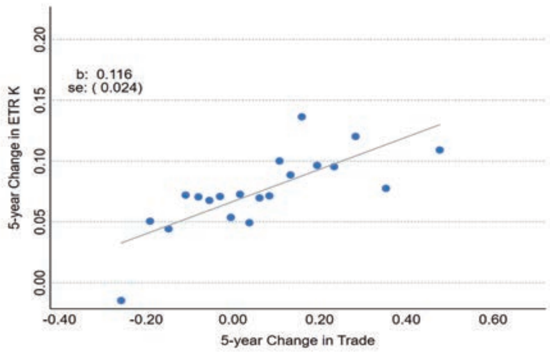
(d) Large vs small countries (non-oil rich)



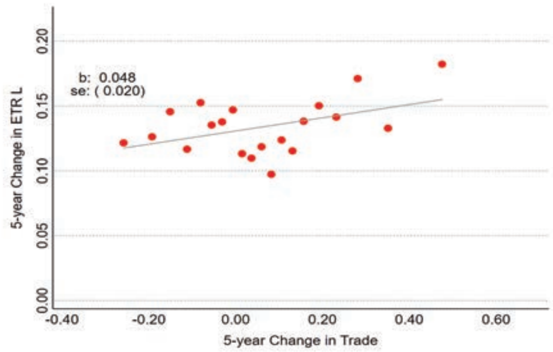
Notes: These panels show the evolution of the effective tax rate on capital, ETR_K, for major developing countries and sub-samples of developing countries. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. Panel (a) plots the ETR_K series for the four largest (most populous) developing countries: Brazil, China, India, Indonesia. Panel (b) compares our benchmark series to the series that excludes China. Panel (c) plots the ETR_K series for a sample of oil-rich countries (countries with more than 7% of GDP from oil in 2018), and the benchmark ETR_K series without these countries. Within the sample of non-oil rich developing countries, panel (d) compares large countries to small countries. Large countries are defined as having a population above 40 million in 2018. This figure is discussed in Section 4.3.

Figure 5: Within-Country Associations between Effective Tax Rates and Trade

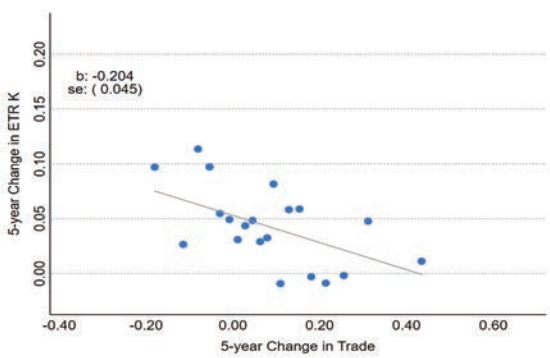
(a) ETR_K: All countries



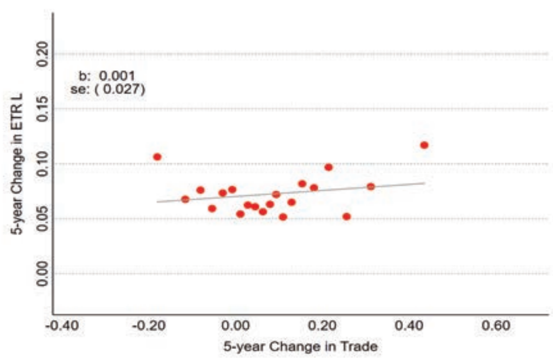
(b) ETR_L: All countries



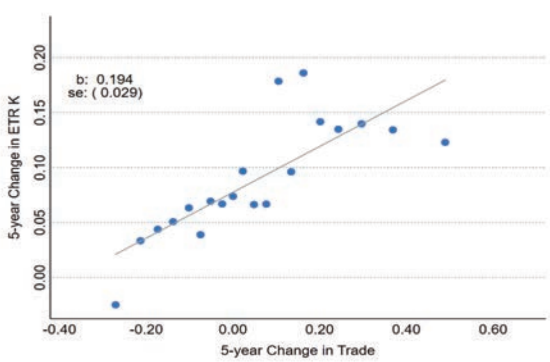
(c) ETR_K: High-income



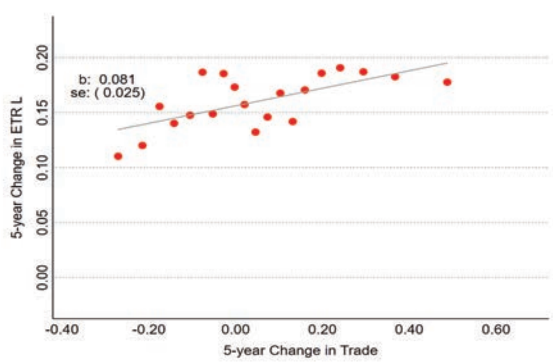
(d) ETR_L: High-income



(e) ETR_K: Low & middle-income

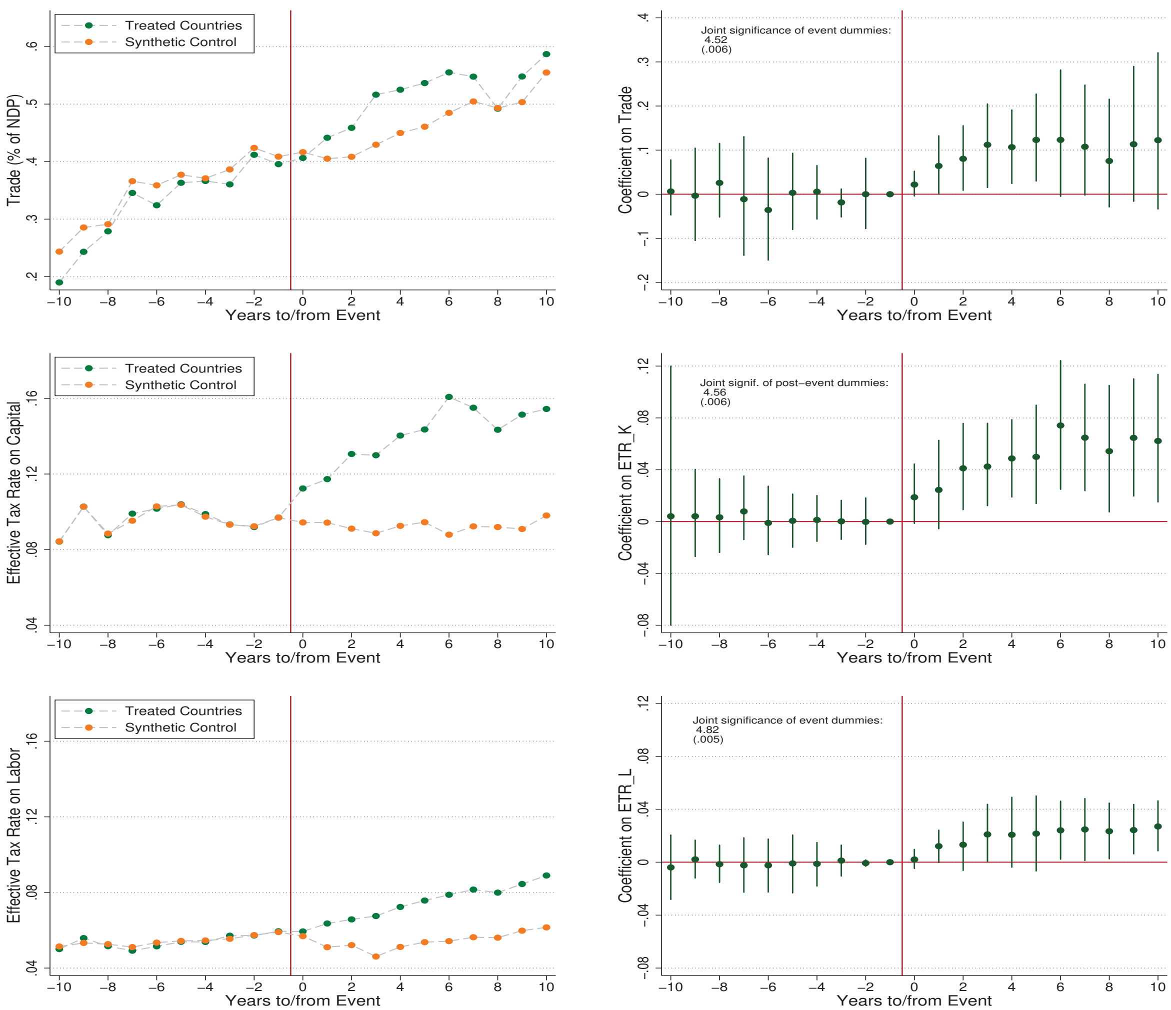


(f) ETR_L: Low & middle-income



Notes: These panels shows the association between trade and effective tax rates. The outcome is the effective tax rate on capital, ETR_K, and on labor, ETR_L, in the left-side and right-side panels, respectively. The top panels show the associations in all countries; the middle panels show the associations in high-income countries (based on World Bank income classification in 2018); the bottom panels show the associations in low and middle-income countries. Trade is measured as the sum of import and exports as a share of net domestic product. Both the x-axis and y-axis are measured as within-country percent changes over 5 years. Each graph shows binned scatter plots of each outcome against trade, after residualizing all variables against year-fixed effects. Each dot corresponds to a ventile (20 equal-sized bins) of the residualized trade variable, with average values of trade and ETR calculated by ventile. In each graph, the line represents the best linear fit based on the underlying country-year data, with the corresponding slope coefficient and standard error reported in the top-left corner. For more details, see Section 4.4.

Figure 6: Event Study of Trade Liberalization Reforms



Notes: These figures show event-studies for trade liberalization in seven large developing countries: Argentina, Brazil, China, Colombia, India, Mexico and Vietnam. The panels correspond to different outcomes: trade (top panels); effective tax rate on capital (middle panels); effective tax rate on labor (bottom panels). The left-side graphs show the average level of the outcome in every year to/since the event for the treated group and for the group of synthetic control countries. The right-hand graphs show the β_e coefficients on the to/since dummies, based on estimating the dynamic event-study regression in equation (4). The bars represent the 95% confidence intervals. Standard errors are clustered at the country-event level and estimated with the wild bootstrap method. The top-left corners report the F-statistic on the joint significance of the post-event dummies, with the p-value in parentheses. Details on methodology in Section 5.1.1.

Table 1: Trade Impacts on Effective Taxation of Capital and Labor in Developing Countries

	Benchmark		Robustness: Specification and covariates					Robustness: $K-L$ assignment to taxes and factor shares				Robustness: Individual instruments	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: ETR_K													
Trade	0.048***	0.151***	0.135***	0.211*	0.141**	0.136***	0.159***	0.161***	0.140***	0.147***	0.158***	0.148***	0.277***
	(0.013)	(0.047)	(0.037)	(0.121)	(0.055)	(0.044)	(0.046)	(0.052)	(0.045)	(0.045)	(0.047)	(0.047)	(0.095)
Panel B: ETR_L													
Trade	0.009*	0.047***	0.052***	0.059	0.037*	0.048***	0.049***	0.048***	0.051***	0.049***	0.042***	0.044***	0.214***
	(0.005)	(0.016)	(0.016)	(0.043)	(0.019)	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)	(0.015)	(0.016)	(0.067)
Specification	OLS	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
1^{st} stage Kleibergen-Paap F-statistic		24.59	24.59	34.51	14.14	23.24	34.84	24.59	24.59	24.59	24.59	45.13	10.75
Modifications to IV in col. (2)			ETR winsorize	NDP weights	Include country-year controls	Include **1**(oil-rich)*year fixed effects	Winsorize trade	Assign based on ILO (2019)	Assign based on corp. K-share	Assign 0% of PIT to capital	Assign 30% of PIT to capital	Only use $Z^{gravity}$ instrument	Only use $Z^{Oil-Dist}$ instrument
N	4916	4916	4916	4916	3938	4916	4916	4916	4916	4916	4916	4916	4916

Notes: This table presents results from estimating the effect of trade on effective tax rates in developing countries. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. The outcome is the effective tax rate on capital, ETR_K, in Panel A and the effective tax rate on labor, ETR_L, in Panel B. Trade is measured as the sum of exports and imports divided by net domestic product (NDP). Column (1) presents the OLS results from estimating equation (5). All other columns use IV; at the bottom of each column, we report the 1^{st}-stage Kleibergen-Paap F-statistic. The benchmark IV specification is in column (2), with the corresponding 1^{st}-stage regression reported in Table A2. The remaining columns modify the benchmark specification of column (2). In column (3), the outcome is non-winsorized, while in column (4) we include country-year NDP weights. In column (5), we include the country-year controls described in Section 5.2.1. In column (6), we include interactive fixed effects between a dummy for oil-rich countries and year dummies. Oil-rich countries derive more than 7% of GDP from oil in 2018. In column (7), we use the trade variable which is winsorized at the 5%-95% percentile on a yearly basis. In columns (8)-(9), we modify the assignment rule for mixed income's capital factor share, respectively by using the ILO (2019) method and by assigning the capital share in the corporate sector. In columns (10)-(11), we assign respectively 0% and 30% of personal income taxes (PIT) to capital taxes. In columns (12)-(13), we estimate the IV using the individual instruments $Z^{gravity}$ and $Z^{oil-distance}$, respectively. For more details, see Section 5.2. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table 2: Trade Impacts on Types of Taxes (% of NDP) in Developing Countries

	Total taxes (1)	CIT (2)	Property and Wealth (3)	PIT (4)	Social Security (5)	Indirect (6)
Panel A: OLS						
Trade	0.036***	0.021***	-0.001	0.003*	0.001	0.010
	(0.011)	(0.003)	(0.001)	(0.002)	(0.001)	(0.006)
Panel B: IV						
Trade	0.101***	0.053***	0.004	0.011**	0.013**	0.018
	(0.033)	(0.014)	(0.003)	(0.005)	(0.006)	(0.023)
1^{st}-stage Kleibergen-Papp F-statistic	24.59	24.59	24.59	24.59	24.59	24.59
N	4916	4916	4916	4916	4916	4916

Notes: This table shows the impacts of trade on collection of types of taxes, expressed as a percent of net domestic product (NDP), in developing countries. OLS results are in Panel A and IV results are in Panel B. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. Trade is measured as the sum of exports and imports divided by NDP. All regressions in Panel B are based on the IV model described in Section 5.2. At the bottom of each column, we report the 1^{st}-stage Kleibergen-Paap F-statistic. The corresponding 1^{st}-stage regression is reported in Table A2. The outcome differs across columns: Column (1) is total taxes, which is the sum of direct taxes on capital and labor and indirect taxes on trade and domestic consumption; column (2) is corporate income taxes (CIT); column (3) is taxes on property, wealth and inheritance; column (4) is personal income taxes (PIT); column (5) is social security and payroll; column (6) is indirect taxes, which combines trade taxes and domestic consumption taxes. For more details on these types of taxes, see Table B2 and Appendix B.1. For more details on the IV, see Section 5.2. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table 3: Trade Impacts on Mechanism Outcomes in Developing Countries

		National income components					Factor shares	
	First-diff. CIT rate (1)	Corporate totl. income (2)	Household mixed income (3)	Corporate profits (4)	Employee compensation (5)	Corporate ETR_K (6)	Capital share natl. income (7)	Capital share corp. sector (8)
Panel A: OLS								
Trade	-0.003***	0.038***	-0.016	0.026***	0.006	0.074***	0.020**	0.029**
	(0.001)	(0.013)	(0.011)	(0.009)	(0.011)	(0.019)	(0.008)	(0.012)
Panel B: IV								
Trade	-0.012*	0.179***	-0.184***	0.176***	-0.014	0.163**	0.150***	0.192***
	(0.007)	(0.044)	(0.041)	(0.035)	(0.036)	(0.075)	(0.034)	(0.050)
1^{st} stage Kleibergen-Paap F-Statistic	24.59	24.59	24.59	24.59	24.59	24.59	24.59	24.59
N	4916	4916	4916	4916	4916	4916	4916	4916

Notes: This table presents results from estimating the effects of trade on mechanism outcomes in developing countries. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. Trade is measured as the sum of exports and imports divided by net domestic product (NDP). Panel A presents OLS results and Panel B presents the IV results, based on the instruments described in Section 5.2. At the bottom of each column in Panel B, we report the 1^{st}-stage Kleibergen-Paap F-statistic. Across the columns, the outcome differs: column (1) is the first-differenced statutory corporate income tax (CIT) rate; column (2) is the corporate income share of net domestic product, where corporate income is the sum of corporate profits and corporate employee compensation; column (3) is the mixed income share of net domestic product; column (4) is the corporate profit share of net domestic product; column (5) is the employee compensation share of net domestic product; column (6) is the average effective tax rate on corporate profits; column (7) is the capital share of net domestic product; column (8) is the capital share of corporate income. For sake of space, we omit showing the insignificant impact of trade on OS_{HH}, the remaining component of national income. For more details on the outcomes, see Section 3.1 and Section 6.2. For more details on the instrumental variables, see Section 5.2. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table 4: Firm-Level Regressions in Rwanda: ETR^K, Trade and Size

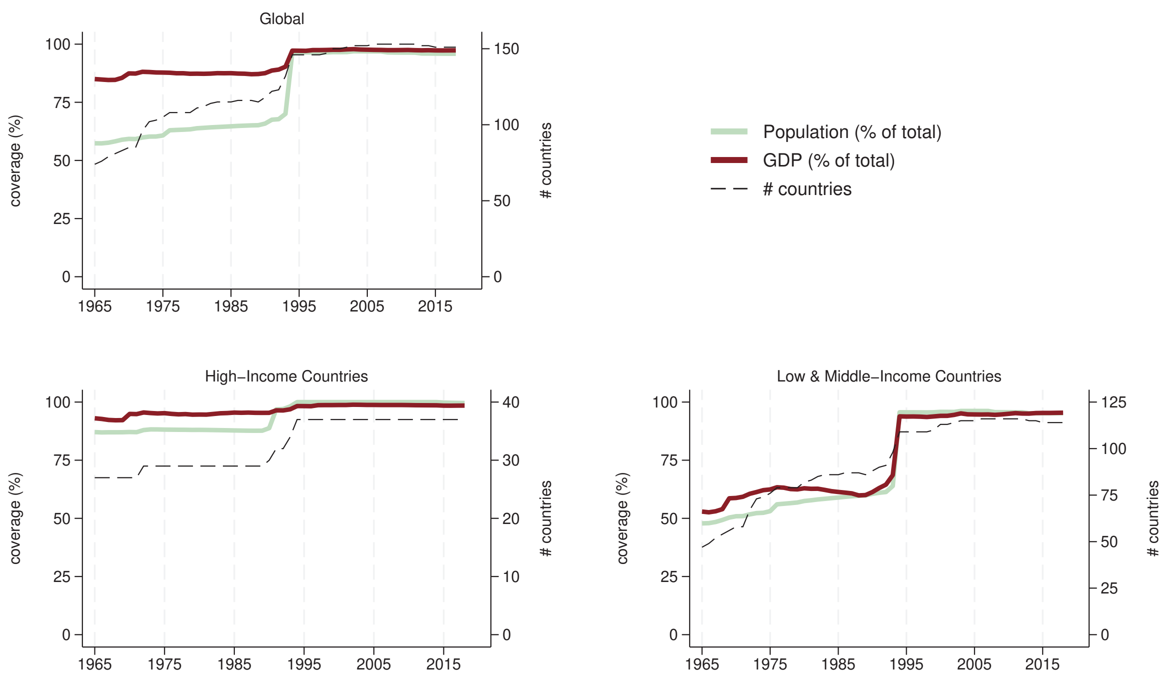
	(1)	(2)	(3)	(4)	(5)
Panel A outcome: ETR^K					
S^{Total}	0.100***	0.087***	0.075***	0.025*	0.133**
	(0.021)	(0.017)	(0.017)	(0.014)	(0.060)
Panel B outcome: Log revenue					
S^{Total}	1.362***	1.351**	1.078**	0.202*	1.444***
	(0.466)	(0.542)	(0.475)	(0.107)	(0.233)
Panel C outcome: ETR^K					
Log revenue	0.040*	0.092***	0.077**	0.029***	-
	(0.023)	(0.029)	(0.027)	(0.003)	-
Estimation	OLS	OLS	OLS	OLS	IV
1^{st}-stage Kleibergen-Paap F-statistic					18.17
Year FEs	Y	Y	Y	Y	Y
Industry-Geography FEs		Y	Y		
Firm controls			Y	Y	Y
Firm FEs				Y	Y
N	18478	18478	18478	18478	18478

Notes: This table presents firm-level regression results from corporate firms in Rwanda between 2015 and 2017. The outcome differs across panels: Panels A) and C) is the effective tax rate on corporate profits, ETR_i^K; Panel B) is log of annual revenue. In Panels A) and B), the reported regression coefficient is for total foreign input share, S^{Total}; in Panel C), it is for log annual revenue. Columns (1)-(4) present OLS results from estimating variations of equation (9): Column (1) includes year fixed effects; column (2) adds industry-geography fixed effects; column (3) adds firm-year controls (number of employees and total number of clients and suppliers); column (4) adds firm fixed effects. Column (5) is the IV estimation where the total foreign input share (S^{Total}) is instrumented with trade-shocks to firms and their supplier network based on the shift-share design of Hummels, Jørgensen, Munch, and Xiang (2014). The instruments are described in detail in Section 6.3 and Appendix D.1. In column (5), we also report the 1^{st}-stage Kleibergen-Paap F-statistic from estimating the 1^{st}-stage in equation (11). Details on the sample are provided in Appendix D.1. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the industry-geography level in columns (1)-(3), and at the firm-level in columns (4)-(5) (results are robust to clustering at firm-level in all columns).

Appendix

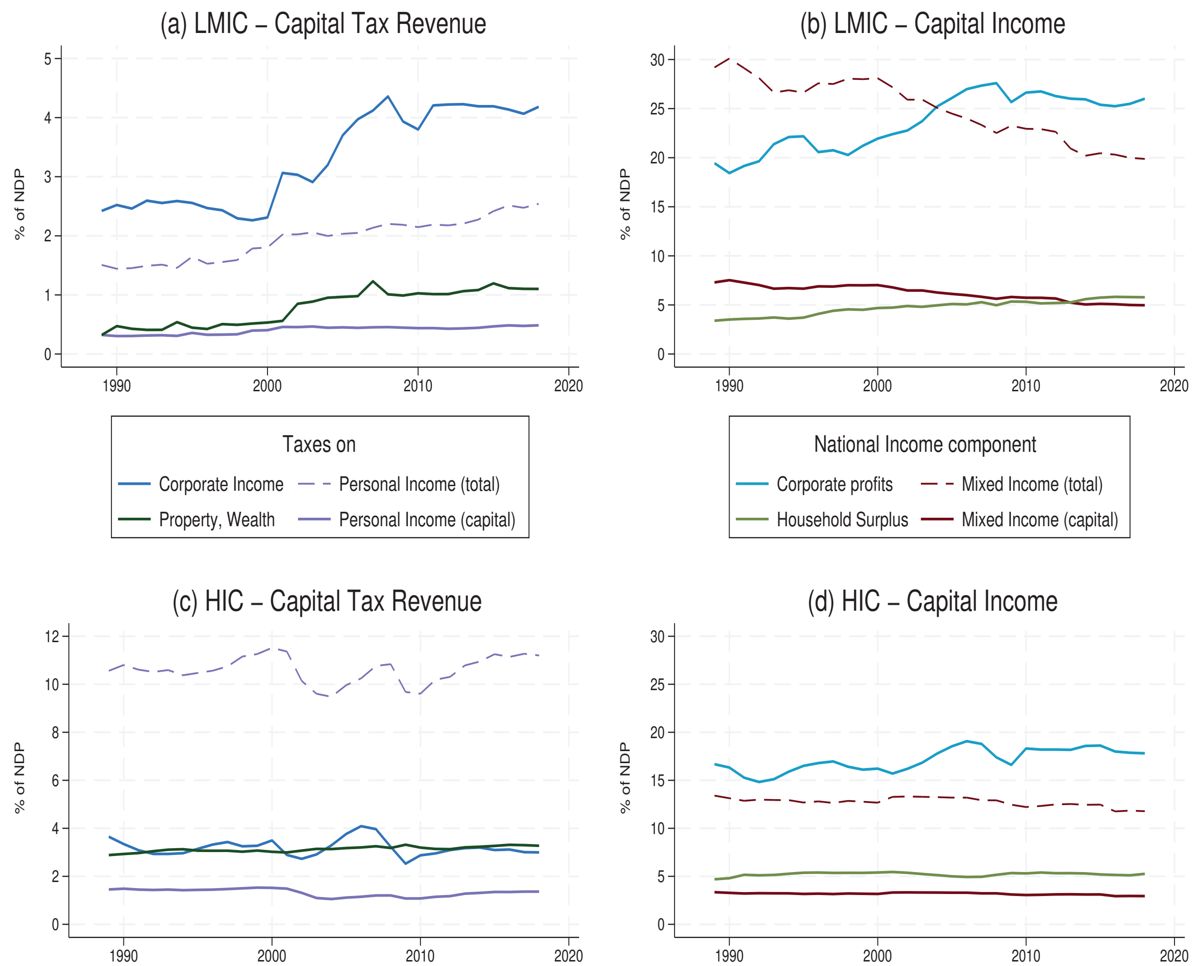
Appendix A Additional Figures and Tables

Figure A1: Data Coverage of Effective Tax Rates



Notes: These panels show the coverage of our effective tax rate data between 1965 and 2018 at the global level (top left panel), in high income countries (bottom left panel), and in low- and middle-income countries (bottom right panel). Low, middle and high-income countries are based on the World Bank income classification in 2018. The solid lines plot the percent of total population and GDP that are covered in our data (left axis). The dashed lines show the number of countries in the data (right axis). The dataset is composed of two quasi-balanced panels. The first covers the years 1965-1993 and excludes communist regimes. The second covers 1994-2018 and integrates former communist countries, in particular China and Russia. See Section 3.3 for more details.

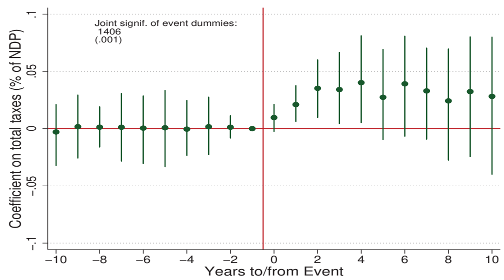
Figure A2: Evolution of ETR_K Components since 1989



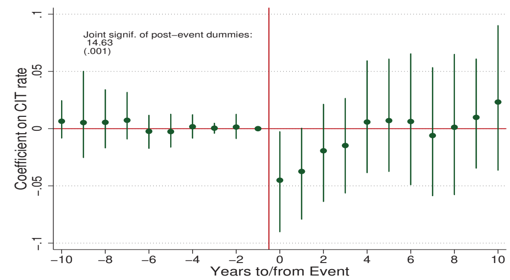
Notes: These panels show the evolution of the components of ETR_K between 1989 and 2018. This period is selected to match the period of rising ETR_K in low and middle-income countries (LMICs). The left-hand side panels correspond to the taxes on capital (numerator of ETR_K): corporate income taxes; taxes on property, wealth and inheritance; and the share of personal income taxes allocated to capital (including capital gains and dividends). The right-hand side panels correspond to the national income components attributed to capital (denominator of ETR_K): corporate profits; operating surplus of households (rents); and the share of mixed-income attributed to capital. The top panels are for LMICs, while the bottom panels are shown, as comparison, for high-income countries (HICs). Series are weighted by countries' national domestic product in 2010. The tax revenue data between 1989-1993 for former command economies (e.g. China, Russia) is missing, and is imputed by assigning the 1994 values 5 years backward. LMICs and HICs are defined according to the World Bank income classification in 2018. This figure is discussed in Section 4.1.

Figure A3: Mechanism Impacts in Trade Liberalization Event Studies

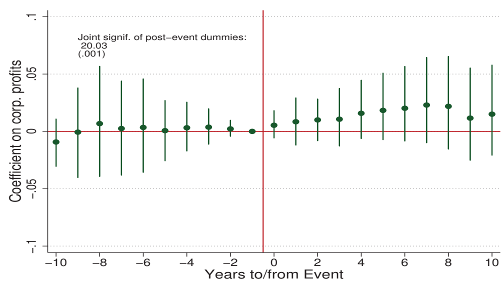
(a) Total taxes (% of NDP)



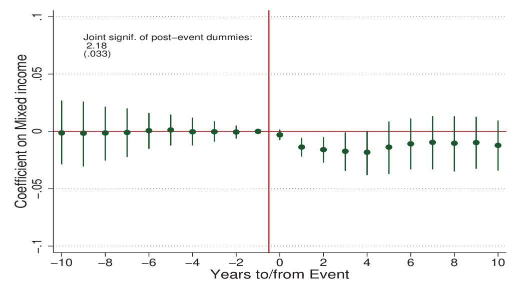
(b) Corp. income tax rate (first-diff.)



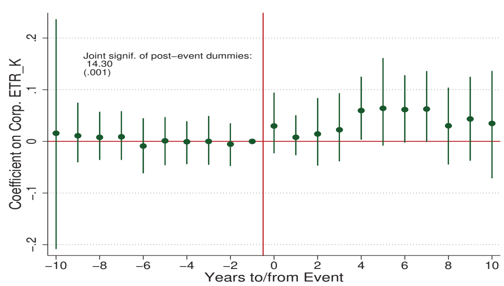
(c) Corporate income (% of NDP)



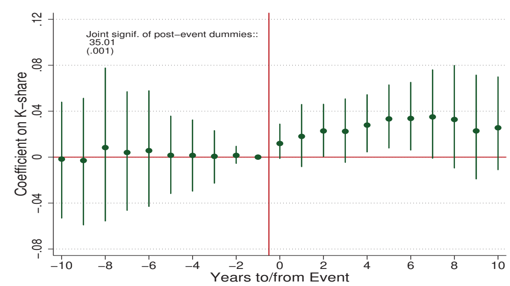
(d) Mixed income (% of NDP)



(e) Corporate ETR_K



(f) Capital income (% of NDP)



Notes: These panels show the impacts of the trade liberalization events on total taxes collected and mechanism outcomes. The panels are constructed using the method in Section 5.1, and similarly to Figure 6. Across panels, the outcome differs: panel a) is total tax revenue, as a percent of net domestic product (NDP); panel b) is the first-differenced statutory corporate income tax rate; panel c) is the corporate income share of net domestic product, where corporate income is the sum of corporate profits and employee compensation; panel d) is the mixed income share of net domestic product; panel e) is the average effective tax rate on corporate profits; panel f) is the capital share of net domestic product. In each panel, the top-left corner reports the F-statistic for the joint significance of post-event dummies, with the p-value reported in parentheses.

Table A1: Synthetic Difference-in-Difference of Trade Liberalization

	Trade	ETR_K	ETR_L
	(1)	(2)	(3)
Panel A: *Synthetic control for each outcome separately*			
Post*Treat	0.064	0.045***	0.020**
	(0.047)	(0.015)	(0.009)
Imputed treatment effect	0.070*	0.047***	0.020***
	(0.039)	(0.009)	(0.005)
Panel B: *Synthetic control for all outcomes jointly*			
Post*Treat	0.092*	0.033*	0.012
	(0.044)	(0.016)	(0.008)
Imputed treatment effect	0.101***	0.033***	0.012***
	(0.028)	(0.006)	(0.004)
Panel C: *Donor pool excluding major trading partners*			
Post*Treat	0.073	0.047***	0.018**
	(0.055)	(0.015)	(0.008)
Imputed treatment effect	0.082**	0.048***	0.018***
	(0.035)	(0.009)	(0.004)
Panel D: *Donor pool restricted to not-yet liberalized*			
Post*Treat	0.054	0.054***	0.013
	(0.058)	(0.014)	(0.008)
Imputed treatment effect	0.062*	0.054***	0.013***
	(0.034)	(0.009)	(0.005)
Panel E: *Donor pool restricted to same region*			
Post*Treat	0.049	0.034*	0.007
	(0.060)	(0.019)	(0.008)
Imputed treatment effect	0.058*	0.035***	0.017***
	(0.031)	(0.012)	(0.005)
Panel F: *Donor pool restricted to LMICs*			
Post*Treat	0.076	0.040**	0.016*
	(0.052)	(0.016)	(0.009)
Imputed treatment effect	0.085**	0.041***	0.016***
	(0.034)	(0.008)	(0.005)
N	294	294	294

Notes: This table shows the results from estimating the difference-in-difference effect and the imputed treatment effect - see Appendix C.2 for details. In Panel A, the synthetic control is created separately for each outcome (trade, ETR_K, ETR_L) and each liberalization country-event. In Panel B, the synthetic control is created for all three outcomes jointly for each country-event. In Panel C, the donor pool for each country-event excludes the 5 major import and export trading partners of the country, measured in terms of total volume of trade in the year immediately preceding liberalization. In Panel D, the donor pool excludes all countries that have already liberalized by the time of the event (based on Wacziarg & Welch, 2008). In Panel E, the donor pool is restricted to countries in the same region. In Panel F, the donor pool is all low and middle-income countries (LMICs), based on the World Bank income classification in 2018.* p<0.10 ** p<0.05 *** p<0.01.

Table A2: First-Stage and Reduced Form Regressions

	1^{st}-stage	Reduced form		1^{st}-stage		Reduced form	
	Trade (1)	ETR_K (2)	ETR_L (3)	Trade (4)	Trade$*\mathbb{1}$(High-inc.) (5)	ETR_K (6)	ETR_L (7)
$Z^{gravity}$	0.069***	0.010***	0.003***	0.014	0.040***	0.014*	0.002
	(0.010)	(0.002)	(0.001)	(0.019)	(0.014)	(0.008)	(0.004)
$Z^{oil-distance}$	-0.116***	-0.033***	-0.020**	-0.088***	-0.021	-0.022***	-0.015***
	(0.036)	(0.009)	(0.005)	(0.015)	(0.014)	(0.007)	(0.003)
1^{st}-stage F-statistic	24.59			22.82	11.75		
1^{st}-stage Sanderson-Windmeijer Weak Instruments F-statistic	24.59			41.93	26.60		
1^{st}-stage Kleibergen-Paap F-statistic	24.59			15.34			
Sample	Developing countries only			Developing and developed countries			
N	4916	4916	4916	6489	6489	6489	6489

Notes: This regression table shows the first stage and the reduced form results. The sample is developing countries ($N = 49160$) in cols. (1)-(3), and developing and developed countries ($N = 6489$) in columns (4)-(7). Trade is exports and imports divided by net domestic product. Column (1) corresponds to the first-stage in developing countries, used in Tables 1-2-3. Columns (4)-(5) correspond to the first-stage in the full sample, which estimates heterogeneous effects by development level, and which is used in Table A7. We report several 1^{st}-stage statistics: the F-statistic of excluded instruments; the Sanderson-Windmeijer multivariate F-test of excluded instruments; and, the Kleibergen-Paap F-statistic. When there is only one endogenous regressor (column 1), these three F-statistics are equivalent. Note in columns (4)-(5) that there is only one Kleibergen-Paap F-statistic, which evaluates the overall strength of the first-stage, even though there are two first-stage regressions. Columns (2)-(3) and (6)-(7) report the reduced form regressions of the instruments on the effective tax rates for capital, ETR_K, and labor, ETR_L. Developing (developed) countries are low and middle-income countries (high-income countries) according to the World Bank income classification in 2018. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table A3: Trade Impacts on Effective Tax Rates in Different Samples

	Sample changes related to tax revenue data				Sample changes related to System National Accounts data			Sample changes related to time-periods and balancedness		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: ETR_K										
Trade	0.157***	0.133***	0.215**	0.206***	0.162**	0.138**	0.183***	0.110*	0.205**	0.150***
	(0.054)	(0.049)	(0.098)	(0.068)	(0.064)	(0.062)	(0.052)	(0.060)	(0.098)	(0.052)
Panel B: ETR_L										
Trade	0.051***	0.029***	0.093*	0.028	0.039**	0.037*	0.041**	0.041**	0.056**	0.067***
	(0.017)	(0.011)	(0.049)	(0.022)	(0.018)	(0.018)	(0.020)	(0.015)	(0.021)	(0.020)
Modifications to benchmark sample in Table 1	Remove interpolated tax revenue	Only use HA tax data	Only use ICTD tax data	Only use OECD tax data	Remove composite SNA data	Only use SNA1968 data	Only use SNA2008 data	Only use pre-1994 years	Only use post-1994 years	Fully balanced panel 1965-2018
N	4563	2268	1004	1644	2752	983	1769	2122	2794	2879

Notes: This table presents results from estimating the effect of trade on effective tax rates in different samples across developing countries. The estimation is identical to the benchmark IV model in column (2) of Table 1; across columns, the sample differs from that benchmark sample. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. The outcome is the effective tax rate on capital, ETR_K, in Panel A and the effective tax rate on labor, ETR_L, in Panel B. Trade is measured as the sum of exports and imports divided by net domestic product (NDP). In the first four columns, sample-changes are made to the tax revenue data: interpolated values are dropped in column 1; the only data-source is historical archives (HA) in column 2; the only data-source is ICTD in column 3; the only data-source is OECD in column 4. In the next three columns, sample-changes are made to the system of national accounts (SNA) data: in column (5), the composite SNA values are removed; in column (6), only data from SNA1968 are used; in column (7), only data from SNA2008 are used. In the final three columns, sample-changes are made regarding balancedness: in column (8), the quasi-panel between 1965 and 1993 is used; in column (9), the quasi-panel between 1994 and 2018 is used; in column (10), the fully balanced panel of countries between 1965 and 2018 is used. For more details on the interpolations, imputations and data-sources, see Section 3 and Appendix B.

Table A4: Robustness of Results for Total Taxes and Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Total taxes (% of NDP)						
Trade	0.105*	0.092**	0.097***	0.106***	0.099***	0.170**
	(0.060)	(0.039)	(0.031)	(0.032)	(0.032)	(0.073)
1^{st} stage K-P F-stat	34.51	14.14	23.09	34.84	45.17	10.75
N	4916	3938	4916	4916	4916	4916
Panel B: CIT rate (first-diff.)						
Trade	0.004	-0.007	-0.011*	-0.013*	-0.012*	-0.030*
	(0.011)	(0.009)	(0.006)	(0.007)	(0.007)	(0.016)
1^{st} stage K-P F-stat	34.51	14.14	23.24	34.84	45.13	10.75
N	4916	3938	4916	4916	4916	4916
Panel C: $log(1+\text{CIT rate})$						
Trade	-0.009*	-0.006	-0.009*	-0.010*	-0.009*	-0.026*
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.014)
1^{st} stage K-P F-stat	34.51	14.14	23.24	34.84	45.13	10.75
N	4916	3938	4916	4916	4916	4916
Panel D: Corp. income (% of NDP)						
Trade	0.188***	0.197***	0.173***	0.189***	0.179***	0.211**
	(0.051)	(0.047)	(0.044)	(0.046)	(0.045)	(0.104)
1^{st} stage K-P F-stat	34.51	14.14	23.24	34.84	45.13	10.75
N	4916	3938	4916	4916	4916	4916
Panel E: Mixed income (% of NDP)						
Trade	-0.203***	-0.162***	-0.184***	-0.194***	-0.185***	-0.137
	(0.053)	(0.040)	(0.040)	(0.038)	(0.041)	(0.112)
1^{st} stage K-P F-stat	34.51	14.14	23.24	34.84	45.13	10.75
N	4916	3938	4916	4916	4916	4916
Panel F: Capital share of NDP						
Trade	0.102*	0.112**	0.145***	0.158***	0.152***	0.107**
	(0.052)	(0.044)	(0.032)	(0.033)	(0.035)	(0.052)
1^{st} stage K-P F-stat	34.51	14.14	23.24	34.84	45.13	10.75
N	4916	3938	4916	4916	4916	4916
Panel G: Corp. ETR_K						
Trade	0.238*	0.189*	0.148**	0.172**	0.160**	0.385**
	(0.156)	(0.096)	(0.074)	(0.077)	(0.076)	(0.183)
1^{st} stage K-P F-stat	34.51	14.14	23.24	34.84	45.13	10.75
N	4916	3938	4916	4916	4916	4916
Modifications to IV in Panel B of Table 3	NDP weights	Include country-year controls	Include **1**(oil-rich)*year fixed effects	Winsorize trade at 5%-95%	Only use $Z^{gravity}$ instrument	Only use $Z^{Oil-Dist}$ instrument

Notes: This table presents robustness checks for trade's impacts on several outcomes in developing countries. Developing countries are low and middle-income countries according to the World Bank income classification in 2018. Trade is the sum of exports and imports divided by net domestic product (NDP). The outcome differs across panels, and the specification differs across columns: each cell is the coefficient from a separate IV regression. We report the 1^{st}-stage Kleibergen-Paap F-statistic separately for each IV regression. Panel A is total taxes as a % of NDP. Panel B is the first-differenced corporate income tax (CIT) rate. Panel C is the percent change from log of $(1 + \text{CIT rate})$. Panel D is the corporate income share of NDP. Panel E is the mixed income share of NDP. Panel F is the capital share of NDP. Panel G is the average effective tax rate on corporate profits. The different specifications across columns are the same as in Table 1 - please refer to that table for more details. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table A5: Impacts of Trade in LMICs, Heterogeneity by Enforcement Policy

	ETR_K (1)	ETR_L (2)	Corp. income (3)	Corp. ETR_K (4)
Panel A: Large Taxpayer Unit				
Trade	0.116*	0.013	0.171***	0.117*
	(0.066)	(0.029)	(0.057)	(0.068)
Trade∗𝟙(LTU)	0.089	0.084**	0.019	0.113
	(0.077)	(0.040)	(0.051)	(0.131)
Implied coef. for Trade with LTU	0.205***	0.098***	0.190***	0.230**
	(0.062)	(0.029)	(0.042)	(0.097)
Panel B: Customs-Tax Integration				
Trade	0.121*	0.018	0.172***	0.160*
	(0.064)	(0.038)	(0.052)	(0.094)
Trade∗𝟙(Customs-Tax)	0.208	0.198*	0.046	0.183
	(0.185)	(0.109)	(0.112)	(0.249)
Implied coef. for Trade with Customs-Tax	0.330**	0.217**	0.219**	0.344*
	(0.153)	(0.090)	(0.089)	(0.202)
Panel C: Value-Added Tax				
Trade	0.116**	0.015	0.171***	0.156*
	(0.058)	(0.025)	(0.054)	(0.089)
Trade∗𝟙(VAT)	0.101	0.096**	0.022	0.085
	(0.081)	(0.043)	(0.054)	(0.115)
Implied coef. for Trade with VAT	0.218***	0.111***	0.194***	0.241***
	(0.064)	(0.032)	(0.045)	(0.087)
Panel D: International Accounting Standards				
Trade	0.132**	0.023	0.160***	0.183**
	(0.054)	(0.022)	(0.051)	(0.088)
Trade∗𝟙(IAS)	0.122	0.111**	0.017	0.124
	(0.087)	(0.042)	(0.055)	(0.135)
Implied coef. for Trade with IAS	0.255**	0.134***	0.177***	0.307***
	(0.077)	(0.036)	(0.050)	(0.110)
N	4916	4916	4916	4916

Notes: This table estimates heterogeneous IV effects of trade in developing countries (low and middle-income countries according to the World Bank income classification in 2018). Trade is the sum of exports and imports divided by net domestic product (NDP). Outcomes differ across columns: column (1) is the effective tax rate on capital, ETR_K; column (2) is the effective tax rate on labor, ETR_L; column (3) is the corporate income share of NDP; column (4) is the average effective tax rate on corporate profits. We estimate

$$y_{ct} = \mu \cdot trade_{ct} + \kappa \cdot trade_{ct} \cdot \mathbb{1}(A)_{ct} + \theta \cdot \mathbb{1}(A)_{ct} + \pi_c + \pi_t + \epsilon_{ct}$$

where $\mathbb{1}(A)_{ct}$ is an indicator variable which takes a value of 1 in all years after the administrative reform has been implemented. We instrument for $trade_{ct}$ and $trade_{ct} \cdot \mathbb{1}(A)_{ct}$ using the two instruments (Section 5.2). The coefficient on $\mathbb{1}(A)_{ct}$ is also estimated, but is not reported in the table. In Panel A, the administrative reform is the existence of a large taxpayer unit (LTU); this variable is coded based on the USAID's 'Collecting Taxes Database' (website link) and country-sources. In Panel B, the administrative reform is the integration of the customs authority and the domestic tax authority in a single revenue agency; this variable is coded based on USAID's 'Collecting Taxes Database' (website link), the OECD Tax Administration Comparative Series (website link), and country-sources. In Panel C, the administrative reform is the implementation of a value-added tax (VAT); this variable is coded based on Keen and Lockwood (2010) and country-sources. In Panel D, the administrative reform is the adoption of international accounting standards (IAS); this variable is coded based on the IAS country-profiles (website link). At the bottom of each column and panel, we report the implied coefficient and estimated standard error based on the linear combination of the $trade_{ct}$ and $trade_{ct} \cdot \mathbb{1}(A)_{ct}$ coefficients. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table A6: Impacts of Trade Outside of Periods of Tax Revenue Loss

	ETR_K (1)	ETR_L (2)	Corp. income (3)	Corp. ETR_K (4)
Panel A: Excluding Trade-Induced Tariff Revenue Loss Periods (based on Cage and Gadenne, 2018)				
Trade	0.151***	0.047**	0.183***	0.203**
	(0.056)	(0.020)	(0.045)	(0.089)
N	3954	3954	3954	3954
Panel B: Excluding Periods of Indirect Tax Revenue Loss				
Trade	0.189***	0.053***	0.197***	0.225***
	(0.051)	(0.016)	(0.044)	(0.083)
N	3011	3011	3011	3011
Panel C: Excluding Periods of Total Tax Revenue Loss				
Trade	0.174***	0.048***	0.174***	0.203**
	(0.050)	(0.015)	(0.042)	(0.081)
N	3016	3016	3016	3016

Notes: This IV specification is the same as column (2) in Table 1, but modifications are made to the sample of developing countries. In Panel A, we exclude all country-year observations which belong to an episode of trade revenue loss, based on Cagé and Gadenne (2018). In a dataset of 130 countries between 1792 and 2006, the authors define such an episode by a fall in trade tax revenues as a percentage of GDP of at least 1 percentage point from a local yearly maximum to the next local yearly minimum that is accompanied by a non-decrease in the volume of imports as a share of GDP. In Panels B and C, we consider alternative definitions of revenue loss periods. In Panel B, we calculate the within-country yearly change in indirect taxes collected as a share of net domestic product (NDP), and take the three-year moving average. We then create terciles of this variable, separately for each country. We define periods of indirect tax revenue loss to be the observations which lie in the bottom tercile of this distribution, and exclude these country-year observations from the sample. In Panel C, we calculate the same revenue-loss variable, but based on changes in total taxes collected rather than indirect taxes collected. Trade is the sum of exports and imports divided by NDP. The outcome differs across columns: column (1) is the effective tax rate on capital, ETR_K; column (2) is the effective tax rate on labor, ETR_L; column (3) is the corporate income share of NDP; column (4) is the average effective tax rate on corporate profits. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table A7: Heterogeneous Impacts of Trade by Development Level

	ETR_K (1)	ETR_L (2)	First-diff. CIT Rate (3)	Corp. Totl. Income (4)	Mixed Income (5)	Corp. Profits (6)	Employee Comp. (7)	Corp. ETR_K (8)	Natl. K-Share (9)	Corp. K-Share (10)
Trade	0.253**	0.116**	-0.020	0.279**	-0.183*	0.176***	0.056	0.445**	0.131**	0.158**
	(0.127)	(0.051)	(0.021)	(0.119)	(0.106)	(0.049)	(0.104)	(0.193)	(0.054)	(0.064)
Trade$*\mathbb{1}$(High-inc.)	-0.293	0.014	-0.064*	-0.502**	0.340**	-0.312***	-0.214**	-0.289	-0.197**	-0.239**
	(0.215)	(0.110)	(0.033)	(0.218)	(0.138)	(0.099)	(0.114)	(0.320)	(0.086)	(0.110)
Implied coef. for Trade in High-inc.	-0.040	0.130	-0.084***	-0.223	0.160	-0.135*	-0.158	0.156	-0.066	-0.081
	(0.127)	(0.095)	(0.020)	(0.154)	(0.135)	(0.072)	(0.117)	(0.173)	(0.056)	(0.081)
1^{st}-stage Kleibergen-Papp F-statistic	15.34	15.34	15.34	15.34	15.34	15.34	15.34	15.34	15.34	15.34
N	6489	6489	6489	6489	6489	6489	6489	6489	6489	6489

Notes: This table presents IV results from estimating the effects of trade on ETR and mechanism outcomes in the full sample of developing and developed countries. Trade is measured as the sum of exports and imports divided by net domestic product (NDP). We run the following IV regression: $y_{ct} = \mu \cdot trade_{ct} + \kappa \cdot trade_{ct} \cdot \mathbb{1}(HighIncome)_c + \Theta \cdot X_{ct} + \pi_c + \pi_t + \epsilon_{ct}$ The first-stage regression is reported in Table A2. At the bottom of each column, we report the implied coefficient and estimated standard error based on the linear combination of the $Trade$ and the $Trade * \mathbb{1}$(High-inc.) coefficients. High-income is based on the World Bank income classification in 2018. We also report the 1^{st}-stage Kleibergen-Paap F-statistic. Each column is a different outcome: column (1) is the effective tax rate on capital; column (2) is the effective tax rate on labor; column (3) is the first-differenced statutory corporate income tax rate; column (4) is the corporate income share of net domestic product, where corporate income is the sum of corporate profits and corporate employee compensation; column (5) is the mixed income share of net domestic product; column (6) is the corporate profit share of net domestic product; column (7) is the employee compensation share of net domestic product; column (8) is the average effective tax rate on corporate profits; column (9) is the capital share of net domestic product; column (10) is the capital share of corporate income. For more details on outcomes, see Section 3.1 and Section 6.2. For more details on the instrumental variables, see Section 5.2. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Appendix B Data & Construction of Effective Tax Rates

This appendix section provides an overview of the data sources used to create our tax revenue and national income series (Section B.1). Additionally, we discuss the methodology to measure effective tax rates (Section B.2).

B.1 Data sources

Tax revenue data Our tax revenue data draws from three key sources:

(i) **OECD Government Revenue Statistics** (website link): OECD revenue statistics take precedence in our data hierarchy as it contains all types of tax revenues already arranged in the OECD taxonomy of taxes. While it covers all OECD countries, it only covers a subset of developing countries which typically start in the early 2000s.

(ii) **ICTD Government Revenue Dataset** (website link): ICTD data covers many developing countries, but only begins in the 1980s. ICTD at times does not separate income taxes into personal vs. corporate taxes and often does not contain social security contributions.

(iii) **Archival data**: The main archival data collection corresponds to the digitization of the Government Documents section in the Lamont Library at Harvard University (website link). For each country, we scanned, tabulated and harmonized official data from the public budget and national statistical yearbooks, to retrieve official tax revenue statistics. The supplementary appendix lists the main historical documents used in each country's time-series. In the case where the document is a statistical yearbook, the initial listed source is always a report produced by the finance ministry or the national tax authority. To complement hard-copy archival data, we retrieved countries' online reports, usually published by their national statistical office or finance ministry. We also used complementary sources, including offline archival Government Finance Statistics data from the IMF which covers the period 1972-1989. For social security contributions, we relied on two additional sources: the 'D61' statistic on social contributions in the household sector in SNA-1968 and SNA-2008, and data from Fisunoglu, Kang, Arbetman-Rabinowitz, and Kugler (2011).

To increase the credibility of the tax revenue series based on newly digitized historical documents, we base our approach on the following four guiding rules:

1. We seek to build long time-series from the archival records in order to overlap with pre-existing sources (OECD, ICTD, IMF). We use the overlapping years to inspect that the different sources provide similar estimates of the overall levels of taxes collected and to verify that they report the same set of taxes in place. If discrepancies exist when data sources overlap, we inspect the accuracy of each source with additional information. For this reason, switches in data-source rarely lead to a significant change in trend.

2. In historical time-periods where no overlap exists with pre-existing sources, we find academic publications and policy reports to compare the estimated overall levels of

tax/GDP. When discrepancies exist, we investigate its causes (e.g. inclusion of non-tax revenues, differences in estimated GDP numbers).

3. We take note of instances where the overall tax take, or individual tax types, see sudden and large changes. We use additional sources to try to determine the proximate causes as they relate to policy changes, political transitions or economic shocks. We flag cases where we cannot find the proximate cause or where the political or economic events induce very significant volatility in the time-series.

4. We aim to be conservative in our inclusion of countries and time-periods. Specifically, we exclude countries in time-periods where data exists but where significant concerns remain about its reliability (and where it proves difficult to find corroborating sources). These instances are often in periods of significant political or economic change. For example, we exclude Afghanistan in the late 1970s and early 1980s; Cambodia in the late 1980s and early 1990s; Dominican Republic in the early 1960s; and, Namibia in 1990.

The supplementary appendix contains a table which summarizes our decisions as they relate to these four guiding rules in each country in our sample. The table emphasizes the uncertainty that exists for specific countries in specific time periods and we flag instances where we assess the data to be worthy of inclusion but where it should still be interpreted with caution and additional investigations would be helpful. We confirm that none of our main results change if we exclude these flagged instances. Moreover, the supplementary appendix provides case-studies with additional details on our decisions and direct links to the initial historical documents for each country. The case studies are currently limited to 67 countries but will ultimately cover the entire sample.[63]

Equipped with the historical time series, we have to construct long-run panels across sources. Below, we outline the instructions used to harmonize across sources and to improve data quality for the measurement of each type of tax. We flag instances where we consider the series to be legitimate, but where harmonization proved more challenging due to coinciding economic or political changes. For each country, the main decisions related to harmonization and data-quality are provided in the supplementary appendix.

1. We first rely on OECD data whenever it exists. Archival data is initially second in priority, but we revise this based on whether ICTD data provides a long time series and separates personal from corporate income taxes. We also study if ICTD has the better match in overlapping time-periods with OECD data. When possible, we aim to use no more than two data sources per country.

2. We exclude country-years for communist/command economies. This implies that our panel size jumps in 1994, including when China and Russia first appear. The year 1994 is a few years removed from the dissolution of the Soviet Union but, as discussed below, arguably corresponds to China's establishment of a modern tax system (World Bank, 2008).

[63]We invite comments from researchers to improve the accuracy of the series as we build the case studies and expand the data to recent years.

3. When none of the data sources separate PIT from CIT, we use academic sources and tax legislation to assign values.

4. To guard against omitting significant values of decentralized tax revenues, we use the OECD database on subnational government finance (link) to find the countries with significant state and local taxes, and we attempt to collect further data for these countries if necessary.

5. We linearly interpolate data when a given tax type is missing, but for no more than 4 years in a time-series and without extrapolation. We check for significant socio-economic changes that could cast doubt on the continuity of the tax revenue series and do not interpolate in such years.

6. We only use actual amounts of taxes collected, and do not rely on estimated values.

China's establishment of a modern tax system in 1994

In our benchmark setting, we only include formerly communist economies into our data starting in 1994. Given China's weight in the global economy, it is worth reviewing the reason for that choice. The tax revenue data for China covers most of our sample period although its quality improves markedly in the 1980s. Official statistics are available online: link here.

Prior to the 1980s, China had a command economy model of 'profit delivery,' in which the state directly received the revenues of profitable SOEs, and subsidized unprofitable ones. A corporate income tax first appears in China in 1983-84, but the majority of the base continues to be state-owned enterprises. In 1985, the tax system was further reformed into a 'fiscal contracting' system whereby firms negotiated a fixed lump-sum payment (regardless of economic outcomes), which cannot be split into labor versus capital taxes (nor into consumption taxes). We therefore exclude the 'pseudo'-CIT revenue dating from 1985 through 1993.

Rather, we consider that China's modern tax system began in 1994. The World Bank (2008) shows that, in 1994, China established for the first time a central tax administration; reformed the 'fiscal contracting' system; unified the PIT; created a VAT; and reduced 'extra budgetary' (non-tax) revenues. Thus from 1994 onward we can categorize tax revenue precisely by type, assign them to capital or labor, and estimate our ETRs.

National accounts data To compute factor incomes of net domestic product, we combine two main datasets from the United Nations Statistics Division. The first is the 2008 System of National Accounts (SNA) online data repository. The second is the 1968 SNA archival material. The 2008 and 1968 SNAs initially have different reporting classifications; to the best of our knowledge, our project is the first to harmonize national accounts across these two sources.

To estimate capital and labor factor incomes requires information on the 4 main sub-components that make up net domestic product (see equation 3). However, in some country-years where we have information on domestic product from an SNA dataset, there may not be data on all four sub-components at the same time. This is more frequently the case for the 1968 SNA than for the 2008 SNA and it is most frequent for mixed income (OS_{PUE}). In these cases, we first attempt to recover the value of the missing component using data from the other SNA dataset and national accounting identities with non-missing values for other components within the same country-year. For the remaining cases after applying this process, we impute values for the component. All of the regressions in Sections 5-6 include dummy variables for these composite cases; our main results also hold without the imputed values (Table A3). For the imputation, we follow the procedure from Blanchet et al. (2021). The World Inequality Database uses this procedure to impute consumption of fixed capital (depreciation) when it is missing in countries' series. For example, applying this procedure in our setting means that we model OS_{PUE} as a function of log national income per capita, a fixed country characteristic, and an AR(1) persistence term.

Table B1 summarizes the national accounts coverage in our dataset. The 'Complete SNA2008' row refers to country-years where all components of net domestic product are extracted from the 2008 SNA; similarly for the 'Complete SNA1968' row. The 'Composite' row counts instances where one component (or more) of net domestic product is initially missing from an SNA dataset and is retrieved from the other SNA dataset, is calculated via accounting identities, or is imputed.

Table B1: Main Data Sources

	Country-year obs.	%
Panel A: Tax revenue data		
OECD	2875	42.3%
Archives	2678	39.4%
ICTD	1246	18.3%
N	*6799*	*100%*
Panel B: Factor income data		
Complete SNA2008	2455	36.1%
Complete SNA1968	1360	20.0%
Composite	2984	43.9%
N	*6799*	*100%*

Notes: See Section B.1 for more details on the data-sources for tax revenue and factor income.

B.2 Construction of ETR

By combining data on disaggregated tax revenues and national income components, we construct effective tax rates on capital and labor (equations 1 and 2 in Section 3.1). Here we provide further details on the definitions of ETR. Computing ETR_L and ETR_K requires the following information for country c in year t:

$$ETR_{L,ct} = \frac{T_{L,ct}}{Y_{L,ct}} = \frac{\lambda_{PIT,ct} \cdot T_{1100,ct} + \lambda_{socsec,ct} \cdot T_{2000,ct}}{CE_{ct} + \phi_{ct} \cdot OS_{PUE,ct}}$$

$$ETR_{K,ct} = \frac{T_{K,ct}}{Y_{K,ct}} = \frac{(1-\lambda_{PIT,ct}) \cdot T_{1100,ct} + (1-\lambda_{CIT,ct}) \cdot T_{1200,ct} + (1-\lambda_{assets,ct}) \cdot T_{4000,ct}}{(1-\phi_{ct}) \cdot OS_{PUE,ct} + OS_{CORP,ct} + OS_{HH,ct}}$$

For each type of tax j, there is a $\lambda_{j,ct}$ allocation of the tax to labor which may vary by country-year (and $1-\lambda_{j,ct}$ is the allocation to capital). The allocation for each type of tax is described in Table B2, where the types of taxes follow the OECD classification. In our benchmark assignment, these allocations are time- and country-invariant for all types of taxes, except for personal income taxes ($\lambda_{PIT,ct}$) which we discuss in detail below. Further, in our benchmark assumption, we assume that the labor share of mixed income, ϕ_{ct}, is fixed at 75% in all country-years ($\phi_{ct} = 0.75$). In robustness checks, we let ϕ_{ct} vary at the country-level, based on ILO (2019), or at the country-year level by using the labor share in the corporate sector. In our benchmark assignment, replacing the invariant parameters with their fixed numerical values, we therefore have:

$$ETR_{L,ct} = \frac{T_{L,ct}}{Y_{L,ct}} = \frac{\lambda_{PIT,ct} \cdot T_{1100,ct} + T_{2000,ct}}{CE_{ct} + 0.75 \cdot OS_{PUE,ct}}$$

$$ETR_{K,ct} = \frac{T_{K,ct}}{Y_{K,ct}} = \frac{(1-\lambda_{PIT,ct}) \cdot T_{1100,ct} + T_{1200,ct} + T_{4000,ct}}{0.25 \cdot OS_{PUE,ct} + OS_{CORP,ct} + OS_{HH,ct}}$$

The parameter values are described in Table B2, both for the tax revenue numerator and the national income denominator. We now provide more details on λ_{PIT} and ϕ.

Labor share of personal income taxes: λ_{PIT} As discussed in Section 3.1, the level of personal income tax (PIT) that derives from capital versus labor income is rarely directly observed.[64] Thus, within PIT, an important parameter is the share of revenue assigned to labor, denoted λ_{PIT}. In the United States, Piketty et al. (2018) find that approximately 85% of PIT revenue is from labor and 15% from capital. To construct country-year specific $\lambda_{PIT,ct}$, we start from the US benchmark ($\lambda_{PIT} = 85\%$) and make two adjustments:

(a) First, the location of the PIT exemption threshold in the income distribution impacts λ_{PIT}, since the capital income share is higher for richer individuals. We retrieve PIT exemption thresholds from Jensen (2022). We assume countries with a higher PIT exemption threshold have a higher λ_{PIT}. Since the US has a low exemption threshold

[64] PIT revenue from capital income includes taxes on dividends and capital gains and on the capital share of self-employment income. OECD revenue data occasionally reports tax revenue from capital gains, which was on average 4% of PIT in the period 2010-2018 (7.5% in the US).

with $\lambda_{PIT} = 85\%$, we assign 85% of PIT to labor in countries where the PIT at least half of the workforce (mainly high-income countries). For countries where the PIT covers 1% or less of the workforce (lowest-income countries), we assign a maximum PIT capital share of 30%. For PIT thresholds with a coverage between 1% to 50% of the workforce, we linearly assign λ_{PIT} between 70% and 85%.

(b) Second, we assume that countries where a dual PIT system is in place have a larger λ_{PIT}. Dual PIT systems set capital income taxation to a lower—often flat—rate, while labor income is taxed with progressive marginal tax rates. We compute the measure of the percent difference between the tax rate on dividends and the top marginal tax rate on labor income. Data on dividend vs wage income tax rates are taken from OECD Revenue Statistics and country-specific tax code documents. Since we only have dividend rates, we assume that 50% of capital income in PIT benefits from the lower rate (e.g., capital gains might not benefit). For this 50%, we multiply λ_{PIT} by the percent difference in dividend versus top marginal tax rates.

Labor share of mixed income: ϕ Section 3.1 noted the difficulty of estimating the labor share of mixed income (unincorporated enterprises). We assume a benchmark measure of $\phi = 75\%$. The implied capital share is lower than the 30% used in Distributional National Accounts guidelines (Blanchet et al., 2021). However, since the global average corporate capital share is 27%, assuming that the capital share of unincorporated enterprises is slightly lower appears reasonable (see Guerriero, 2019).

We implement two robustness checks. First, we set the labor share of mixed income equal to that of the corporate sector at the country-year level; specifically, $\phi_{ct} = \frac{CE_{ct}}{CE_{ct}+OS_{CORP,ct}}$. This procedure follows Gollin (2002).

Second, we implement the ILO (2019) method which relies on harmonized household surveys and labor force surveys in developing countries between 2004 and 2017. Estimation of the relative labor income of self-employed is based on the observable characteristics of those workers and their comparison with employees. Relevant variables, including industry, occupation, education level and age, are used in a regression to uncover the determinants of labor income of employees. Given the estimated relationship between employee labor income and the explanatory variables, labor income is extrapolated to self-employed, generating a coefficient of relative earnings to employees, denoted γ_q. The method estimates a separate γ_q for different groups q of self-employed: self-employed workers; own-account workers; and, contributing family members. A correction procedure is implemented to reduce the bias from selection into self-employment. Total labor income in a given country-year is then determined as $Y_L^{ILO} = CE + \sum_q w_{emp} \cdot \gamma_q \cdot b_q$, where CE is the total compensation of employees in SNA, w_{emp} is the average employee wage (which relates CE to the total employee workforce), b_q is self-employed group q's count in the workforce, and γ_q is the q-specific earnings coefficient relative to the average employee wage. Equipped with the Y_L^{ILO} estimate, we calculate the 'implicit' labor mixed income ($OSPUE_L$) as the difference between Y_L^{ILO} and the value of compensation for employees CE observed in the national accounts. Then, we compute the mixed income share allocated to labor. Specifically, ϕ^{ILO} is computed as follows: $\phi^{ILO} = \frac{(Y_L^{ILO} - CE)}{OSPUE} = \frac{OSPUE_L^{ILO}}{OSPUE}$

Finally, we compute the average ϕ^{ILO} for each country during 2004-2017 and assign this value to all years. We assign a country-specific but time-invariant value for two reasons. First, prior to 2004, the ILO lacks the required data to compute Y_L^{ILO} on a country-year basis. Second, when measured at the country-year level during the 2004-2017 period, ϕ^{ILO} varies little within country across years. Assigning a country-specific but time-invariant mixed income factor share may therefore be reasonable.

The main challenge is that the estimation framework for γ_q is not disciplined by the country's actual values in SNA. In particular, nothing prevents $\sum_q w_{emp} \cdot \gamma_q \cdot b_q > OS_{PUE}$ - such that estimated labor mixed income is larger than the SNA actually observed entire mixed income. This would, implausibly, imply that $\phi^{ILO} > 100\%$. To remedy this concern, we winsorize ϕ^{ILO} at 100%. In cases where γ_q and b_q are not from ILO (2019), we also winsorize ϕ^{ILO} from below with the lowest observed country value in ILO (2019), which is 36%. While the ILO (2019) method generates important country-level variation, the global average value for ϕ^{ILO}, at 80%, does not differ much from our benchmark value $\phi = 75\%$.

Mixed income in China and the US We make mixed-income adjustment to the benchmark series for China and the United States. For China, Piketty, Yang, and Zucman (2019) (PYZ) show that Chinese national accounts systematically underestimate mixed income and overestimate other factor incomes: for example, the income of self-employed agricultural workers is attributed to employee compensation in the SNA 2008 data and not to mixed income (as in other countries). We base our mixed income series on PYZ.

Following PYZ, we define mixed income as the sum of the income attributed to self-employed workers from agriculture and individual businesses. PYZ covers the period 1992-2014. For years before and after, we extend the series as follows:

(a) For agriculture, relevant data is available dating back to 1952. We extend the series back to 1965 relying on the price deflator available at World Inequality Database. For more recent years (2014-2018), we predict the trend based on sources used in PYZ (National Bureau of Statistics, link).

(b) For individual businesses, PYZ computes the income of this sector by combining several data sources. Unfortunately, a crucial part of it is not available prior to 1992, namely the 'flow of funds' data. Instead, our assumption is that, prior to 1980, Chinese individual businesses accounted for a negligible share of the economy. This observation is consistent with facts on self-employment structure in China at the micro and macro levels, and the trends presented in PYZ for the 1990s.[65] For recent years (2014-2018), we predict the trend based on sources used in PYZ (National Bureau of Statistics, link).

The estimated series of mixed income in China follows the same trend as for the rest of LMICs, although it starts from a slightly higher initial level.

For the US, we use the factor shares from Piketty et al. (2018), which (i) assumes a higher capital share of income for partnerships vs. other non-corporate businesses; and (ii) accounts for the rising capital intensity of partnerships since the 1980s.

[65] At the micro level, self-employed workers represent less than 2% of workforce in the 1980s, but had similar income per capita as wage earners (Gustafsson & Zhang, 2022). At the macro level, very small-scale industries represented 0.4% of industry output in the 1970s, reaching 7% only in 1989 (Yusuf, 1994).

Table B2: Main Tax Revenue and National Accounts Concepts

Panel A: Tax Revenue

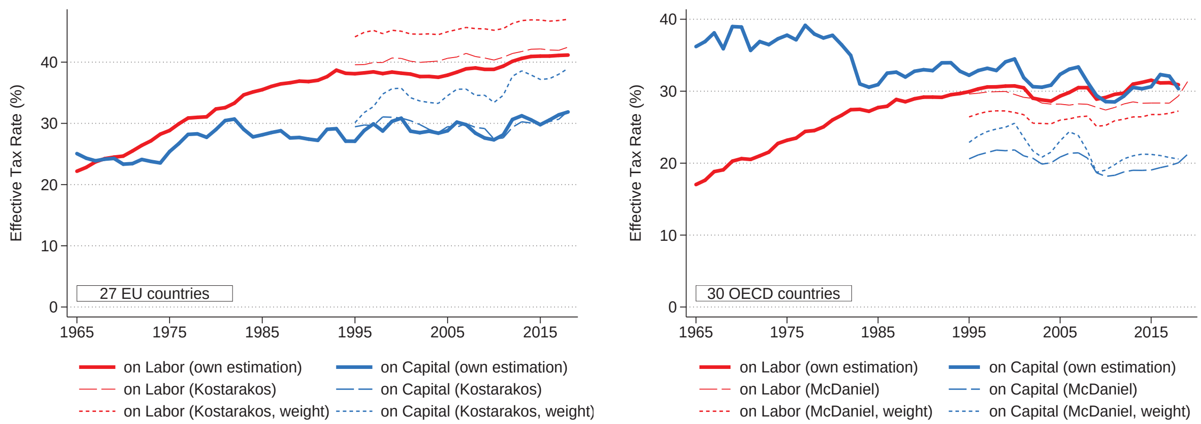
OECD revenue classification	type of tax j	incidence λ_j on labor	notes
1100	personal income tax (PIT)	$68\% \leq \lambda_{PIT} \leq 93\%$	Taxes on individuals (wages, capital income, capital gains). $\lambda_{PIT,ct}$ varies by country and year: see Section B.2 for details
1200	corporate income tax (CIT)	$\lambda_{CIT} = 0\%$	Taxes on corporate profits. Unallocable income taxes (OECD category *1300*) are split between PIT and CIT based on information from additional sources (see supplementary appendix)
2000 / 3000	social security & payroll	$\lambda_{soc.sec.} = 100\%$	Includes all social security contributions as well as payroll taxes
4000	property & wealth taxes	$\lambda_{assets} = 0\%$	Includes property, wealth and financial transaction taxes
5000	indirect taxes	excluded	Includes trade taxes, value-added taxes and other sales taxes and excise taxes. We consider these taxes as prior to factor income returns, such that they can be excluded from factor income taxation (Browning, 1978; Saez and Zucman, 2019).
6000	other taxes	excluded	Rare in occurrence and often quantitatively small
7000	non-tax revenue	excluded	Does not meet definition of taxation, can be quantitatively significant

Panel B: National Accounts

Natl. accounts acronym	national income component	benchmark allocation	notes
CE	compensation of employees	labor	Includes wages and salaries, employer and employee social contributions, and all payments from employers to their employees
OS_{PUE}	mixed income	$\phi = 75\%$ labor	'Operating surplus of private unincorporated enterprises' includes income from self-employment, household business owners, and informal or unincorporated enterprises
OS_{HH}	imputed rent	capital	'Operating surplus of households' is imputed rental income accruing to homeowners who live in their own home
OS_{CORP}	corporate profits	capital	'Operating surplus of corporations' includes all corporate income after paying employees and expenses, and can be thought of as corporate-sector capital income
OS_{GOV}	government operating surplus	—	$OS_{GOV} = 0$, by construction in national accounts
NIT	net indirect taxes	excluded	'indirect taxes, net of subsidies' usually comprise 8-15% of national income.
NFI	net foreign income	—	We treat domestic income without balancing the accounts to foreign earned income: many countries tax income earned domestically, regardless of citizenship, whereas net foreign income is taxed only with difficulty
CFC	depreciation	excluded	Factor income and our ETR are expressed net of 'consumption of fixed capital'

Figure B1: Comparing ETR Evolution in Our Data and Existing Studies

(a) Kostarakos and Varthalitis (2020) (b) McDaniel (2020)



Notes: These graphs provide a comparison of our ETR estimations with the recent literature. The left-hand graph compares our estimations with Kostarakos and Varthalitis (2020), based on EU-27 members from 1995 to 2019. The right-hand graph compares our estimations with the updated dataset in McDaniel (2020) that includes 30 OECD countries from 1995 to 2018. This extension is based on McDaniel (2007) (Table B3), and covers the largest OECD countries, including the US, as well as Mexico and Turkey. The solid line represents the results using our ETR measures and weights, but based on the exact country samples in the respective studies. The long-dash line replicates the ETR measures from the two studies. The short-dash line extends their ETR series but using our country-year weights. For a discussion of the differences between series, see Section 4.2, Table B3 and the supplementary appendix.

Table B3: Effective Tax Rates: Existing Databases

Paper	**Time**	**Countries**	**Source**	**Notes on methodological differences with our approach**
Mendoza et al (1994)	1965-1988	G7 members	OECD	Difference: All mixed income is allocated to capital income. Difference: Labor and capital in the PIT are taxed at the same rate
Carey and Rabesona (2004)	1975-2000	25 OECD biggest members	OECD	Difference: Mixed income allocation where self-employed pay themselves the annual salary earned by the average employee. Similarity: Labor and capital in PIT are not taxed at same rate, measure preferential tax treatment of pension funds and dividends. Difference: Social security contributions deducted from household income.
McDaniel (2007) (McDaniel 2020)	1950-2003 (updated: 1995-2018)	15 OECD biggest members (updated 30 OECD biggest members)	OECD	Difference: Mixed income imputed to capital based on rest-of-economy share. Difference: Labor and capital in PIT are taxed at the same rate
Kostarakos and Varthalitis (2020)	1995-2019	EU-27 members	Eurostat	Follows Carey and Rabesona (2004)

Appendix C Trade Liberalization Event Studies

C.1 Description of liberalization events

Our selection of trade events is determined by three criteria. First, the event is related to measurable policy reforms; this improves the transparency of the event-study design which is based on a well-defined policy event. Second, the policy reforms induced large changes in trade barriers; this increases the likelihood of observing sharp breaks in macroeconomic outcomes around the event-time. Third, the event has been studied in academic publications; this allows us to rely on events for which the positive effects on openness have previously been established.

These criteria led us to focus on the six trade liberalization events referenced in review articles by Goldberg and Pavcnik (2007, 2016) to which we add China's WTO accession event (studied in Brandt et al., 2017). Most of these selected events feature reductions in tariff rates: many of the countries did not participate in the early GATT/WTO negotiation rounds, making reductions in tariffs an available policy lever. The tariff reductions were large: Brazil cut tariff rates from 59% to 15%, India from 80% to 39% percent, and China from 48% to 20%. Mexico reduced tariff rates from 24% to 12% and import license requirements went from covering 93% of national production to 25%; Colombia's tariffs were reduced from 27% to 10% and import requirements dropped from 72% of national production coverage to 1%. In the selected countries, "tariff reductions constitute a big part of the globalization process" (Goldberg & Pavcnik, 2016). The timing of the events and academic references are provided in the supplementary appendix.

Below are narrative analyses for some of the events:

- **Brazil** The liberalization event of 1988 is detailed in Dix-Carneiro and Kovak (2017). The authors note: "In an effort to increase transparency in trade policy, the government reduced tariff redundancy by cutting nominal tariffs... Liberalization effectively began when the newly elected administration suddenly and unexpectedly abolished the list of suspended import licenses and removed nearly all special customs regimes."

- **Colombia** Similarly to Brazil, tariff reductions in Colombia in 1985 were driven by the country's decision to impose uniform rates across products and industries under the negotiation commitments to the WTO. Goldberg and Pavcnik (2007) note that this reform objective makes "the endogeneity of trade policy changes less pronounced here [in Colombia] than in other studies."

- **China** Brandt et al. (2017) note that trade openness reforms had gradually been implemented in China prior to the country's WTO accession in 2001, but that the tariff reductions implemented upon accession were large, "less voluntary" and largely complied with the pre-specified WTO accession agreements. Importantly, the potential accession to WTO contributed to the timing of privatization initiatives, in which the Chinese government restructured and reduced its ownership in state-owned enterprises. While the privatization efforts began in 1995 and were incremental, it is possible that additional sell-offs in the post-WTO years contribute to the observed medium-run trends in our outcomes.

- **India** The 1991 event in India occurred as a result of an IMF intervention that dictated the pace and scope of the liberalization reforms. Under the IMF program, tariff rates had to be harmonized across industries, which, like in Brazil and Colombia, led to a large average reduction in tariffs. Topalova and Khandelwal (2011) argue the Indian reform "came as a surprise" and "was unanticipated by firms in India." The reforms were implemented quickly "as a sort of shock therapy with little debate or analysis." The IMF program was in response to a set of events including "the drop in remittances from Indian workers in the Middle East, the increase in oil prices due to the Gulf War, and political uncertainty following the assassination of Rajiv Gandhi".

- **Vietnam** The 2001 reform was implemented as a broad trade agreement that did not involve negotiations over specific tariffs (McCaig & Pavcnik, 2018). The reform was driven by the American government's decision to reclassify Vietnam from 'Column 2' of the US tariff schedule to 'Normal Trade Relations'. Column 2 was designed in the early 1950s for the 21 communist countries, including Vietnam, with whom the US did not have normal trading relations.

These descriptions of reform timing do not suggest that the liberalization events were directly triggered by changes in domestic taxation or factor incomes.

Goldberg and Pavcnik (2007) note other cross-border reforms that occurred during post-years of the liberalization events. Argentina's 1989 event and Brazil's 1988 event were followed by accession to Mercosur in 1991; India's 1991 event was followed by foreign direct investment liberalization in 1993; and Mexico's 1985 WTO accession was followed by a removal of capital inflow restrictions in 1989. These reforms occurred with some lag to the trade liberalization events.

C.2 Event study methodology

Our sample is constructed by applying a synthetic matching procedure to every treated country for each outcome of interest. The donor pool has to be fully balanced in all pre-event periods. To estimate the event study in equation (4) for a given outcome, the sample pools the seven treated countries and their synthetic control countries for 10 years before and after the events (yielding 294 observations). We estimate the event-study in equation 4 and the DiD model: $y_{ct} = \beta^{DiD} \cdot \mathbb{1}(e \geq 0)_t \cdot D_c + \theta_t + \kappa_c + \pi_{Year(t)} + \epsilon_{ct}$. The DiD model uses the same notation as equation (4). Moreover, we use the imputation method by Borusyak et al. (2021) to report average treatment effects comparable to β^{DiD} with a technique that deals with issues with two-way fixed effects and heterogeneous event timing. Details are provided in the supplementary appendix. All the DiD average treatment effects are reported in Table A1. We test if our results hold with a more restrictive synthetic control, by using our three main outcomes—trade, ETR_K and ETR_L— to construct one synthetic control group per treated country. The results are reported in Panel B of Table A1.

C.3 Alternative trade liberalization event study

We present results based on an alternative measure of trade liberalization events. We use the events from Wacziarg and Welch (2008), which cover 141 countries at all levels of development between the 1950s and 1998. When merged with our data, the sample covers

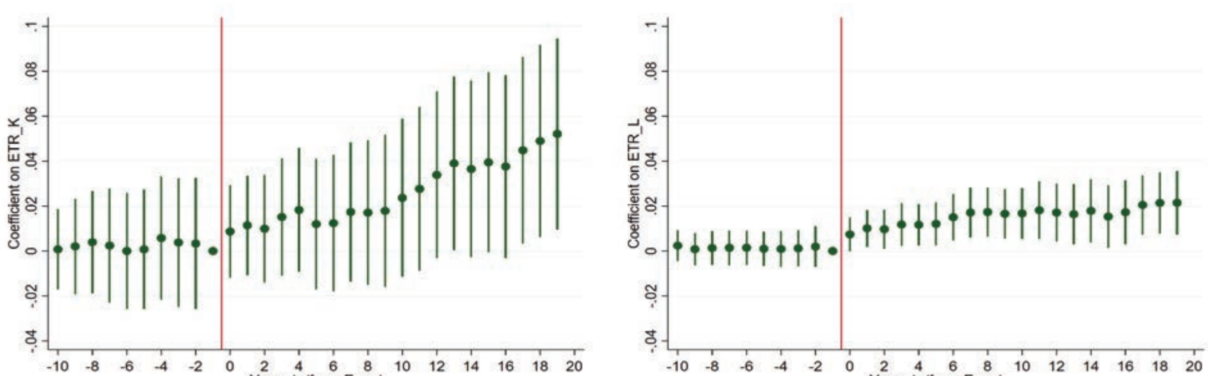
68 liberalization events that occurred between 1965 and 1998 in developing countries. A trade liberalization event is defined to occur when all five of the following conditions no longer hold: (i) average tariff rates are above 40%; (ii) non-tariff barriers cover at least 40% of trade; (iii) the black market exchange rate is at least 20% lower than the official exchange rate; (iv) there is a state monopoly on major exports; (v) there is a socialistic system in place. These conditions are broader than our main liberalization event criteria (Section 5.1 and C.1). At the same time, our main events are covered in this expanded event sample (with the exception of China and Vietnam, whose events are after the end of the sample period); this occurs because the reduction in tariff rates, one of our main event criteria, was the remaining event-condition to be satisfied in Wacziarg and Welch (2008). We estimate the effects of the liberalization events using the DiD model: $y_{ct} = \beta^{DiD} \cdot E_{ct} + \theta_t + \theta_c + \epsilon_{ct}$. y_{ct} is the outcome of interest in country c in year t, E_{ct} is the event indicator which takes on a value of 1 in all periods after a country has a liberalization event (and 0 otherwise), and θ_t and θ_c are year and country fixed effects, respectively. ϵ_{ct} is clustered at the country level. Estimation issues arising from heterogeneous treatment-timing may be important; for this reason, we focus on the imputed treatment effects based on Borusyak et al. (2021). We restrict the sample to developing countries between 1965 and 2008.

Panel A of Table C1 reports the β^{DiD} impacts on trade, ETR_K and ETR_L. Despite being based on broader criteria, the trade liberalization events produce qualitatively similar results to the main event-study (Section 5.1), with positive impacts on openness and both ETRs, and a larger magnitude-impact on ETR_K than ETR_L. Figure C1 estimates the dynamic event-study. Liberalized and control countries are on parallel trends until the event onset; both ETRs start to increase in the immediate post-event years. Panel B shows that the results are robust to estimating the effects in a fully balanced panel 10-years post-reform. In Panel C, the results hold when the control group is formed within-region. Panel D shows the results are robust to excluding countries which have cross-border capital liberalization events at any point during the sample-period (Bekaert, Harvey and Lundblad, 2000). Finally, Panel E shows the results hold when we exclude countries with concurrent domestic reforms (Wacziarg & Wallack, 2004).

Figure C1: Event-Study of Trade Liberalization Based on Wacziarg & Welch (2008)

(a) Effective tax rate on capital

(b) Effective tax rate on labor



Notes: These graphs show event-study impacts of the trade liberalization events from Wacziarg and Welch (2008) on ETR_K (left panel) and ETR_L (right panel).

Table C1: Trade Liberalization Event-Study Based on Wacziarg & Welch (2008)

	Trade	ETR_K	ETR_L
	(1)	(2)	(3)
Panel A: *Benchmark*			
Post*Treat	0.030	0.021	0.006
	(0.048)	(0.017)	(0.006)
Imputed treatment effect	0.090*	0.043**	0.021***
	(0.049)	(0.016)	(0.005)
N	4032	4032	4032
Panel B: *Fully balanced panel, 10-year post-reform*			
Imputed treatment effect	0.110**	0.031**	0.018***
	(0.054)	(0.014)	(0.005)
N	3082	3082	3082
Panel C: *With region-year fixed effects*			
Imputed treatment effect	0.084**	0.042**	0.021***
	(0.041)	(0.016)	(0.005)
N	4032	4032	4032
Panel D: *Excluding countries with capital liberalization*			
Imputed treatment effect	0.101*	0.028*	0.014**
	(0.057)	(0.017)	(0.006)
N	2651	2651	2651
Panel E: *Excluding countries with domestic reforms*			
Imputed treatment effect	0.056	0.040**	0.015***
	(0.051)	(0.016)	(0.005)
N	3551	3551	3551

Notes: This table shows the results from estimating the difference-in-difference regression and the imputed treatment effect of the 68 trade liberalization events from Wacziarg and Welch (2008), between 1965 and 2008. The sample is low and middle-income countries, based on the World Bank income classification in 2018. In Panel A, the post*treat coefficient corresponds to the β^{DiD} based on estimating the equation in Section C.3. The imputed treatment effect is based on the method in Borusyak, Jaravel, and Spiess (2021). In Panel B, the sample is restricted to the fully balanced set of countries in the 10 years after the liberalization event. In Panel C, the estimation is augmented with region-by-year interactive fixed effects. In Panel D, the sample excludes all countries that have a capital liberalization reform at any point during the sample-period, based on Bekaert, Harvey and Lundblad (2000). In Panel E, the sample excludes all countries with domestic reforms which coincide in timing with their trade liberalization event, based on Wacziarg and Wallack (2004). Standard errors are clustered at the country level. For more details on the liberalization events, see Appendix C.3. * p<0.10 ** p<0.05 *** p<0.01.

Appendix D Results on Tax Capacity Mechanism

D.1 Firm-level analysis in Rwanda

Data-sources and sample Our analysis draws on three administrative datasets from the Rwanda Revenue Authority (RRA), for the years 2015-2017. These data sources can be linked through unique tax identifiers for each firm, assigned by the RRA for the purpose of collecting customs, corporate income and value-added taxes. The first data source is the customs records, which contain information on international trade transactions made in each year by each firm. We use this data to measure each firm's direct imports. The second data is the firms' corporate income tax (CIT) declarations merged with the firm registry. These data contain detailed annual information on firms' profits, revenue and costs. We use these data to measure each firm's effective tax rate. The third data source is the business-to-business transactions database. These data are retrieved through the electronic billing machines (EBM) that all firms registered for VAT are legally required to use (Eissa and Zeitlin, 2014). For a given seller, EBMs record the transactions to each buyer identified by the tax firm-ID. We use this data to measure buyer-seller relationships.

When combined, these data allow us to construct the buyer-supplier relationships of the Rwandan formal economy and document firms' total trade exposure. Importantly, since the network data is based on tax-IDs, we cannot observe transaction linkages with informal, non-registered firms. This sample selection on formal firms also features in most recent network studies, by virtue of relying on administrative data, including in Chile (Huneeus, 2020); Costa Rica (Alfaro-Ureña et al.); Ecuador (Adao et al., 2022); India (Gadenne et al., 2022); Turkey (Demir et al., 2021); and Uganda (Almunia et al., 2023).

Our sample is the set of firms that are registered for CIT and that report positive income during the years 2015-2017. Note that only a small number of firms are registered for CIT or VAT but not both, meaning that the overlap with the EBM transactions data is strong. However, restricting the sample to positive income is consequential, as a significant number of registered CIT firms are 'nil filers' that report zero income ('nil filers' are common in developing countries: Keen, 2012). We measure each firm i's yearly effective tax rate on corporate profits, corresponding to corporate ETR_i^K in equation (6), as the ratio of corporate taxes paid divided by net profit. Net profit is revenue minus material, labor, operational, depreciation and financial costs.

The EBM data is meant to improve the enforcement of corporate taxes and VAT, and the reporting of linkages is more comprehensive for the relatively larger firms that are registered for these tax bases. For smaller incorporated firms that are instead registered to simplified tax bases (flat-amount or turnover), only a few of them are registered for VAT. Consequently, these firms are most likely to be recorded in the EBM data as clients in a particular transaction, making the coverage of their linkages less comprehensive. It is in principle also possible to measure ETR_i^K amongst these smaller, incorporated firms. However, the information on their tax returns regarding cost items is less detailed and additional assumptions on the relationship between turnover and profit are required, which makes the profit measure in the denominator of ETR_i^K less precise. With these data-challenges in mind, we can include these additional tax-registered firms in the analysis; we find qualitatively similar results (available upon request).

Exposure to trade To measure a firm's total exposure to trade, we follow Dhyne et al. (2021) who use similar administrative datasets as ours to measure trade exposure of Belgian firms. We define firm i's total foreign input share as the share of inputs that it directly imports (s_{Fi}), plus the share of inputs that it buys from its domestic suppliers l (s_{li}), multiplied by the total import shares of those firms:

$$s_i^{Total} = s_{Fi} + \sum_{l \in V_i} s_{li} \cdot [s_{Fl} + \sum_{r \in V_l} s_{rl} \cdot (s_{Fr} + ...)]$$

where V_i is the set of domestic suppliers of firm i, and V_l is the set of domestic suppliers of firm l. The denominator of the input shares is the sum of purchases from other firms and imports. Note that s_i^{Total} is recursive: a firm's total foreign input share is the sum of its direct foreign input share and the share of its inputs from other firms, multiplied by those firms' total foreign input shares. We limit the calculation to the inputs from a firm's immediate suppliers l as well as the suppliers to their suppliers r (adding more network-levels only marginally increases s_i^{Total}). s_i^{Total} reflects the direct import share of firm i's suppliers and the suppliers' suppliers, each weighted by the share of inputs that each firm buys from other domestic firms. We focus on firms' exposure to imports through their supplier network; in an extension, we find qualitatively similar results when studying firms' exposure to exports through their client network (results available).

Figure D1 displays a histogram of s_i^{Total} and s_{Fi} for all formal Rwandan firms. While just under 30% of firms import directly, 93% rely on trade either directly or indirectly through their suppliers. In the median firm, the total foreign input share is 48% (it is 39% for the median Belgian firm in Dhyne et al., 2021).

Impacts of trade exposure on ETR^K and size To visualize the association between trade exposure (s_i^{Total}) and ETR_i^K, we plot binned scatters of the variables against each other, after residualizing both against year fixed effects. In Figure D2, the dots correspond to equal-sized bins of the residualized trade variable. The line corresponds to the best linear fit regression on the underlying firm-level data ($N = 18478$). Figure D2 reveals a positive and strongly significant association: firms that are more exposed to international trade, both through direct imports and through links to importers in the supply network, have higher effective tax rates on corporate profits.

We investigate this association in a regression form in Table 4, deploying both OLS and IV. The IV applies the design in Dhyne et al. (2021) that extends the shift-share approach of Hummels et al. (2014) to a setting with shock pass-through via network linkages. The empirical strategies and the main results are described in Section 6.3.

In additional regressions (not shown but available), we find that the results are robust to controlling for trade shocks to firm i's potential suppliers (firms that operate in the same industry and geographical area as i's current suppliers but are not currently supplying to i) and firm i's horizontal suppliers (firms that are suppliers to firm i's current clients).

We focus on firms' exposure to imports through their supply network, but firms may also be impacted by imports through their clients. In an extension, we find that increased output exposure to imports through the client network has positive effects on ETR^K (results available), though this average effect could mask heterogeneity across firms depending on the complementarity between imports and domestic inputs.

Because the estimation is within the corporate sector, this exercise cannot speak to the magnitude of trade's net impact on sector-level $\overline{ETR}_C^K$. These firm-level results on corporate ETR_i^K are therefore complementary to the country-level results on $\overline{ETR}_C^K$. An additional limitation is that the network linkage measures are derived from administrative data which, by construction, only exist for tax registered firms (Atkin & Khandelwal, 2020). This sample restriction implies that this firm-level regression is not suited to study the impacts of trade on the size of informal firms.

D.2 Type of trade analysis

We investigate whether trade has differential impacts on ETR and mechanism outcomes depending on the nature of the trade variation (Section 6.4). We use our two instruments to investigate the impacts of: (i) imports versus exports (of trade in both intermediate G-S and final G-S); (ii) trade in intermediate G-S versus final G-S (summed across imports and exports). We use UN's Broad Economic Categories (Rev. 5) to classify final versus intermediate goods-services (G-S), combining capital goods with the latter. For the imports versus exports IV, the two 1^{st}-stage regressions are

$$log(imp_{ct}) = \beta_1 \cdot Z_{ct}^{gravity} + \beta_2 \cdot Z_{ct}^{oil-dist} + \mu_c + \mu_t + \epsilon_{ct}$$
$$log(exp_{ct}) = \pi_1 \cdot Z_{ct}^{gravity} + \pi_2 \cdot Z_{ct}^{oil-dist} + \eta_c + \eta_t + \iota_{ct}$$

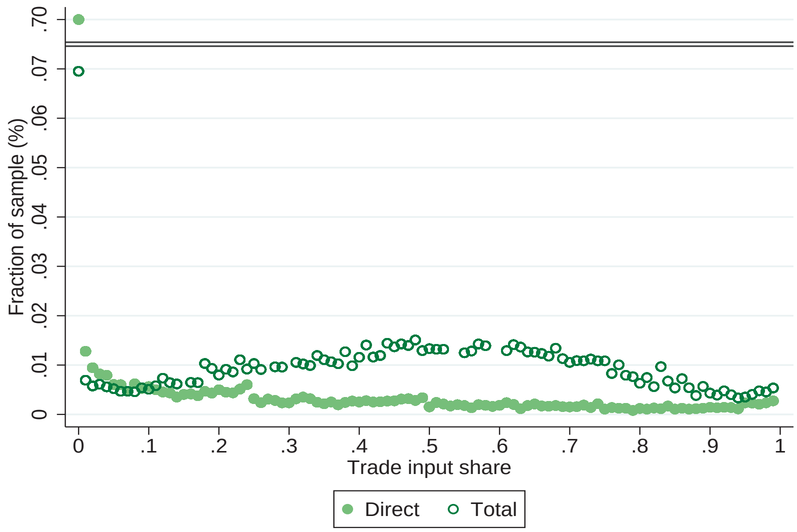
where $log(imp_{ct})$ and $log(exp_{ct})$ are the logs of total imports to NDP and total exports to NDP, respectively, in country c in year t. The log-transformation improves the 1^{st}-stage (results without logs are qualitatively similar). The 2^{nd}-stage is

$$y_{ct} = \theta_1 \cdot log(imp_{ct}) + \theta_2 \cdot log(exp_{ct}) + \kappa_c + \kappa_t + \phi_{ct}$$

The set-up is similar for the second IV (intermediate G-S vs final G-S) where we replace $log(imp_{ct})$ and $log(exp_{ct})$ with the log of total trade in intermediate G-S to NDP and the log of total trade in final G-S to NDP. IV results for developing countries are in Panel A of Table D1, with 1^{st}-stage regressions in Panel B. Note that it is ex ante unclear if the two instruments generate a strong overall first-stage. We gauge this by inspecting the Kleibergen-Paap F-statistics, which are not well above conventional threshold levels (13.56 and 8.21). Given this challenge, we limit our scope to studying whether the coefficient signs for the different types of trade are consistent with our simplified predictions (and whether they statistically differ from each other). The exclusion restriction requires that the regressors add up to total trade openness. For this reason, we cannot implement an IV which focuses on the impacts of final versus intermediate G-S for, say, imports only. This also implies that, for a given outcome, the hypotheses in our two IVs (final versus intermediate G-S; imports versus exports) will be correlated. We accordingly adjust the p-values for multiple hypotheses testing using the Romano-Wolf method.

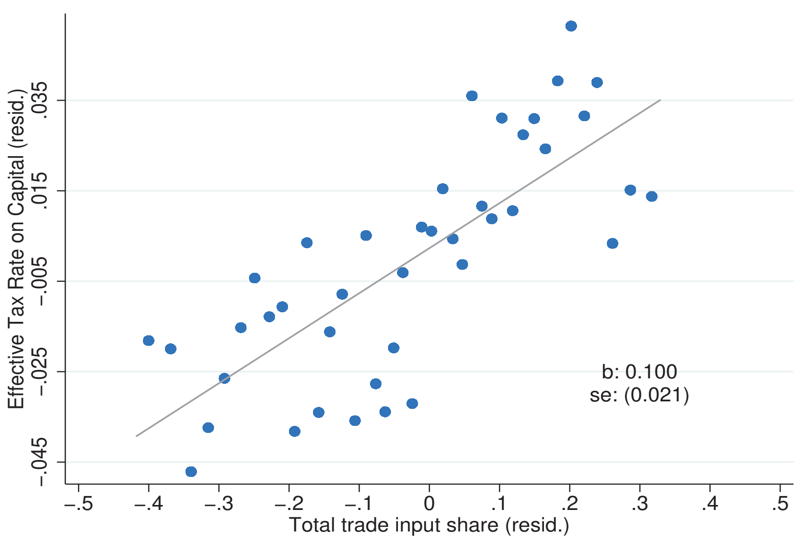
The results are described in Section 6.4. Since we only have 2 instruments, we cannot decisively conclude on the impacts for the 4 types of trade (imports of intermediate G-S, exports of intermediate G-S, imports of final G-S, exports of final G-S). Notwithstanding, the estimated IV coefficients are consistent with imports of final G-S decreasing ETR_K and mechanism outcomes (μ_C,$\overline{ETR}_C^K$), and imports of intermediate G-S increasing them.

Figure D1: Rwandan Firms' Direct and Total Exposure to Trade in Imports



Notes: This figure shows the distribution of direct foreign input share, s_{Fi}, and total foreign input share, s_i^{Total}, for all corporate firms in Rwanda between 2015 and 2017. The measures are calculated annually, and the figure pools all firm-year observations. The horizontal line represents a scale break in the vertical axis. More details are in Section D.1.

Figure D2: Rwandan Firms' Trade Exposure and Corporate Effective Tax Rate



Notes: This figure shows the firm-level association between total foreign input share, s_i^{Total}, and the corporate effective tax rate for all corporate firms in Rwanda between 2015 and 2017. The graph plots binned scatters of the variables against each other, after residualizing both variables against year-fixed effects. The dots correspond to equal-sized bins of the residualized trade exposure variable. The line corresponds to the best linear fit regression on the underlying firm-level data ($N = 18478$), which is also reported in column (1) of Table 4.

Table D1: Type of Trade Analysis in Developing Countries

Panel A: IV	ETR_K		ETR_L		Corporate Income		Mixed Income		Corporate ETR_K	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Export of G-S	0.487*		0.225**		0.214*		-0.159*		0.611*	
	(0.263)		(0.096)		(0.123)		(0.091)		(0.339)	
	[0.066]		[0.019]		[0.039]		[0.119]		[0.076]	
Import of G-S	-0.358***		-0.184***		-0.126*		0.069		-0.442***	
	(0.126)		(0.044)		(0.074)		(0.049)		(0.158)	
	[0.059]		[0.013]		[0.045]		[0.145]		[0.033]	
Intermediate G-S		0.303***		0.133***		0.147**		-0.119***		0.385***
		(0.095)		(0.038)		(0.070)		(0.045)		(0.122)
		[0.053]		[0.012]		[0.033]		[0.048]		[0.031]
Final G-S		-0.245***		-0.125***		-0.089**		0.050**		-0.302***
		(0.051)		(0.023)		(0.044)		(0.024)		(0.056)
		[0.013]		[0.006]		[0.019]		[0.119]		[0.006]
F-test: Equality of coefficients [p-value]	4.82	14.78	8.55	19.06	2.73	4.33	2.55	5.98	4.60	15.35
	[0.030]	[0.004]	[0.004]	[0.001]	[0.096]	[0.039]	[0.113]	[0.016]	[0.034]	[0.000]
N	4572	4572	4572	4572	4572	4572	4572	4572	4572	4572

Panel B: 1st-stage	Import of G-S	Export of G-S	Intermediate G-S	Final G-S
	(1)	(2)	(3)	(4)
$Z^{gravity}$	0.287***	0.252***	0.282***	0.268***
	(0.034)	(0.060)	(0.034)	(0.052)
$Z^{oil-distance}$	-0.077***	0.003	0.008	-0.116***
	(0.011)	(0.018)	(0.013)	(0.019)
1st-stage F-statistic	134.47	15.75	54.76	75.85
1st-stage Sanderson-Windmeijer Weak Instrument F-statistic	36.49	34.02	65.33	70.59
1st-stage Kleibergen-Papp F statistic	8.21		13.56	
N	4572	4572	4572	4572

Notes: The sample is developing countries, which are low and middle-income countries according to the World Bank income classification in 2018. Panel A presents IV results, while Panel B presents 1st-stage results. In Panel A's odd-numbered columns, imports and exports are the regressors while in even-numbered columns it is trade in intermediate goods and services (G-S) and trade in final G-S. Outcomes differ across columns in Panel A: in cols. (1)-(2), effective tax rate on capital, ETR_K; in cols. (3)-(4), effective tax rate on labor, ETR_L; in cols. (5)-(6), corporate income share of net domestic product; in cols. (7)-(8), mixed income share of net domestic product; in cols. (9)-(10), average effective tax rate on corporate profits. For details on the outcomes and the instruments, see Table 1 and 3. Relative to those tables, the drop in sample size in this table is due to availability of the type of trade classification. For each coefficient, we report in brackets the p-values which correct for multiple hypotheses testing, using the Romano-Wolf method. Multiple hypothesis testing is accounted for within each outcome between the two IV estimations (exports and imports; final G-S and intermediate G-S). At the bottom of each column in Panel A, we report the F-test for the equality of coefficients. In Panel B, cols. (1)-(2) correspond to the first-stage regression that instruments simultaneously for imports and exports; cols. (3)-(4) is the first-stage regression which instruments simultaneously for intermediate G-S and final G-S. In Panel B, we report the F-statistic of excluded instruments; the Sanderson-Windmeer multivariate F-test of excluded instruments; and, the Kleibergen-Paap F-statistic. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level. For more details, see Section D.2.

Appendix E Capital Liberalization Events

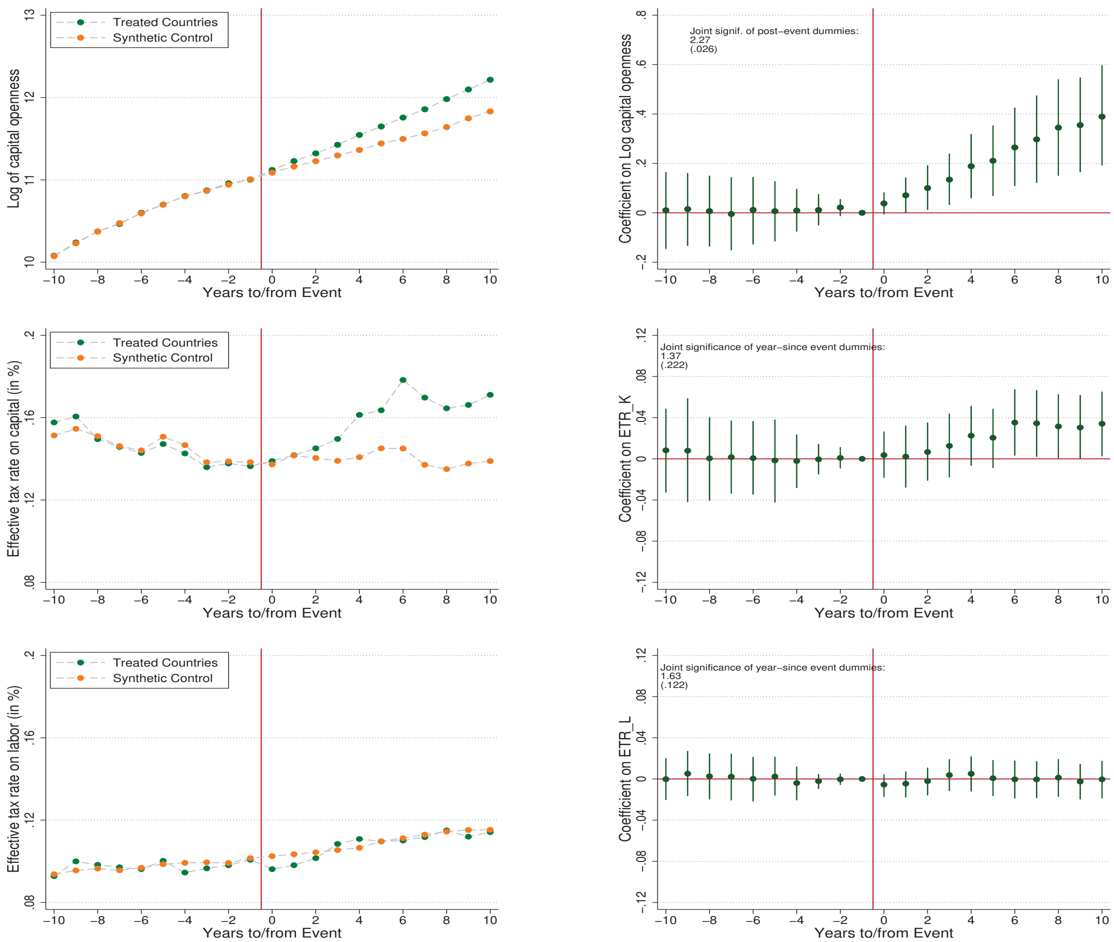
To attempt to investigate the impact of capital liberalization on effective tax rates, we draw on Chari et al. (2012). The authors measure capital liberalization events in 25 developing countries as the date when foreign investment in the domestic stock market was first allowed. They show that these events significantly increase foreign capital inflows, including foreign direct investment (FDI) and import of capital goods.[66] Compared to other policies aimed at lifting FDI restrictions, liberalizing the domestic stock market occurs at a precise point in time, is not marked by policy-reversal or net capital outflow, and is unambiguously related to capital liberalization (Eichengreen, 2001). We employ the empirical design of Section 5.1 and create a synthetic control country for each of the 25 treated countries and for each outcome. We measure capital openness as the total sum of the stocks of foreign assets and liabilities (Gygli et al., 2019). We find similar results when using alternative measures of capital openness, including portfolio equity assets and liabilities and the KOF financial globalization index (Gygli et al., 2019).

Figure E1 reports the event-study results. Relative to a stable pre-trend, we observe a sustained rise in capital openness precisely at the time of the event. ETR_K also increases, with a small lag to the timing of the capital liberalization event; in the medium-run, the positive effect on ETR_K is significant at the 5% level. There is no discernible effect on ETR_L. Similar to the reasoning for the trade tax-capacity mechanism, the inflow of foreign capital, as well as any subsequent increase in capital goods imports and aggregate investment, may positively impact ETR_K by contributing to general growth of firms or by causing an expansion of initially larger firms. Consistent with this interpretation, we find that the capital liberalization events led to increases in the corporate output share and the average corporate effective tax rate (results not shown but available).

One important limitation is that the events considered here remove restrictions on capital *inflows* and are not informative of the impacts of increased capital *outflows*. In general, more work is needed to understand the determinants of policies that impact cross-border capital flows in developing countries and their effects on ETRs.

[66] FDI includes green field investments (building plants from scratch) and cross-border mergers and acquisitions (M&A). Chari et al. (2012) note that M&A is impacted by stock market liberalization, makes up to 40-60% of FDI in developing countries, and can trigger subsequent green field investments.

Figure E1: Event Study of Capital Liberalization Reforms



Notes: These panels show event-studies for capital liberalization reforms in the 25 developing countries of Chari, Henry, and Sasson (2012). The panels correspond to different outcomes: capital openness (top panels); effective tax rate on capital (middle panels); effective tax rate on labor (bottom panels). Capital openness is the total sum of the stocks of foreign assets and liabilities, in constant USD. We use the log transformation for this outcome; results where the total sum is expressed as a percent of GDP are similar. The left-hand graphs show the average level of the outcome in every year to/since the event, for treated countries and for synthetic control countries. The right-hand graphs show the estimated β_e coefficients on the to/since dummies, based on equation (4) but where the trade liberalization events are replaced with capital liberalization events. The bars represent the 95% confidence intervals. Standard errors are clustered at the country level and estimated with the wild bootstrap method. The top-left corners report the F-statistic on joint significance of the post-event dummies, with the p-value in parentheses. Details are in Appendix E.

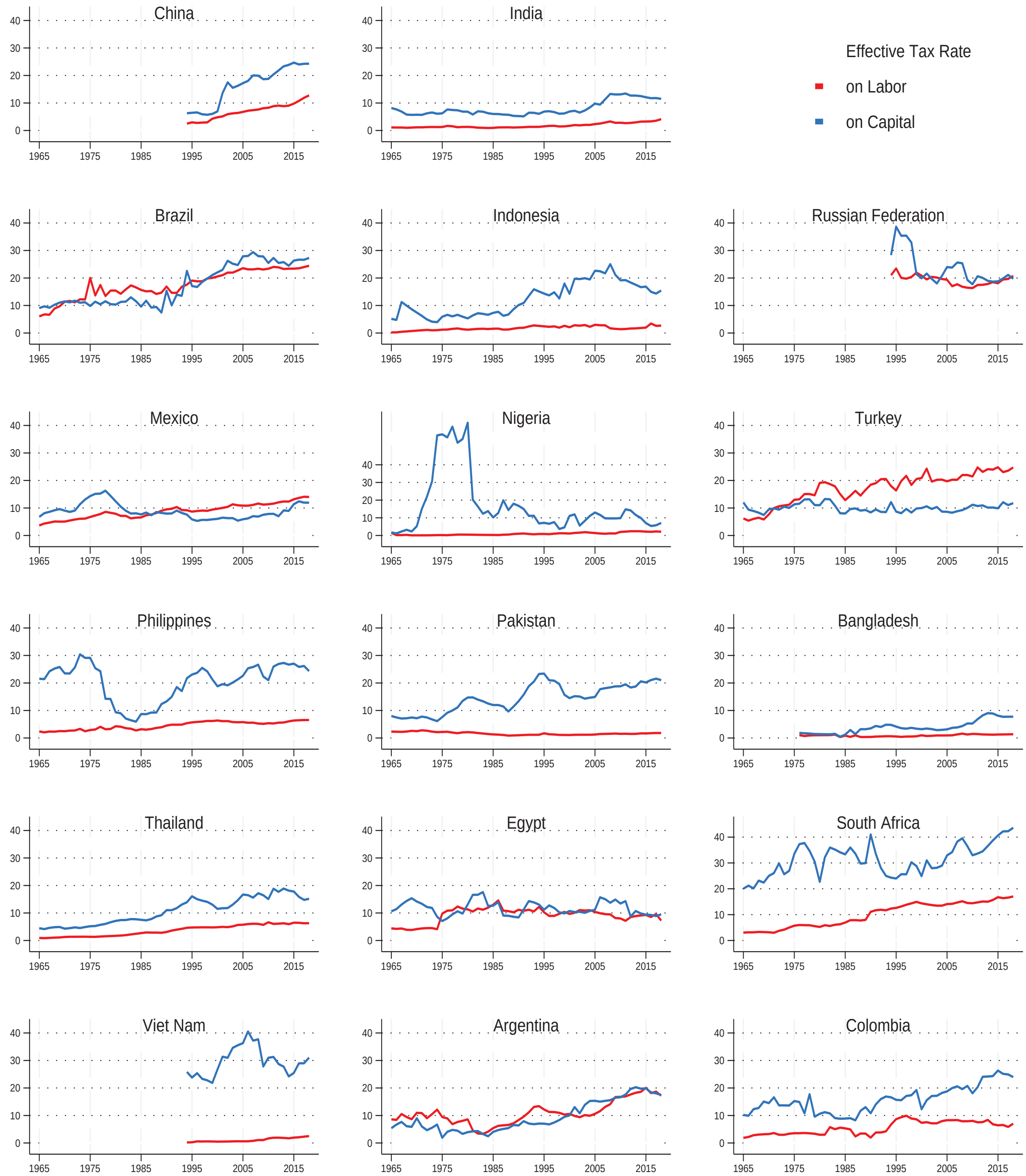
Appendix

Online Supplementary Appendix, Not for Publication "Capital Taxation, Development and Globalization: Evidence from a Macro-Historical Database"

(Please note: This appendix contains the same material as the 108-page online supplementary appendix, including the case-studies report, circulated under the title "Globalization and Factor Income Taxation")

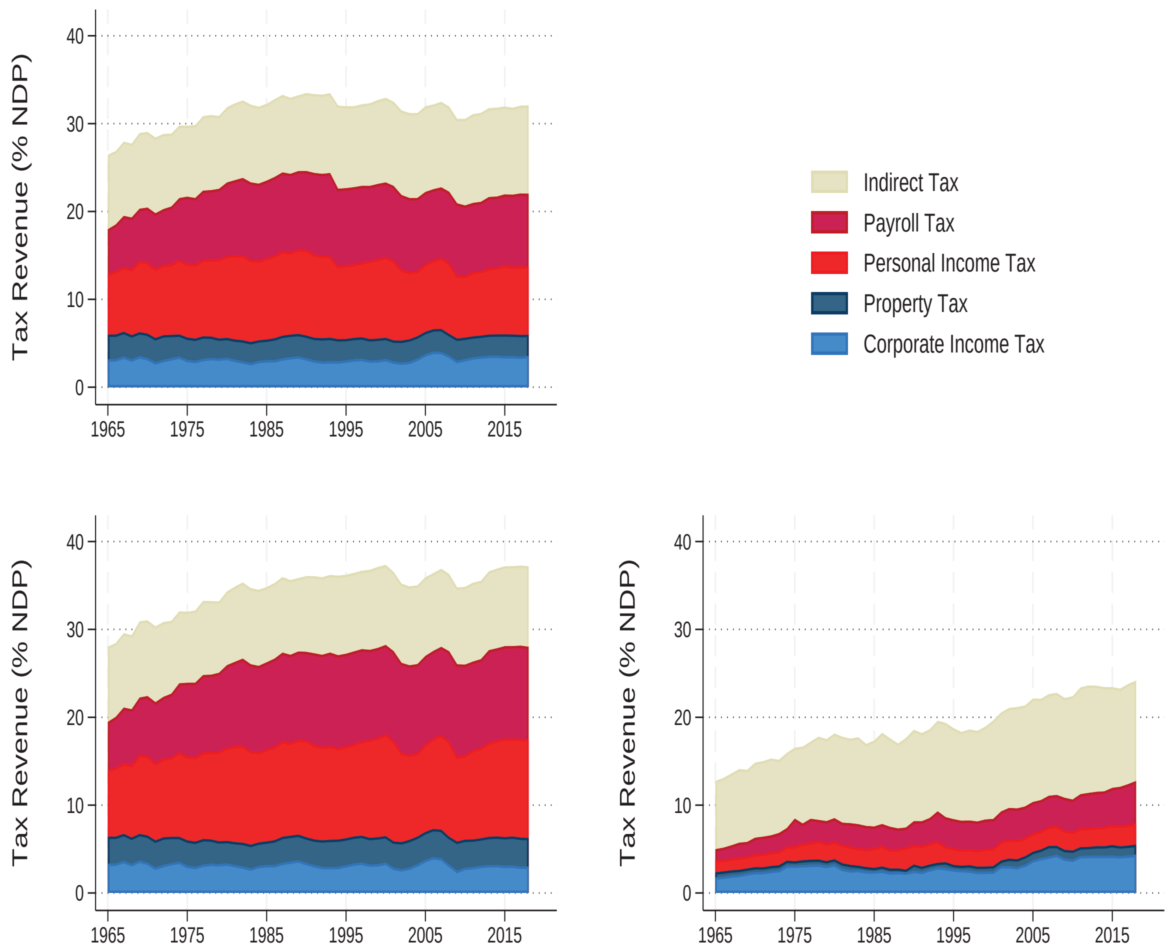
by Pierre Bachas, Matthew-Fisher Post, Anders Jensen and Gabriel Zucman

Figure O1: Effective Tax Rates in Large Developing Countries



Notes: This figure shows the evolution of effective tax rates on labor and capital for the 17 largest low and middle-income countries. Countries are displayed when they rank in the top 20 both in terms of population and Net Domestic Product (NDP) in 2018. Low and middle-income countries are based on the World Bank income classification in 2018.

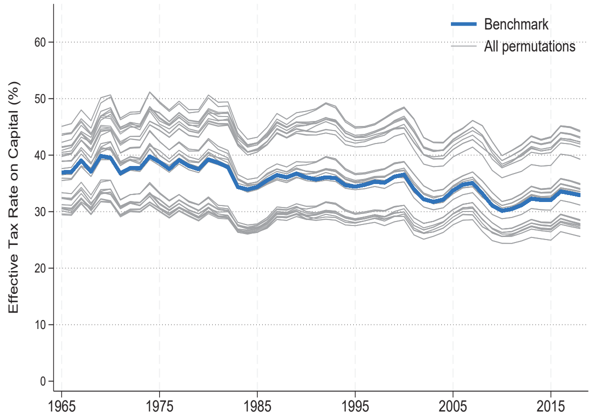
Figure O2: Tax Revenue as a Share of Net Domestic Product



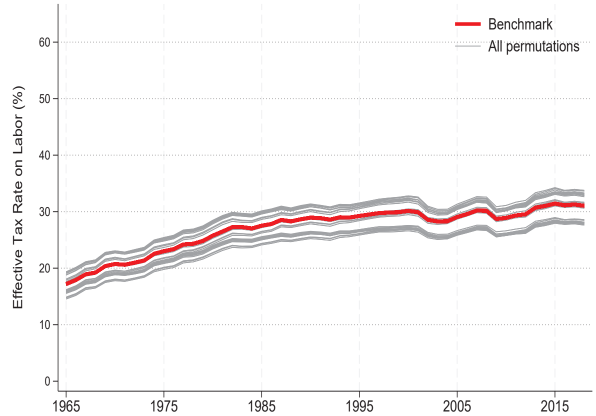
Notes: This figure plots the time series of tax revenue as a share of net domestic product (NDP), separated into five revenue sources. The top left panel corresponds to the global average, weighting country-year observations by their share in that year's total NDP, in constant 2019 USD (N=154). The bottom-left panel shows the results for high-income countries, and the bottom right for low- and middle-income countries. Low, middle and high-income countries are based on the World Bank income classification in 2018. Tax revenues are separated into five main categories: indirect taxes (domestic consumption taxes and international trade taxes), payroll and social security taxes, taxes on personal income, taxes on property and wealth, and taxes on corporate income. The dataset is composed of two (quasi) balanced panels: the first covers the years 1965-1993 and excludes communist regimes. The second covers 1994-2018 and integrates former communist countries, in particular China and Russia.

Figure O3: Robustness of ETR_K and ETR_L Trends by Development Levels

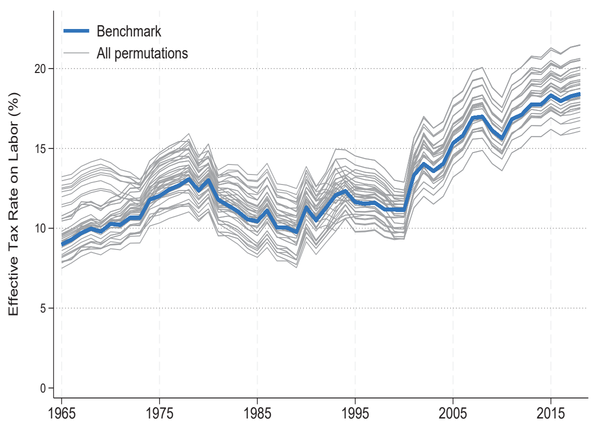
(a) ETR_K in HICs



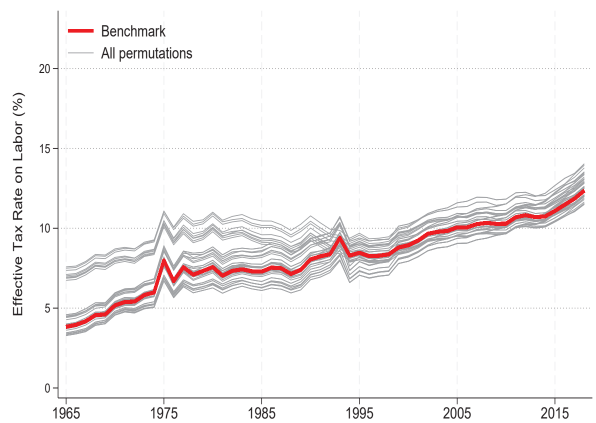
(b) ETR_L in HICs



(c) ETR_K in LMICs



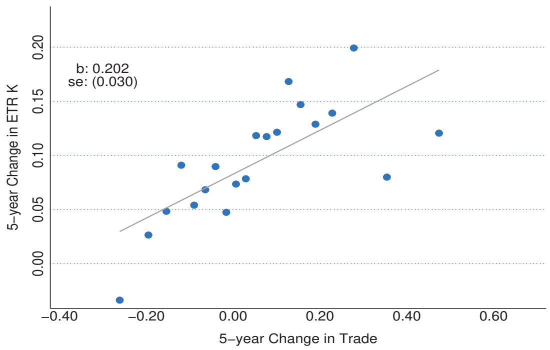
(d) ETR_L in LMICs



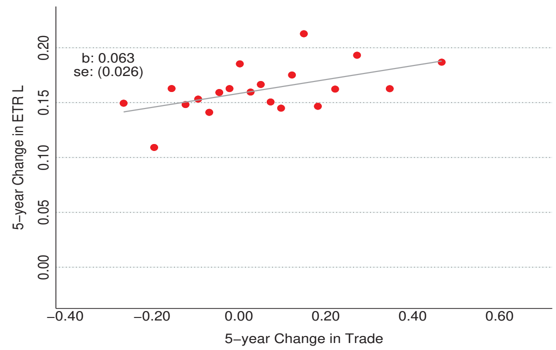
Notes: These panels show trends in the effective taxation of capital and labor for high-income countries (HICs, top panels) and low and middle-income countries (LMICs, bottom panels). Low, middle and high-income countries are based on the World Bank income classification in 2018. The benchmark series are denoted by the thick colored lines and the grey lines denote all 54 possible permutations of the series when varying the four key methodological choices (detailed in Section 4.2): the allocation of personal income tax revenue to capital vs labor; the allocation of mixed income to capital vs labor; presenting results for an unbalanced panel of countries vs a balanced panel via imputations; and, how to weight individual countries' series when aggregating them. Panel (c) corresponding to the ETR_K for low and middle-income countries is further decomposed in Figure 3.

Figure O4: Associations between ETR and Trade, Conditional on GDP per Capita

(a) ETR_K: Low & middle-income

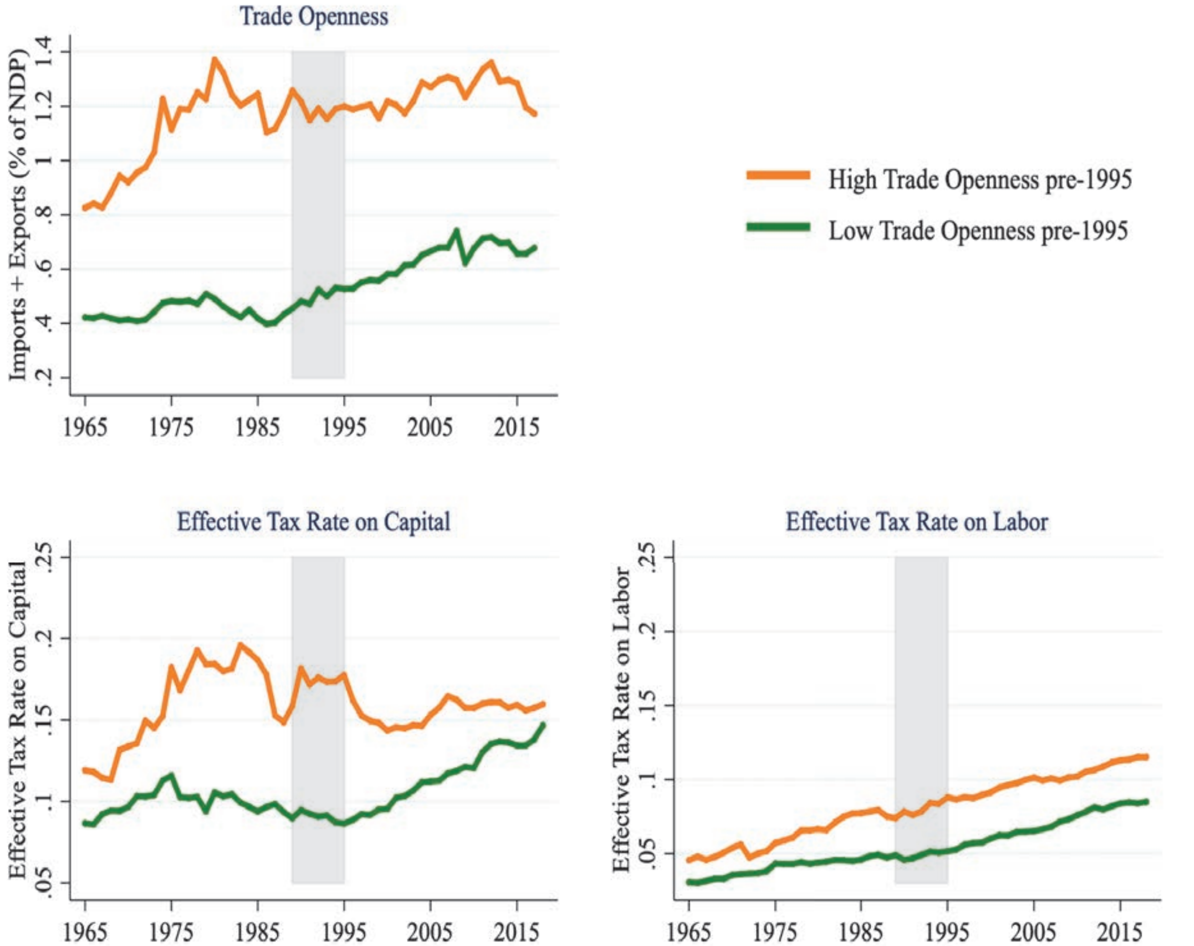


(b) ETR_L: Low & middle-income



Notes: These panels show the association between trade and effective tax rates, conditional on GDP per capita, in developing countries. The panels are created exactly in the same way as in Figure 5, except the variables are additionally residualized against the growth rate in GDP per capita. Developing countries include low-income and middle-income countries, where categories are based on the World Bank income classification in 2018. The outcome is the effective tax rate on capital, ETR_K, and on labor, ETR_L, in the left-side and right-side panels, respectively. Trade is measured as the sum of import and exports as a share of net domestic product. Both the x-axis and y-axis are measured as within-country percent changes over 5 years. Each graph shows binned scatter plots of each outcome against trade, after residualizing all variables against year fixed effects and the 5-year growth rate in GDP per capita. Each dot corresponds to a ventile (20 equal-sized bins) of the residualized trade variable, with average values of trade and ETR calculated by ventile. In each graph, the line represents the best linear fit based on the underlying country-year data, with the corresponding slope-coefficient and standard error reported in the top-left corner. For more details, see Section 4.4.

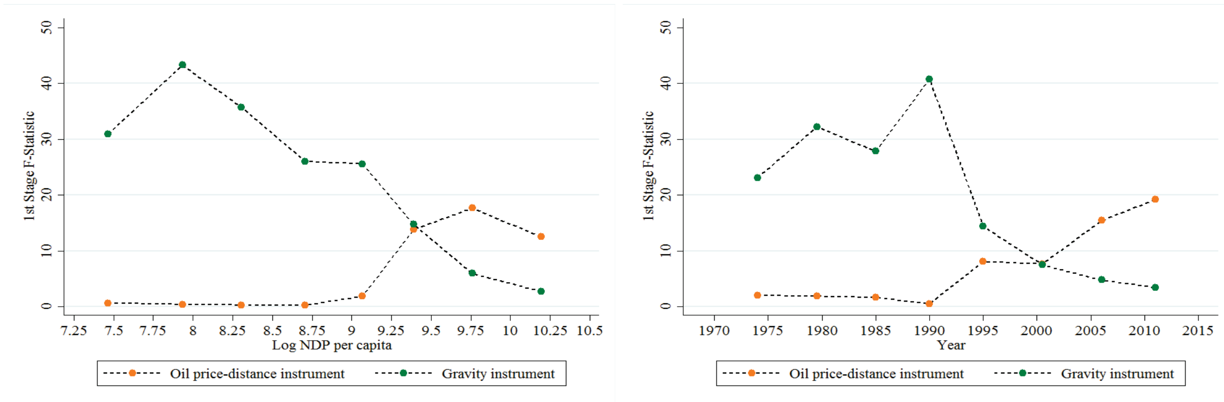
Figure O5: Trends by Initial Trade Openness in Developing Countries



Notes: These panels plot the time series of trade openness (top-left panel), effective tax rate on capital (bottom-left panel) and effective tax rate on labor (bottom-right panel). The sample is limited to low- and middle-income countries, according to the World Bank income classification in 2018. Within each panel, the orange line (green line) traces the evolution of the group which had relatively high (low) trade openness prior to 1995. Specifically, high (low) trade openness is defined as having average trade openness which lies above (below) the global average between 1965 and 1995. Trade openness is measured as the share of imports and exports in national domestic product; note that this share can exceed a value of 1. Each line plots the year fixed effects from an OLS regression in the relevant sub-sample of the outcome on country and year fixed effects. The inclusion of country fixed effects limits the influence of countries entering and leaving the sample. The fixed effects are normalized to equal the level of the outcome variable in the relevant sub-sample in 1965. The shaded area highlights the notable 1990-1995 period, which marks the beginning of the 'second wave' of globalization that featured a proliferation of bilateral and multilateral trade agreements (Egger, Nigai, & Strecker, 2019).

Figure O6: Strength of Individual Instruments Across Subsamples

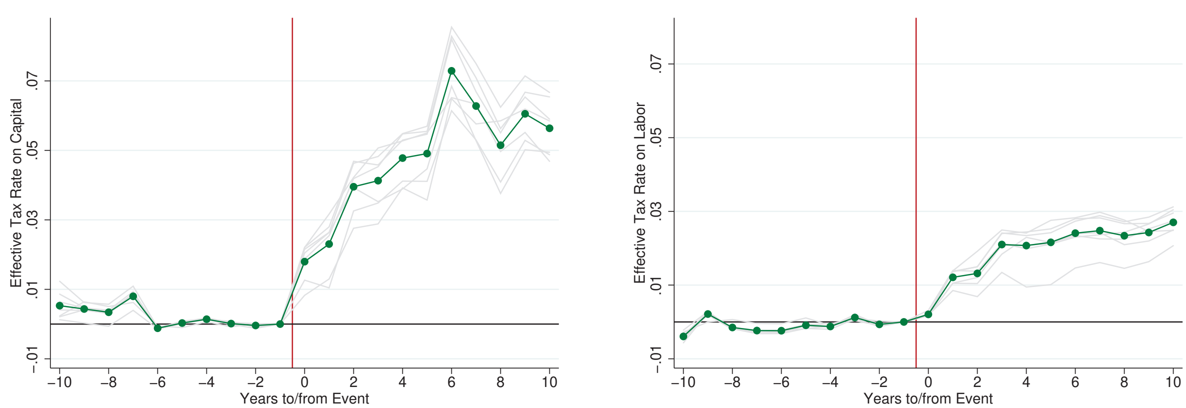
(a) Sub-samples of NDP per capita (b) Sub-samples of time-periods



Notes: These figures show the statistical strength of the instruments $Z^{oil-distance}$ and $Z^{gravity}$ in developing countries (low and middle-income countries based on World Bank classification in 2018, $N = 4916$). The outcome is the first-stage F-statistic from a regression of trade openness on each individual instrument, in subsamples of log NDP per capita (panel a) and years (panel b). The x-axis variable is partitioned into ten deciles, and the estimation is done in increments of one decile with a bandwidth of one additional decile of on either side. To maintain equal sample sizes, estimation centered on the first and the tenth decile are dropped. This figure is discussed in Section 5.2.

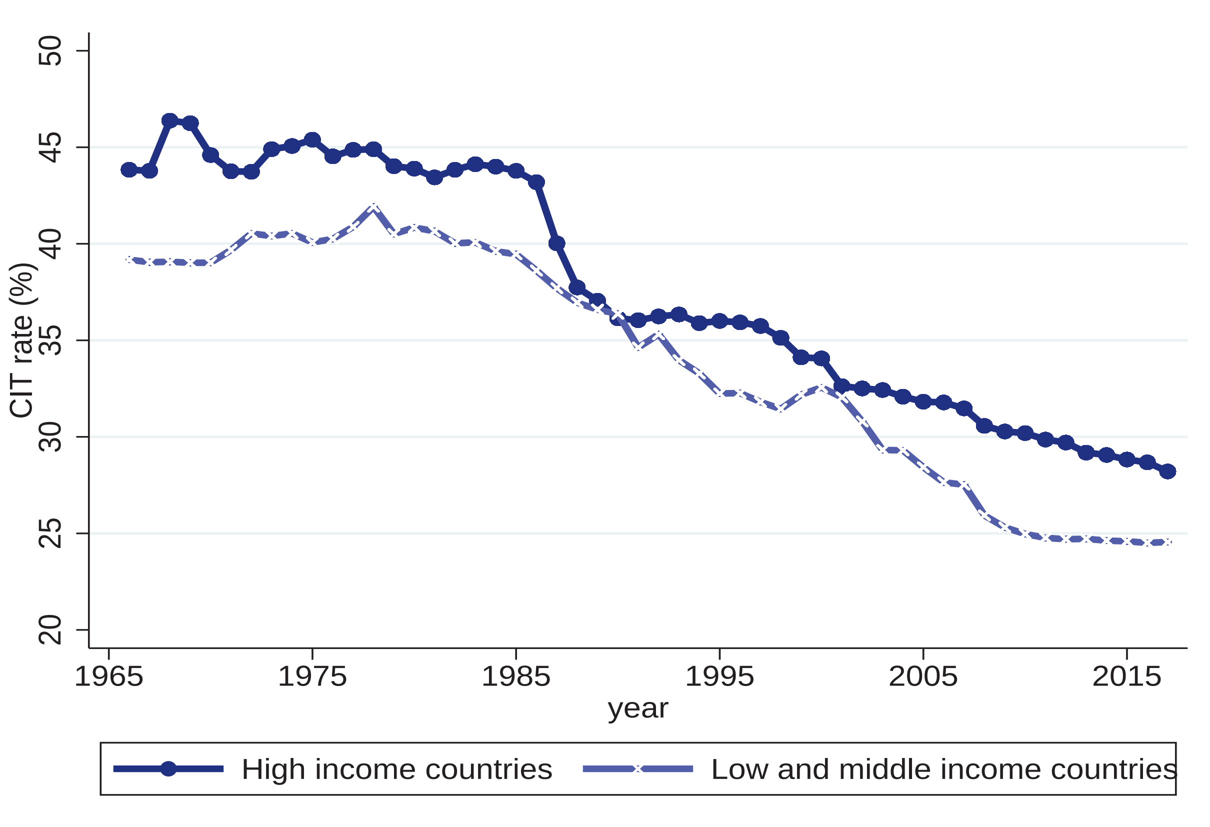
Figure O7: Robustness of Trade Liberalization to Changing Events-Sample

(a) Effective tax rate on capital (b) Effective tax rate on labor



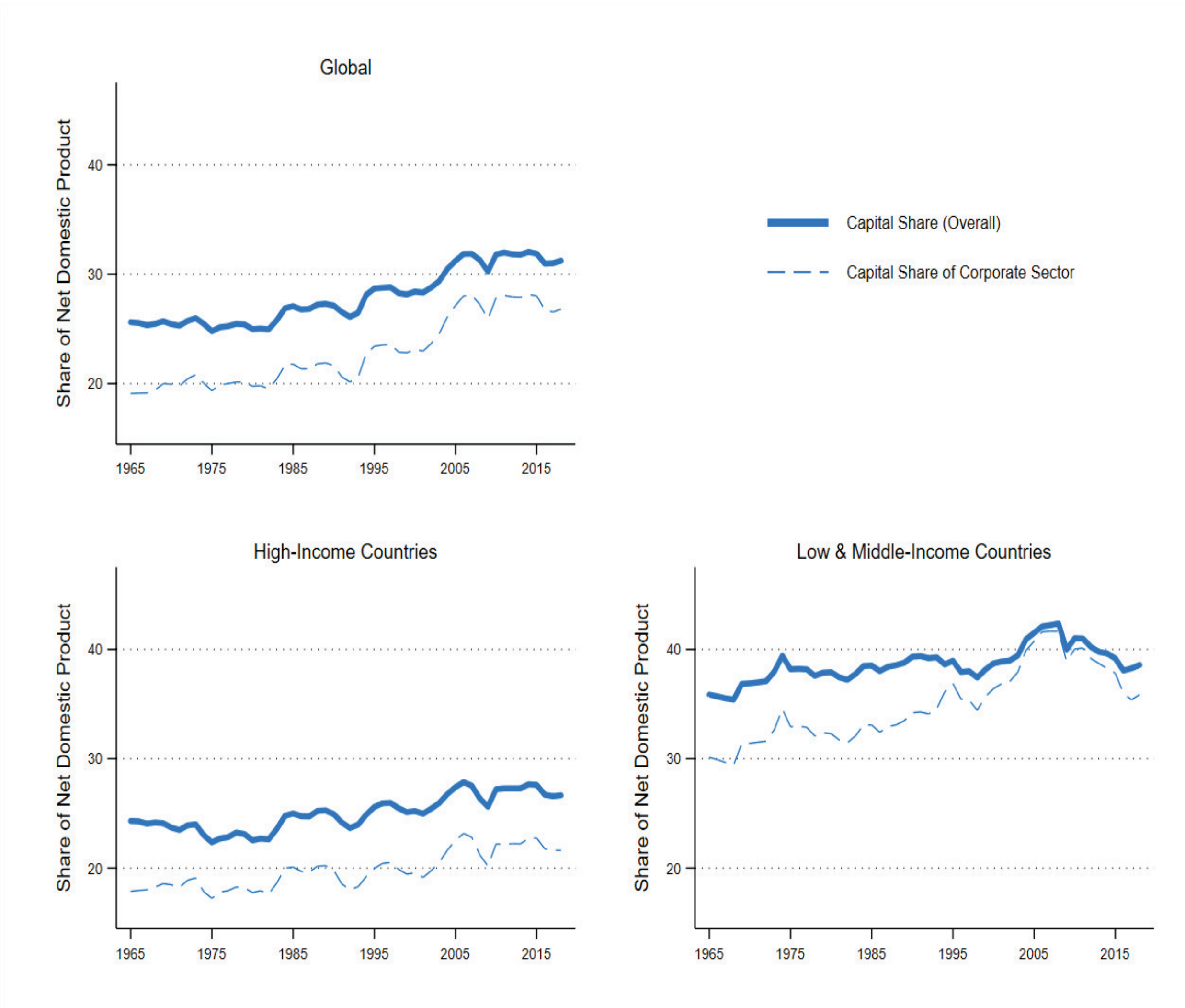
Notes: These figures show event study impacts of trade liberalization on the effective tax rate on capital (panel a) and the effective tax rate on labor (panel b). The solid green line displays the dynamic event-study coefficients β_e estimated in the full sample of 7 liberalization event-countries (Figure 6); the gray lines present the event-study coefficients estimated in samples that remove one event-country one at a time. More details in Section 5.1.1.

Figure O8: Trends in Corporate Income Tax Rates



Notes: This figure plots the time series of the statutory corporate income tax (CIT) rate, separately for high-income countries and for middle and low-income countries. Low, middle and high-income countries are based on the World Bank income classification in 2018. Each line plots the year-fixed effects from an OLS regression of the CIT rate on country and year-fixed effects in the relevant sub-sample. The inclusion of country-fixed effects helps alleviate the influence of countries entering and leaving the sample. The fixed effects are normalized to equal the level of the CIT rate in the relevant sub-sample in 1965. Country-year observations are weighted by their share in the year's total net domestic product in constant 2019 USD.

Figure O9: Capital Share of Net Domestic Product



Notes: This figure plots the time series of the capital share as a percentage of net domestic product (NDP). The solid line corresponds to the overall capital share, and the dotted line to the capital share within the corporate sector. The top left panel corresponds to the global average, weighting country-year observations by their share in that year's total NDP, in constant 2019 USD (N=154). The bottom-left panel shows the results for high-income countries, and the bottom right for low- and middle-income countries. Low, middle and high-income countries are based on the World Bank income classification in 2018. The dataset is composed of two (quasi) balanced panels. The first covers the years 1965-1993 and excludes communist regimes. The second covers 1994-2018 and integrates former communist countries, in particular China and Russia.

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Afghanistan	2003-2018	ICTD (2003-2018)		Exists HA series (1973-1978), but series is short and tax/GDP figures appear too volatile (could not find corroborating evidence); historical ICTD data exists (1982-1989), but no disaggregation by tax type	
Albania	1994-2018	ICTD (1994-2018)		Data begins in post-communist period; existing HA data nested in coverage in ICTD, so prefer ICTD source (and levels of tax revenues are comparable between sources)	Polackova (1996) tax/GDP estimate in 1993-1994 is slightly higher than ours, but likely includes non-tax revenues; World Bank (2020) tax/GDP matches our estimates from 1995 to 1998.
Algeria	1965-2017	HA (1965-2017)	*Annuaire Statistique de l'Algerie*	Taxe sur chiffre d'affaires classified as unallocable between PIT and CIT in HA, but it is a tax on firms; 'Contributions diverses' left as an excise tax in 1960s; interpolate 1967, 1970-1971, 1974.	
Argentina	1965-2018	HA (1965-1989); OECD (1990-2018)	*Sintesis Estadistica Mensual, Boletin Mensual de Estadistica*	Historically stable tax/GDP series, despite multiple political changes, until growth in tax take in 1990s when indirect tax expands; interpolate direct tax split (PIT vs CIT) between 1961-1969 and 1984-89, based on ratios on adjacent years; social security data from Alvaredo (2010) is comparable to OECD in overlapping years, so is preferred historical source.	Tax/GDP numbers comparable to historical time-series from Alvaredo (2010).
Armenia	1994-2018	HA (1994-2018)	Data provided by Statistical Committee of the Republic upon request	Independence in 1991, but official published revenue data begins in 1994; notable dip in social security in mid-2010s is genuine, results from several reforms (IMF, 2019; Asatryan, 2014).	Polackova (1996) tax/GDP estimate in 1993-1994 is slightly higher than ours, but likely includes non-tax revenues
Australia	1965-2018	OECD (1965-2018)			
Austria	1965-2018	OECD (1965-2018)			
Azerbaijan	1994-2018	ICTD (1994); HA (1995-2018)	Data retrieved from State Statistical Committee online data website	ICTD data is more accurate in 1994, from 1995 matches in trends and levels with HA data; independence in 1991, but unrest ensued until 1994 and limited government records (HA records unreliable, GDP numbers hard to corroborate); non-tax revenues are significant, especially since early 2000s; spike in CIT revenue in late 2000s reflects genuine economic shock (Aliyev and Gasimov, 2016)	
Bahamas	1973-2018	HA (1973-1989); OECD (1990-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	Historical HA data is based on IMF sources; social Initial difference between HA and OECD in overlapping years is due to social security contributions (missing in HA).	
Bahrain	1974-2018	HA (1974-1987); ICTD (1988-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	Historical HA data is based on IMF sources; change in CIT revenues in 1970s corresponds to nationalization and expropriation events (Kobrin, 1984) and there was no major change to oil production during this period (Ross and Mahdavi, 2015)	Comparable tax/GDP numbers in recent periods based on World Bank (2020), though the data in latter source stops in 2004.
Bangladesh	1976-2018	HA (1976-2000); ICTD (2001-2018)	*Budget Book, Statistical Digest of Bangladesh*	Independence in 1971, but reliable government data begins in 1976. Interpolate 1980-1981; very low direct taxes collected on firms prior to 1986 reform, and significant CIT drop in 2003.	Comparable tax/GDP numbers with ICTD in overlapping periods.
Barbados	1972-2018	HA (1972-1990); OECD (1991-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	HA data is based on IMF historical reports; use social security as reported in initial sources, corroborated with data from Fisunoglu et al. (2011)	
Belarus	1992-2018	ICTD (1993-2018)		ICTD data exists in 1991 but it is not disaggregated; decrease in CIT and increase in indirect taxes in early 2000s, may be due both to Russian financial crisis and to ICTD switching its source from IMF Article IV to IMF GFS [flagged]	Consistent tax/GDP when comparing to World Bank (2013), after adjusting ICTD for existence of social security contributions
Belgium	1965-2018	OECD (1965-2018)			
Belize	1982-2018	HA (1982-1989); OECD (1990-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	HA data is based on historical IMF data; interpolate 1986-1987; social security contributions missing in HA, we take it from Fisunoglu et al. (2011); social security started in 1979 (SSA, 2015)	
Benin	1965-2018	HA (1965-2018)	*Comptes de la Nation, Statistiques Finances Publiques*	Social security first implemented in 1970 (SSA, 2017); interpolate between 1988 and 1990.	Historical sources are hard to find. HA series comparable to historical IMF series in early periods (1960-1970), and dip in late 1980s exists across sources.
Bolivia	1965-2018	HA (1965-1989); OECD (1990-2018)	*Anuario Estadistico, Bolivia en Cifras*	Use historical data from Fisunoglu et al. (2011) for social security, which started prior to our time-coverage (SSA, 2017); unclear what 'complementaria' tax (1960-1970) refers to, we assign it equally to PIT and CIT; large decline in mid-1980s appears genuine (Kehoe et al., 2019)	Historical tax/GDP numbers comparable to Kehoe et al. (2019) and Sachs (1990), though larger than numbers reported in Thirsk (1997) in 1970s [flagged].
Bosnia and Herzegovina	1999-2018	ICTD (1999-2018)		War ends in 1995 but reliable data only starts in 1999; important role of local taxation (Fox and Wallich, 1997; Kandeva, 2001), compare data from ICTD with IMF GFS and Zorn et al. (1999) which suggest local tax sources are adequately covered.	Comparable tax/GDP numbers with World Bank (2020) after 2005, but higher tax/GDP reported in Ding and Sherif (1999) for historical period [flagged].
Botswana	1967-2018	HA (1967-1989); ICTD (1990-2003); OECD (2004-2018)	*Annual Statements of Accounts, Statistical Abstract, Statistical Bulletin, Financial Statistics*	OECD data is missing trade taxes, which we bring in from ICTD in overlapping years; GDP estimates differ in the pre-1990 period between IMF, World Bank and UN-SNA sources [flagged], we use World Bank source; CIT value in ICTD in 1990 appears too large, interpolate based on surrounding years [flagged]; 'mineral tax' in HA data appears to partly include CIT, predict CIT-share based on precise split between CIT and resource tax in other sources [flagged]; social security program starts in 1996 according to SSA (2017), however we observe contributions prior to that date (Fisunoglu et al., 2011) which may correspond to a non-contributory pension benefit (Arza and Johnson, 2006) [flagged]; large economic shocks in 1980s which affected public finances (O'Connell, 1988).	Comparable historical tax/GDP based on Takirambudde (1995), O'Connell (1988), Bonu (1995).

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Brazil	1965-2018	HA (1965-1989); OECD (1990-2018)	*Anuario Estatistico do Brasil*	Challenging to find reliable GDP data in historical periods, use reported national price index from Ayres (2019) prior to 1990 which in turns is based on Brazilian Institute of Geography and Statistics; OECD data appears to be high quality, including with respect to sub-national tax collection; corroborate sub-national taxes in HA data using detailed information from Afonso and Araujo (2004), which discusses local public finance since 1960s, and Varsano (1999); use Fisunoglu et al. (2011) prior to 1980 for social security contributions. Interpolate income tax between 1968 and 1971 and 1973 and 1975.	Historical tax/GDP numbers comparable to Afonso and Araujo (2004), Chelliah (1971) - both of which cover national and sub-national revenues.
Bulgaria	1993-2018	HA (1993-2018)	*Statistical Yearbook, Monthly Statistical Reviews*	Excess tax revenue' category in 1995-96 is contribution from previous year taxes collected, we check that it has not been double-counted [flagged]; some difference in social security contributions between UN-SNA data and Fisunoglu et al. (2011), prefer latter source as it compares to ICTD in overlapping years	Comparable tax/GDP numbers in early periods based on Bogetic and Hassan (1997), match World Bank (2020) in later years.
Burkina Faso	1965-2018	HA (1965-1999); OECD (2000-2018)	*Comptes Econonomiques de la Haute-Volta, Comptes Nationaux de la Haute-Volta, Les Comptes Economiques de la Nation*	Several periods of political instability (1974, 1980-1983, 1987) where interpolate data; large tax from property rights registration tax in late 1970s, appears genuine.	Comparable tax/GDP numbers with ICTD in overlapping periods.
Burundi	1965-2018	HA (1965-2018)	Annuaire Statistique, Bulletin Statistique	Data interpolated 1970-1973 but concerns remain about data quality (violence) [flagged]; IMF (1973) data suggests little change in composition of taxes, though change in overall tax take; historical IMF source lists a tax on property, which cannot be found in HA (which instead records a transaction tax)	
Cambodia	1994-2018	ICTD (1994-2018)		HA data exists from 1987 to 1993, but the data has quality concerns (given political transition), so prefer not to use that data; social security contributions begin in 1997 (SSA, 2018), we draw on data from Fisunoglu et al. (2011).	
Cameroon	1965-2018	HA (1965-1992); OECD (1993-2018).	*Note Trimestrielle sur la Situation Economique, Note Annuelle de Statistique*	Interpolate 1969-1970, 1989; classify the 'taxe unique' as an indirect tax, rather than direct firm tax, based on information from Gauthier et al. (2002); drop in overall revenue in 1980s confirmed to be mainly related to dwindling trade taxes (Gauthier and Gersovitz, 1997); decline in wealth taxes between 1968 and 1993 is not accounted for via additional sources [flagged]; social security contributions start in 1968 (SSA, 2017), we draw data from Fisunoglu et al. (2011); significant general volatility of revenues likely driven by reliance on volatile commodities (de Herdt, 2002).	Tax/GDP numbers comparable to Gauthier, Soloaga and Tybout (2002).
Canada	1965-2018	OECD (1965-2018)		Vaillancourt and Kerkhoff (2019) for additional information on the capital share of PIT	
Central African Republic	1965-2018	HA (1965-2007); ICTD (2008-2018)	*Bulletin de Statistique, Bulletin Mensuel de Statistique, Annuaire Statistique*	Political unrest in early historical periods create uncertainty around data [flagged]; the 'tax additionnelle' in the 1960s was a direct tax on firms rather than individuals (Mbounou-Ngopo, 2019); dips in tax collection in mid-1990s coincide with political transitions. HA data features change in terminology 1994-1997, which could erroneously be interpreted as a substitution from direct to indirect taxes; social security started in 1963 (SSA, 2019), we draw on Fisunoglu et al. (2011) for entire HA period.	Observed decrease in tax/GDP in recent periods also confirmed in IMF reports (IMF, 2016). Difficult to find historical sources to corroborate.
Chad	1965-2018	HA (1965-1982); ICTD (1983-2009); OECD (2010-2018)	*Bulletin Mensuel de Statistique, Budget General de l'Etat*	Military rule from 1975-1978 and civil war from 1979-1982, interpret data with caution during these periods [flagged]; social security program began in 1977 (SSA, 2017), draw on data from Fisunoglu et al. (2011) prior to OECD coverage; volatility in recent years is notable but appears to be genuine (found in both OECD and ICTD data).	Tax/GDP estimate in early historical period (1965) approximately 1.5 percentage points lower than reported in Lotz and Morss (1967).
Chile	1965-2018	HA (1965-1979); ICTD (1980-2018)	*Informe Economico Anual, Statistical Profile of Chile,*	The more recent data (1990-2018) is from Inter-American Development Bank; interpolate 1978-79; data quality is challenged during 1970-1973 period, the transition years for Allende [flagged]; use information from Mamalakis (1978) and Corbo (1989) to confirm HA data split between CIT and PIT, the extent of sub-national taxes, and the existence of social security contributions.	Social security data from Fisunoglu et al. (2011) in agreement with historical data from Cerda (2005). Historical tax/GDP ratio comparable to values reported in World Bank (1980).
China	1994-2018	HA (1994-2007); OECD (2008-2018)	Statistical Yearbook, online data from National Bureau of Statistics	See Appendix B for more details on sources and tax system; data exists prior to 1994, but we start the series after the transition away from central planning (conceptual difficulties with defining certain revenue sources as taxes in the pre-transition period); HA and OECD match well in overlapping years.	Tax/GDP comparable to values reported in Lou and Wang (2008) and ICTD.
Colombia	1965-2018	HA (1965-1989); OECD (1990-2018)	*Estadisticas Fiscales, Cifras Fiscales, Informe Financiero*	Good match in levels when sources overlap; special 'pro equity income tax', initially classified as unallocable between PIT and CIT in OECD data, is in fact a tax on corporate income (World Bank, 2014); concern if we are capturing all sub-national taxes in the HA period (based on Arroyo-Abad and Lindert, 2016) [flagged]; social security data missing in HA period [flagged].	McClure cites a tax/GDP in 1980 which is 3 percentage points higher than our HA estimate, though our estimates agree with Mitchell (2003) in the period 1965-1989.
Congo	1972-2018	HA (1972-1981); ICTD (1982-1998); OECD (1999-2018)	*Annuaire Statistique, Bulletin Mensuel de Statistique, Economic Survey*	Interpolate between 1977 and 1979; year of overlap between ICTD and OECD coincides with period of genuine drop in revenues, due to violence.	Historical tax/GDP values are broadly in line with Tait et al. (1979) for the 1970s.
Costa Rica	1965-2018	HA (1965-1987); OECD (1988-2018)	*Anuario Estadistico, Memoria Annual*	Use IMF historical data between 1974 and 1987, matches will in overlapping years with both HA and OECD; low CIT revenue collected in the 1980s is confirmed in Shome (1992); social security contributions are from Fisunoglu et al. (2011).	Comparable historical tax/GDP compared to Tait et al. (1979).

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Cote d'Ivoire	1965-2018	HA (1965-1989); OECD (1990-2018)	*Bulletin Mensuel de Statistique, Les Comptes de la Nation, Budget General de Fonctionnement,*	Interpolate 1977-1979, 1987-1989; in HA years where individual versus firm split exists for 'tax on benefices', assume same ratio in all other years where same tax does not have breakdown; OECD data appears to under-estimate PIT in recent years, use information from IMF and ICTD to adjust level [flagged].	
Croatia	1996-2018	ICTD (1996-2018)		HA data exists but has less complete coverage than ICTD by type of tax; ICTD also appears to have captured well sub-national taxes (IMF, 2020).	Comparable levels of tax/GDP in World Bank (2020).
Cuba	1990-2018	OECD (1990-2018)		HA data exists in some of the historical OECD periods, but prefer to draw all data from a single source	
Cyprus	1972-2018	HA (1972-2018)	*Annual Budget*	Historical IMF data between 1972 and 1989, then HA for remaining periods; supplement HA with Fisunoglu et al. (2011) for social security contributions; Lent (1977) confirms existence of corporate income tax in 1970s.	Comparable tax/GDP levels to Lent (1977).
Czech Republic	1993-2018	OECD (1993-2018)			
Democratic Republic of the Congo	1968-2018	HA (1968-1990); ICTD (1991-2018)	*Conjoncture Economique*	Some difference between sources in overlapping years, coincides with period of high inflation and significant seigniorage tax (De Herdt, 2002; Nachega, 2005) [flagged]; between 1977 and 1990, tax type called 'divers' which is initially unallocable between PIT and CIT [flagged]; likely that we capture local tax revenues in historical periods. Interpolate 1973, 1992-1995.	Trends in taxation by source in historical HA periods is consistent with Emziet (1997).
Denmark	1965-2018	OECD (1965-2018)		Specificity in how social security contributions are levied (through the PIT, unlike in many other countries)	
Dominican Republic	1968-2018	HA (1968-1989); OECD (1990-2018)	*Ejecucion Presupuesto*	Omission of indirect tax categories in HA in late 1970s and late 1980s - interpolate based on surrounding years [flagged]; given unrest in early 1960s, we begin our series in 1968.	Historical IMF data agrees with HA estimates in 1970s.
Ecuador	1973-2018	HA (1973-1989); OECD (1990-2018)	*Cuentas Nacionales del Ecuador*	Ministry of Finance (2016) data includes breakdown between resource and non-resource revenues in historical periods, which suggest that our HA data sometimes includes resource revenues [flagged] - use historical IMF data in limited sets of years to correct the direct firm income tax numbers in HA.	CEPAL (1991) corroborates tax revenue levels in 1979-1987, including the spike in 1985. Garcia and Uquilles (1992) confirm levels in 1989-1992 period.
Egypt	1965-2018	HA (1965-1989); ICTD (1990-2001); OECD (2002-2018)	*Annuaire Statistique*	HA and ICTD data match very well in overlapping year; for periods prior to OECD, use Fisunoglu et al. (2011) to measure social security contributions; interpolate 1965-67, 1970-1971, and 1973-1974; sharp drop in revenue in 1980s is genuine; difficult to ensure CIT does not in part capture resource revenues, as they grew in importance in 1970s [flagged]; in years 2002-2008, 'tax on movable capital revenues from Central Bank' appears to be unallocable between CIT and PIT, we assign shares based on information in Waterbury (2014).	Smith (1970) and Nyrop (1976) corroborate low level of PIT in 1960s and early 1970s. Mitchell (2003) differs on average by 10% from our tax/GDP estimates.
El Salvador	1965-2018	HA (1965-1989); OECD (1990-2018)	*Anuario Estadistico, Indicadores Economicos y Sociales*	Significant currency reform in 2001, interpret tax/GDP number with caution in that year [flagged], and we adjust the currency in prior years to be comparable; social security contributions began in 1959 (SSA, 2017), we use data from Fisunoglu et al. (2011) in years prior to OECD coverage, which was substantial (Grosh, 1990).	Estimates of tax/GDP in 1980s and 1990s very close to numbers reported in Cardemil et al. (2000).
Equatorial Guinea	1981-2018	ICTD (1981-2004); OECD (2005-2018)		Very limited direct taxes collected prior to early 2000s, corroborated in Same (2008); historically strong reliance on revenues from commodity exports (Human Rights Watch, 2017); social security first implemented in 1947 (SSA, 2017), we use data from Fisunoglu et al. (2011).	
Estonia	1993-2018	ICTD (1993-1994); OECD (1995-2018)		Russian presence only phased out by 1993; good match between sources in overlapping years	Our tax/GDP numbers in early transition period matches well with Polackova (1996).
Ethiopia	1965-2018	HA (1965-1992); OECD (1993-2018).	*Statistical Abstract*	Income tax only separates personal from business income starting in 1975; per Schwab (1970), income tax schedules prior to then included both individual and firm income - so we keep the initially unallocable category; our HA record of land use fees suggests we are capturing sub-national taxes; Chloe (1984) notes the reliance on commodity exports, which induces volatility in tax revenues; very limited quantity of social security contributions in comparison with Mengistu et al. (2017) [flagged]. Interpolate 1989, 2005.	Compares well with Mascagni (2016) for historical tax/GDP in 1960s-1980s, series are within one percentage point of each other.
Fiji	1972-2018	HA (1972-1989); ICTD (1990-2007); OECD (2008-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	HA data derives from historical IMF data; use Fisunoglu et al. (2011) for social security contributions, which began in 1966 (SSA, 2016); disappearance of property taxes between 1992 and 2010, though consistently missing in both ICTD and OECD sources [flagged]	
Finland	1965-2018	OECD (1965-2018)			
France	1965-2018	OECD (1965-2018)			
Gabon	1965-2018	HA (1965-1985); ICTD (1986-2018)	*Annuaire Statistique du Gabon, Tableau de Bord de l'Economie*	Confident that HA sources excludes resource revenues, comparison with historical IMF data; interpolate 1977-1980; several historical taxes unallocable between CIT and PIT, including 'impot general sur le revenue', for which historical sources are not informative (Abdel-Rahman, 1965); social security contributions began in 1963 (SSA, 2019), we draw on data from Fisunoglu et al. (2011).	Limited historical sources, though drop in revenue in 1980s is corroborated in Gaulme (1991) and Yates (1996).
Gambia	1972-2018	HA (1972-1987); ICTD (1988-2018)	*Estimates of Recurrent Revenue and Expenditure, Gambia Statistical Yearbook*	Start data in 1972, as currency change in 1971 introduces measurement challenges; at same time, continued macro-economic volatility means data prior to 1990 should be interpreted with caution [flagged]	Tax/GDP estimates in HA are comparable to Ansari (1982), about 18% lower on average. Jallow (2016) provides historical account of tax system, but no data on tax/GDP.

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Georgia	1994-2018	ICTD (1994-2018)		Social security contributions, in their current form, began in 1990 (SSA, 2010), we draw on data from Fisunoglu et al. (2011); interpolate social security between 1995 and 1997.	Barbone and Polackova (1996) have comparable tax/GDP estimates for 1995, once our series are adjusted for social security contributions.
Germany	1965-2018	OECD (1965-2018)			
Ghana	1967-2018	HA (1967-1999); OECD (2000-2018)	*Quarterly Digest of Statistics*	Very good match between sources in overlapping years, once account for social security contributions; remove revenue from government properties in HA historical periods (initially counted within 'tax on interest and profits); volatile patterns in 1980s confirmed in Darko-Osei and Telli (2017), coinciding with economic turbulence and IMF recovery and structural adjustment agreements.	Reasonable comparison for tax/GDP numbers in historical periods, as reported in Chelliah (1971), Lotz and Morss (1967), Killick (1978) and Darko-Osei and Telli (2017).
Greece	1965-2018	OECD (1965-2018)			
Guatemala	1965-2018	HA (1965-1989); OECD (1990-2018)	*Estadisticas de Finanzas Publicas de Guatemala*	OECD classifies solidarity tax as unallocable, since 1995, but it is a tax on corporations (PriceWaterhouseCooper, 2020); historical data on social security contributions from Fisunoglu et al. (2011) seem unreliable in this case, so we report such taxes from 1978 onward [flagged].	Tax/GDP estimates in early historical years (1965-1967) lie within one percentage point of numbers from Lotz and Morss (1967).
Guinea	1980-2018	ICTD (1980-2018)		Revenue movements in 1980s coincide with political transition and new fiscal regime under Conte (Yansane, 1990); Topouzis (1989) notes the rising importance of resource revenue, which may be captured inside our CIT category between 1985 and 1992 [flagged].	
Guyana	1972-2018	HA (1972-1986); ICTD (1987-1989); OECD (1990-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	HA is drawn from IMF sources; drop in revenue at end of 1980s is genuine, likely reflects economic turbulent times (consistent numbers across sources); initial difference across sources in overlapping years is due to ICTD data not accounting for social security contributions.	
Haiti	1975-2018	HA (1975-1989); ICTD (1990-2018)	*Tableau des Operations Financieres de l'Etat*	HA data draws on historical IMF reports; interpolate 1988-1989, though should be interpreted with caution given violence at time [flagged]; spike in 1987 is driven by collapse in underlying value of GDP; social security begins in 1965 (SSA, 2017), we draw on Fisunoglu et al. (2011); property and transaction taxes exist in HA, small in magnitude, but are missing from ICTD data [flagged].	Dioda (2012) estimates similar tax/GDP in 1990, Tanzi (2000) estimates slightly higher tax/GDP between 1993 and 1999.
Honduras	1973-2018	HA (1973-1989); OECD (1990-2018)	*Anuario Estadistico*	1974 is a complicated year due to missing types of taxes, so we interpolate it; jump in revenues around year of overlap between OECD and HA, but Herrera (1994) corroborates significant changes in tax performance at that time; use social security data from Fisunoglu in all years prior to OECD coverage.	Historical IMF data are comparable to HA numbers for tax/GDP in the 1980s.
Hungary	1994-2018	OECD (1994-2018)			
Iceland	1965-2018	OECD (1965-2018)		World Bank (2020) corroborates spike in non-recurrent property taxes observed in 2016; infer, based on Herd and Thorgeirsson (2001) and Karlsson (2014), assign portion of initially unallocable OECD revenue to individuals.	
India	1965-2018	HA (1965-2019)	*Monthly Abstract of Statistics, Indian Public Finance Statistics*	In HA, 'income tax other than CIT' is not exclusively a tax on individuals, we use additional information to assign this category to firms versus individuals (including Rao, 2005); HA data carefully records a comprehensive set of wealth and property taxes, including for land; reasonably confident HA captures sub-national taxes; social security contributions, very small in magnitude, appear to be reported inside individual income tax category [flagged].	Very comparable tax/GDP estimate in earliest periods with Rao (2005); also consistent with recent estimates in World Bank (2019).
Indonesia	1965-2018	HA (1965-1996); OECD (1997-2018)	*Statistik Indonesia*	Strong match for data in overlapping years between sources; 1983 reform collapsed multiple taxes (CIT, PIT, other direct income) into a single schedule, so we use shares of capital versus labor direct income taxes in 1983 and assign such shares until 1997 [flagged]; reasonable ability to exclude resource revenues in HA data. Interpolate 1968-1971, 1994.	Multiple sources estimate very low historical tax/GDP ratios (1960s and 1990s), between 2 and 8 percentage points (Prasetyo, 2018; Gillis, 1985; Amir et al., 2013), generally consistent with our estimates
Iran	1969-2018	HA (1969-2018)	*Annual Budget, Iran Statistical Yearbook*	Reasonable ability to exclude resource revenues; social security data in HA and in ICTD are unreliable, we instead draw data from UN-SNA, starting in 1996 [flagged]; data strictly based on central government, but no documentation suggests sub-national taxes are quantitatively important.	In overlapping period (1979-1989), our tax/GDP estimates and trend match very closely with Mazarei (1996). Generally limited studies in English on historical tax system in country.
Ireland	1965-2018	OECD (1965-2018)			
Israel	1972-2018	HA (1972-1992); OECD (1993-2018)	*Accountant General's Report, IMF Government Finance Statistics*	HA draws in part from historical IMF data; interpolate 1992-1994; some IMF data reported in 1970s seems approximate [flagged]; historical IMF misses property tax in some years, supplement with ICTD data.	Historical trends in 1980s and 1990s are corroborated in Brender (2007), though the level of tax/GDP is approximately 15% higher than our estimates.
Italy	1965-2018	OECD (1965-2018)			
Jamaica	1965-2018	HA (1965-1989); OECD (1990-2018)	*Statistical Yearbook of Jamaica, Abstract of Statistics*	Property taxes dip in 1997, but this appears genuine (in the OECD data, based on local public finance records); unallocable part of direct income taxes is significant in the 1980s, comprising a mix of taxes on dividends, interest and an 'education tax' since 1983 (Government of Jamaica, 1988), and we use additional information from Inter-American Center of Tax Administrators to assign it to firms versus workers;	Historical estimates of tax/GDP in 1960s and 1970s are 10-18% larger than ours (Chelliah, 1971; Shome, 1992).
Japan	1965-2018	OECD (1965-2018)			
Jordan	1973-2018	HA (1973-1989); ICTD (1990-2018)	*Annual Report, Yearly Statistical Series, Monthly Statistical Bulletin*	HA uses historical IMF data; at year of merge between sources, Abu-Hammour (1997) confirms a large increase in tax/GDP; non-tax revenues are significant in the country, but our sources can reasonably exclude them.	HA data matches very closely the numbers in Abu-Hammour (1997).

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Kazakhstan	1993-2018	ICTD (1993-1998); OECD (1999-2018)		Independence in 1991, but ICTD coverage begins in 1993; social security contributions began in 1991, we use data from Fisunoglu et al. (2011); volatility in indirect tax revenues is significant but plausible (World Bank, 2017).	
Kenya	1965-2018	HA (1965-2000); OECD (2001-2018)	*Statistical Abstract*	To assign initially unallocable direct tax between firms and individuals, we use information from Jetha (1966) and Wanjala (2006) on the income tax schedules; gradual increase in tax/GDP over the long run observed in Macha et al. (2018) and Omondi et al. (2014); possible that we are not capturing sub-national taxes in HA data [flagged]	Our tax/GDP estimates are systematically smaller by 2 percentage points in the 1960s and 1970s compared to other estimates (Wanyagathi, 2015; Kanji and Wanjala, 2005), but this may also be due to differences in the underlying estimate of GDP (our GDP estimates based on WID are larger than World Bank estimates).
Korea	1965-2018	HA (1965-1971); OECD (1972-2018)	*Korea Statistical Yearbook*	Good match in levels for years of overlap; interpolate from 1968 to 1971; overall low levels of revenue in the 1960s are genuine and reflect government policy; only after major tax reform in 1967 did tax collection start to significantly grow (Yoo, 2000); observed changes in capital taxes collected in 1970s are genuine (Kwack and Lee, 1992)	Tax/GDP estimate reported in mid-1960s is approximately 1 percentage point higher than our data.
Kosovo	2008-2018	HA (2008-2015); OECD (2016-2018)	Data retrieved from Department of Finance and General Services	Government data prior to 2008 is scarce; according to Koshutova (2004) and Kritzer (2005), pension system is funded through general taxation; level of CIT as a share of GDP is confirmed in Hernandez et al. (2019)	
Kuwait	1972-2018	HA (1972-1989); ICTD (1990-2018)	*Government Finance Statistics*	Interpolate 1975-1976; do not observe social security contributions in HA, but it is place in historical years (SSA, 2016), so we use data from Fisunoglu et al. (2011); large resource nationalization in 1975, firms' income tax data from 1972 to 1975 should be interpreted with caution [flagged]	Good historical match in tax/GDP and sources for 1972-1976 when compared to Nyrop (1977).
Kyrgyzstan	1994-2018	ICTD (1994-2018)		Data coverage of property and wealth taxes only begins in 1995, but we could not find a historical source to confirm if this reflects a policy implementation [flagged]	Estimates of tax/GDP in the 1990s are very close to data reported in Bokros and Dethier (1998) and Barbone and Polackova (1996).
Laos	1982-2018	ICTD (1982-2009); OECD (2009-2018)		While social security contributions have existed since 2001 (SSA, 2016), there is no data covering these contributions [flagged]; significant non-tax sources of revenue in the 1980s (Saignasith, 1997)	Historical estimates in Saignasith (1997) are larger than our data, but those figures also report for total revenues (rather than total taxes).
Latvia	1994-2018	OECD (1994-2018)		ICTD data exists prior to OECD coverage, but it is not disaggregated and difficult to reconcile with OECD numbers in overlapping years.	High levels of tax/GDP in 1994-1995 are corroborated in Polackova (1996).
Lebanon	1965-2018	HA (1965-2018); ICTD (1988-2001)	*Recueil de Statististiques Libanaises, Statistical Yearbook*	Series uses HA data, with ICTD data between 1988 and 2001; turbulent tax collection during civil war period (1975-1990), where information in Dimashkieh (1993) and Houry (1997) confirm the levels of taxes by type; social security contributions began in 1963, we use data from Fisunoglu et al. (2011); use information from Eken et al. (1995) to confirm level of CIT collected prior to 1993.	Historical estimates in Saleh (2004) are comparable to our series during turbulent period (1975-1990).
Lesotho	1965-2018	HA (1966-1981); ICTD (1982-2018)	*Statistical Bulletin*	Data missing in 1978-1981, but due to unrest we leave data empty (rather than interpolate); spike in revenue in 1977 is corroborated across data-sets; licensing fees constitutes large source of revenue in earliest periods (Cobbe, 1981).	Tax/GDP estimates from our data are 2 percentage points lower in the 1970s than the numbers reported in Cobbe (1981).
Liberia	1970-2018	HA (1970-1988); ICTD (2000-2018)	*Economic Survey, Quarterly Statistical Bulletin of Liberia, Statistical Bulletin of Liberia*	Important gap in coverage between 1988 and 2000 - a turbulent period during which revenues were collected but diverted from official use and GDP decreased by 90% (Atkinson, 1997); drop in revenue in 1973-1974 is genuine.	Levels of tax/GDP, both before and after the data-gap, are comparable to estimates in Davies and Dessy (2016).
Lithuania	1991-2018	HA (1991-1994); OECD (1995-2018)	*Lithuania Statistics Yearbook*	Social security begins in 1991, we use data from Fisunoglu et al. (2011).	High levels of tax/GDP in 1993-1994 is corroborated in Polackova (1996).
Luxembourg	1965-2018	OECD (1965-2018)			
Macedonia	1993-2018	ICTD (1993-2018)		Interpolate income taxes from 2003 to 2005; SSA (2005) confirms that social security contributions are significant in the country.	Excluding social security, tax/GDP levels are comparable to estimates since 2005 in World Bank (2020).
Madagascar	1965-2018	HA (1965-1989); ICTD (1990-2018)	*Inventaire Socio-Economique, Malagasy Budget, Budget General de l'Etat*	Interpolate 1969-1971, noting the political instability in 1972, 1974-1976, and 1981-1983; use social security data from Fisunoglu et al. (2011), starting in 1965 (SSA, 2018) though data from 1965 to 1969 may be estimated in original source [flagged].	
Malawi	1965-2018	HA (1965-2004); OECD (2005-2018)	*Malawi Statistical Yearbook, Compendium of Statistics, Economic Report*	Social security contributions missing in OECD, so we use data from Fisunoglu et al. (2011) for entire period, which were significant in more historical periods; spike in PIT revenue in 2001 is genuine, likely reflects tax enforcement reforms; Shalizi and Thirsk (1990) emphasize that direct income taxes were a significant share of total taxes in the 1960s.	Historical estimates of tax/GDP are slightly lower than in Chipeta (1998) and Shalizi and Thirsk (1990), though this may also be due to differences in GDP values.
Malaysia	1965-2018	HA (1965-1989); OECD (1990-2018)	*Economic Report*	Interpolate 1979-1980 and 1988, due to missing HA data; use social security contributions data from Fisunoglu et al. (2011), but unreliable in period from 1972 to 1980 [flagged]; total levels match well in overlapping years between OECD and HA, highlights that the stamp duty, which OECD classifies in as 'other tax', is classified as unallocable income tax in HA; drop in indirect taxes in late 1980s is genuine.	Limited existence of studies for historical comparison.
Mali	1965-2018	HA (1965-1979); ICTD (1980-1999); OECD (2000-2018)	*Comptes Economiques du Mali, Annuaire Statistique*	Levels match well in overlapping years between HA and ICTD, lends confidence to HA sources even though we cannot find multiple historical sources to corroborate (Founou-Tchuigoua, 1989); social security contributions begin in 1961 (SSA, 2019), we use data from Fisunoglu et al. (2011); OECD lists no corporate tax in 2000, though ICTD does [flagged].	Limited historical sources; our estimates are comparable to Founou-Tchuigoua (1989) for period 1981-1989.

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Mauritania	1986-2018	HA (1986-2006); OECD (2007-2018)	*Annuaire Statistique, Bulletin de la Direction de la Statistique et des Etudes Economiques*	Levels match well in overlapping years between HA and OECD; 'autres droits' in HA is listed as income tax in OECD, so we follow that assignment in HA; interpolate missing indirect taxes in 2000 and 2007-2008; limited historical sources, 1986 appears to be first year the government prepared a comprehensive budget statement (Handloff, 1990).	Limited historical sources; Oualalou and Jaidi (1986) discuss low overall levels of tax collection historically, and World Bank (2020) provides no data on the country.
Mauritius	1973-2018	HA (1973-1989); OECD (1990-2018)	*Digest of Statistics*	HA draws on historical IMF data; historical government publications match on levels to IMF data, but is less complete in years of coverage; use data from Fisunoglu et al. (2011) for social security contributions prior to OECD coverage.	Limited existence of studies for historical comparison.
Mexico	1965-2018	HA (1965-1979); OECD (1980-2018)	*Anuario Estadistico de los Estados Unidos Mexicanos*	Interpolate social security contributions 1973-1976; large increase in indirect taxes in early 1980s is genuine (driven by policy reforms), as is the drop in income taxes in 1980; initially unallocable income tax 'impuesto sobre la renta' pre-1970 (Aguilar, 2003), assign shares to firms and individuals based on data post-1970; unclear if we capture sub-national taxes pre-1970 [flagged].	Overall levels of taxes match the historical estimates in Martinez-Vasquez (2001) for period 1980-1999.
Moldova	1992-2018	ICTD (1992-2018)		ICTD misses revenue from property-transaction taxes, which we retrieve with HA sources; use social security data from Fisunoglu et al. (2011), starting in 1992.	Tax/GDP level comparable in 1993 to estimates reported in Barbone and Polackova (1996).
Mongolia	1994-2018	ICTD (1994-2018)		ICTD data goes back to 1986, but does not disaggregate income tax; according to IMF sources, capital gains tax is a tax on corporations; incorporate additional data on property-wealth taxes from HA sources, initially missing in ICTD; significant 'other tax' category between 1993 and 2011 (including stamp duties, royalties, land transactions), and important non-tax revenues.	Comparable tax/GDP estimates in World Bank (2020), though this is not surprising given similar initial data-sources.
Morocco	1965-2018	HA (1965-1999); OECD (2000-2018)	*Annuaire Statististique du Maroc*	Spikes in CIT in 1975 and 2009 appear genuine; 'droits d'enregistrement' in HA are classified as property taxes; social security contributions began in 1959, yet are not observable in HA, so we draw data from Fisunoglu et al. (2011); interpolate 1966, 1971-1973, 1995.	CIT levels in 1980s and 1990s are comparable to those reported by Ministry of Finance (2011).
Mozambique	1975-2018	HA (1975-2014); ICTD (2015-2018)	*Informacao Estatistica*	Interpolate 1991 and 2001; use information from Fjeldstad and Heggstad (2010) to assign income taxes to firms versus individuals in period 2003-2009, and Castro et al. (2009) for 1993-2007 period; HA sources for social security raise concerns, so use data from Fisunoglu et al. (2011); drop in revenue in 2014-2015 appears genuine, based on overlap of data between HA and ICTD.	Limited historical sources; our historical estimates for 1993-2007 are comparable to Lemgruber et al. (2010) and Castro et al. (2009).
Namibia	1991-2018	HA (1991-2018)	*Estimate of Revenue and Expenditure for the Financial Year*	Start data in 1991, given political instability in prior years (full independence achieved), though ICTD data exists in more historical years; perfect match in overlapping years with ICTD data for indirect taxes, though HA estimates of income taxes are 1 percentage point of GDP higher [flagged].	Comparable tax/GDP estimates from World Bank (2020).
Nepal	1976-2018	HA (1976-2005); ICTD (2006-2018)	*Statistical Yearbook of Nepal*	Match in overall tax/GDP level between HA and ICTD in overlapping years, but discrepancy in level by type of tax - this is because ICTD classifies 'excise on industrial product' as indirect tax while HA classifies as corporate income tax [flagged]; social security program began in 1962 (SSA, 2017), we use data from Fisunoglu et al. (2011); drop in PIT and CIT in 2005 are significant, have not found additional sources to corroborate; property-wealth taxes are a minuscule fraction of taxes in 2005 (HA), and disappear entirely in ICTD (2006) [flagged].	Overall good match in tax/GDP levels with World Bank (2020), though limited comparisons available in earliest historical periods.
Netherlands	1965-2018	OECD (1965-2018)			
New Zealand	1965-2018	OECD (1965-2018)		Exists initially unallocable income taxes (category 1300 in OECD classification), but as they are taxes on interest, dividends and withholding (on non-residents' passive investment income), we attribute them to capital rather than labor, minuscule in magnitude.	
Nicaragua	1965-2018	HA (1965-1989); OECD (1990-2018)	*Compendio Estadistico*	HA uses historical IMF data in period 1972-1989; multiple currency re-denominations and revaluations that affect tax/GDP estimates in late 1980s and early 1990s [flagged], but we corroborate with historical sources - including the spectacular spike in 1980s (Machado, 2010; Irvin and Croes, 1987; Gibson, 1996); social security policy began in 1956 (SSA, 2017), we use data from Fisunoglu et al. (2011).	Most importantly, comparable tax/GDP estimates in economic turbulent period of 1980s (Ocampo, 1990; Machado, 2010).
Niger	1965-2018	HA (1965-1999); OECD (2000-2018)	*Annuaire Statistique, Budget Annuel,*	Interpolate 1969-1972; use data from Fisunoglu et al. (2011) for social security contributions prior to OECD coverage, first policy implemented in 1967 (SSA, 2017); HA and OECD match in year of overlap, which could suggest low levels in HA in 1999 are genuine; tax 'sur un role' between 1975 and 1998 is assigned to PIT; use information from OECD in overlapping years with HA to assign 'autres recettes fiscales' in HA in 1999.	Limited existence of studies for historical comparison.
Nigeria	1965-2018	HA (1965-1991); ICTD (1992-2009); OECD (2010-2018)	*Annual Abstract of Statistics*	Interpolate personal income taxes between 1987 and 1990, interpolate overall taxes 2008-2009; important concerns about extent to which resource revenues are truly excluded in HA series, but match in levels is reasonable with ICTD in overlapping years [flagged]; personal income tax represents minuscule share of total taxes in 1980s, corroborated in IBFD (2016); drop in indirect taxes between 1965 and 1969 appears genuine, related to policy reforms.	Estimates of tax/GDP in early 1990s are approximately 1 percentage point lower than in Expo and Ndebbio (1996) and Baunsgaard et al. (2012).
Norway	1965-2018	OECD (1965-2018)		Corroborate significant component of corporate income tax which reflects variation in production of oil and gas.	

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Oman	1972-2014	HA (1972-1989); ICTD (1990-2014)	*Statistical Yearbook*	HA uses historical IMF data, general agreement with other historical HA sources but IMF data carefully and consistently excludes resource revenues; missing property-wealth taxes in ICTD [flagged], though HA series suggests very small in magnitude; additional information used to corroborate absence of personal income tax revenues (KPMG, 2013); sharp dip in 1990-2000 appears genuine; significant nationalization reforms in 1972, making data prior to those events hard to harmonize.	Tax/GDP levels from 2000 onward are comparable to estimates in Besley and Persson (2014).
Pakistan	1965-2018	HA (1965-2018)	*Detailed Statement of Revenue Receipts, State Bank of Pakistan Annual Report Statistical Supplement*	Combine several government publications to ensure we capture national and sub-national taxes in all periods, with the latter an important source of total taxes (Hasan, 1997); 'income tax other than corporation tax' is not entirely PIT (similar challenge in classification in India), use additional information from specific years to assign shares within this category to firms versus individuals.	Historical tax/GDP estimates in 1970s and 1980s are systematically larger, by 1-1.5 percentage points, than Hasan (1997, 2015) though this source only captures national taxes.
Panama	1973-2018	HA (1973-1989); OECD (1990-2018)	*IMF Government Finance Statistics, IMF Article IV Report*	HA uses historical IMF data; initially unallocable revenue between PIT and CIT in IMF, use additional information from Gomez and Sabaini (2005) and Shome (1994) to allocate; decline in revenue in mid-1980s is genuine, likely reflects political transition and violence.	Historical tax/GDP estimate comparable to estimates in CEPAL (1978) prior to 1975.
Papua New Guinea	1976-2018	HA (1976-1999); OECD (2000-2018)	*Compendium of Statistics, Summary of Statistics, Estimates of Revenue and Expenditure for the Year*	Social security contributions started in 1980 (SSA, 2016), use data fro Fisunoglu et al. (2011) prior to OECD coverage, though small in magnitude; volatile tax collection in mid-1970s, confirmed in Duc Thac and Lim (1984).	Historical levels and trends in tax/GDP are comparable to estimates reported in Duc Thac and Lim (1984) for period 1965-1997.
Paraguay	1965-2018	HA (1965-1989); OECD (1990-2018)	*Anuario Estadistico*	Decline in taxes in 1980s and uptick in early 1990s appear genuine, likely reflect political transition period; jump in social security contributions in 1998-1999 is genuine, reflects policy reforms (SSA, 2015); initial jump in tax/GDP in 2016-2017 was due to erroneous GDP value provided to WID, has been corrected.	Limited existence of studies for historical comparison.
Peru	1965-2018	HA (1965-1989); OECD (1990-2018)	*Anuario Estadistico*	Historical values (1968-1978) should be treated with caution, given strong changes in currency and macro-economic conditions [flagged]; use data from Fisunoglu et al. (2011) for social security contributions pre-OECD; jump in total revenues in year of transition from HA to OECD is genuine (HA matches well the level of OECD in post-1989 years); OECD lists two types of wealth taxes levied at national level, while HA lists only one [flagged], though small in magnitude compared to local taxes which are captured in both HA and OECD; interpolate 1965-1967.	Limited existence of studies for historical comparison.
Philippines	1965-2018	HA (1965-1993); OECD (1994-2018)	*Philippine Statistical Yearbook, Annual Budget, Annual Report of the Commissioner of Internal Revenue*	Possible that HA fails to capture sub-national taxes, while OECD does [flagged], though such taxes are estimated to represent less than 5% of total taxes; uptick in tax collection starting in mid-1980s reflects important policy reforms (Reside and Burn, 2016).	Historical tax/GDP estimates are systematically within 1 percentage point of estimates in Reside and Burns (2016), though on average 15% smaller than Mitchell (2003).
Poland	1991-2018	OECD (1991-2018)			
Portugal	1965-2018	OECD (1965-2018)		Assign initially unallocable income tax (OECD category 1300) to individuals versus firms on the basis of information provided in OECD data and comparison with historical IMF reports.	
Qatar	2000-2018	ICTD (2000-2018)		Social security contributions do exist (Deloitte, 2019), we use data from Fisunoglu et al. (2011); IMF reports list 'other revenues' which are a significant source of total revenues, but this likely corresponds to resource revenues [flagged]; excise tax was introduced in 2019, corroborates indirect taxes listed in years of coverage.	
Romania	1991-2018	ICTD (1991-2018)		ICTD estimates for social security contributions are comparable to Fisunoglu et al. (2011), so we do not draw on additional data-sources.	Limited existence of studies for historical comparison.
Russia	1994-2018	HA (1994-1999); ICTD (2000-2018)	*HA draws on IMF sources. IMF Government Finance Statistics, IMF Article IV Report.*	Government statistics published during HA coverage should be interpreted with caution (Gale, 2005) [flagged]; complex property tax system in HA, but estimates of levels are corroborated in Martinez-Vasquez and Wallace (1999) and Chua (2003), and comparison with Treisman (2000) suggests HA series meaningfully captures sub-national taxes.	Historical estimates are comparable to numbers in various IMF reports.
Rwanda	1967-2018	HA (1967-1995); OECD (1996-2018)	*Situation Economique et Conjoncturelle, Bulletin de Statistique, Rapport sur l'Evolution Economique et Financiere, Statistical Yearbook*	Social security contributions begin in 1956 (SSA, 2017), we use data from Fisunoglu et al. (2011) prior to OECD coverage; interpolate 1990-1993, though concerns exists about data quality given unrest in country [flagged].	Tax/GDP estimates in early historical period (1966-1968) are very close to those reported in Cheliah (1971).
Samoa	1983-2018	ICTD (1983-2004); OECD (2005-2018)		Interpolate 1984; observe property tax in early ICTD years as well as in all OECD years, but small in magnitude.	Tax/GDP estimates approximately 2 percentage points lower (10%) than estimates in IMF (2006) for late 1990s.
Saudi Arabia	1994-2018	ICTD (1994-2018)		Non-tax revenues contribute significantly to overall revenues; social security contributions drawn from Fisunoglu et al. (2011). Interpolate 2006-2008.	Limited existence of studies for historical comparison.
Senegal	1965-2018	HA (1965-1984); ICTD (1985-1998); OECD (1999-2018)	*Bulletin Statistique et Economique Mensuel, Situation Economique et Sociale du Senegal*	Drop in revenue in HA in 1970s which is hard to account for, so replace with historical IMF data in that period [flagged]; Boye (1990) describes the tax system as generally stable and steadily growing in 1960s and 1970s, also confirms that PIT outweighs CIT in these periods and overall level of income taxes relative to GDP.	Estimates of historical tax/GDP significantly lower than estimates in Cheliah (1971) for period 1966-1968 [flagged].
Serbia	2000-2018	ICTD (2000-2018)		HA data exists in prior years, but series begins in 2000 given political transition.	Estimates of tax/GDP in mid-2000s is consistent with numbers reported in World Bank (2007).

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Seychelles	1980-2018	ICTD (1980-2007); OECD (2008-2018)		Social security program begins in 1971 (SSA, 2017), use data from Fisunoglu et al. (2011) prior to OECD coverage; PIT collections are practically zero in some years between late 1990s and early 2000s, which appears genuine.	Limited historical comparisons to other studies, but additional IMF data provides comparable estimates in 1980s.
Sierra Leone	1965-2018	HA (1965-1989); ICTD (1990-2018)	*Estimates of Revenue and Expenditure*	Use historical IMF data within HA series, between 1974 and 1989, and HA and IMF perfectly match on levels and type of tax in 1974; social security contributions begin in 2001 (SSA, 2017), prior to that observe payroll tax reported in ICTD and HA; interpolate split between PIT and CIT in 1994-1997.	Limited existence of studies for historical comparison.
Singapore	1965-2018	HA (1965-1999); OECD (2000-2018)	*Yearbook of Statistics*	In HA series, use additional information (Asher and Tayabji, 1980; Tanzi and Shome, 1992) to assign initially unallocable income taxes between PIT and CIT; use data from Fisunoglu et al. (2011) for social security contributions prior to OECD coverage; sharp dip in 1980s reflect genuine tax policy reforms (Joon Chien, 1996); based on overlapping years, suggests that HA is missing a tax on financial and capital transactions which is covered in OECD [flagged], but HA covers other, more significant, wealth-property taxes (Haq et al., 1996; Bird, 1991).	Comparable long-run series of tax/GDP as reported in World Bank (2020).
Slovakia	1994-2018	ICTD (1994); OECD (1995-2018)		ICTD and OECD data match perfectly in 1995.	
Slovenia	1995-2018	OECD (1995-2018)		ICTD and OECD data match perfectly in 1995, but ICTD data pre-1995 does not provide sufficient disaggregation by tax type; SSA (2016) confirms significance of social security contributions; drop in CIT revenue in mid-late 2000s appears genuine (World Bank, 2020).	Estimates of tax/GDP in 1990s are comparable with World Bank (2020).
Solomon Islands	1993-2018	ICTD (1993-2007); OECD (2008-2018)		ICTD data exists in 1980s but has no consistent dis-aggregation between PIT and CIT; interpolate in 1996; ICTD likely only covers national taxes [flagged].	Limited existence of studies for historical comparison.
South Africa	1965-2018	HA (1965-1989); OECD (1990-2018)	*Statistical Yearbook*	Higher level in OECD data than HA at year of transitioning data-source, which coincides with important indirect tax reform, which is also tax category where discrepancy lies, but no years of data-overlap to further investigate [flagged]; reasonable confidence in data's ability to exclude resource revenues, corroborated in additional sources (South Africa Revenue Services, 2015); Ndlovu (2017) refers to a pay-as-you-earn social security scheme dating back to 1963, which we do not separately observe in HA but which could be included under PIT category [flagged].	Comparable tax/GDP estimates in 1970s with Ndlovu (2017) and Koch et al. (2003), in 1990s with Central Reserve Bank, though lower estimates than in Glendlay (2008).
Spain	1965-2018	OECD (1965-2018)			
Sri Lanka	1965-2018	HA (1965-2018)	*Statistical Abstract of Ceylon, Statistical Pocketbook*	Social security from 1958, we use data from Fisunoglu et al. (2011); transaction and property taxes prominent from 1982 onward, unable to confirm if absence pre-1982 is due to policy [flagged].	Ravinthirakumaran (2011) reports comparable tax/GDP numbers for the period 1977-2009.
Sudan	1972-2018	HA (1972-1980); ICTD (1981-2018)	*The National Accounts and Supporting Tables*	Challenging to assign initially unallocable income taxes to firms versus individuals in HA data; use some IMF data in HA coverage, but generally difficult to find reliable information in historical periods [flagged]; use data from Fisunoglu et al. (2011) for social security contributions.	Limited existence of studies for historical comparison.
Swaziland	1969-2018	HA (1969-1998); OECD (1999-2018)	*Estimates of Revenue and Expenditure, IMF Government Finance Statistics*	HA draws on historical IMF data between 1972 and 1989; spike in revenues in 1970s appears genuine and related to economic changes; use data from Fisunoglu et al. (2011) for social security contributions prior to OECD; taxes on international trade substantial in 2000s, corroborated in Ayoki (2018).	Limited existence of studies for historical comparison.
Sweden	1965-2018	OECD (1965-2018)			
Switzerland	1965-2018	OECD (1965-2018)		Withholding tax on interest income in financial institutions is unallocable between PIT and CIT without further information.	
Syria	1965-2007	HA (1967-2007)	*Statistical Abstract of Syria*	HA data extends further back, but series starts in 1967 given political transition; drop in PIT in 2004 appears genuine, confirmed in ICTD data; large increase in CIT from 1980 to 1985 reflects changes in resource environment, but unclear if our series entirely excludes resource revenues [flagged]; social security program began in 1959 (SSA, 2018), we use data from Fisunoglu et al. (2011) for entire series.	Limited existence of studies for historical comparison.
Taiwan	1965-2018	HA (1965-2018)	*Statistical Yearbook*	Spike in taxes in 2000 appears genuine, reflects economic changes; no data from any of main sources on social security contributions, and official government records leave the reporting entry for such contributions blank; either social security is funded through other, general tax sources or social insurance schemes are decentralized and no centralized statistics exist (Chow, 2001) [flagged].	Limited existence of studies for historical comparison.
Tanzania	1965-2018	HA (1965-2018)	*Statistical Abstract, Financial Statement and Revenue Estimates*	Interpolate 1972, 1977, 1993-1995; PIT and CIT bundled in one reported category between 1965 and 1974, assume same split as reported in disaggregated data in 1975; uptick in revenue collection in early 1990s attributed to multiple reforms (IMF, 2009).	Estimates of tax/GDP are 1-2 percentage points lower than reported in Fjeldstad (1995) during 1986-1990 period, and lower than in Osoro (1993) for late 1980s period, though good match on levels and trends with IMF (2009) between 1986 and 2008; difference in estimate may partly be driven by differences in estimates of GDP value (WID and World Bank estimates differ by almost 25%) [flagged].
Thailand	1965-2018	HA (1965-1999); OECD (2000-2018)	*Statistical Yearbook*	Good match on level of tax/GDP between 1960s and early 2000s with Jansen Khannabha (2009), approximately 1 percentage point lower on average in our series; good match in levels by tax-type with Matsumoto (2018). Interpolate 1972, 1980, 1985, 1987, 1994, 1996 - more so than in any other country [flagged]	

Table O1: Notes on Tax Revenue Data Series

Country	Years	Main data sources	Historical archive (HA) data sources	Comments on series	Comparison with other studies
Timor-Leste	2006-2018	HA (2006-2018)	*Data retrieved from online annual budgets published by Ministry of Finance*	Data begins after independence; social security contributions began in 2016 (ILO, 2017); initially unallocable income tax is a withholding tax on personal income and hence classified under PIT; country strongly dependent on resource revenues (Doraisami, 2009; Scheiner, 2015).	Comparable tax/GDP with series reported in IMF (2019), though much smaller than estimates in World Bank (2020), the latter may include resource revenues.
Togo	1966-2018	HA (1966-1999); OECD (2000-2018)	*Annuaire Statistique*	Interpolate 1981-1982; 'transaction tax' in HA is classified as indirect tax, rather than property-wealth tax (Ghura, 1998); mid-1960s to late-1960s were were marked by political transition and coups d'etat, caution reliability of data and absence of historical estimates to corroborate [flagged]; social security began in 1968 (SSA, 2017), use data from Fisunoglu et al. (2011) prior to OECD coverage.	Overall tax/GDP levels are comparable to Ghura (1998) and Stotsky and Woldemariam (1997) during 1980s and 1990s.
Trinidad and Tobago	1965-2018	HA (1965-1989); OECD (1990-2018)	*Annual Statistical Digest*	Large increase in early 1970s seems to be driven by economic volatility and increased inflation with strong increase in CIT collection; dip in CIT collection in late 1990s may be genuine (appears in both OECD and official government records), but unclear nature of shock; historical IMF data corroborates levels of PIT and CIT in 1970s.	Close match in levels of tax/GDP during 1960s and 1970s with Lotz and Morss (1967) and Chelliah (1971).
Tunisia	1965-2018	HA (1965-1999); OECD (2000-2018)	*Annuaire Statistique*	Use data from Fisunoglu et al. (2011) for social security contributions prior to OECD coverage, which is otherwise missing in HA; initially unallocable income tax reported in HA in certain years, limited historical sources to specify allocation [flagged].	Tax/GDP levels comparable in late 1960s and early 1970s with historical IMF reports.
Turkey	1965-2018	OECD (1965-2018)		OECD data includes local taxes since late 1970s.	
Uganda	1965-2018	HA (1965-1991); OECD (1992-2018)	*Financial Summary and Revenue Estimates*	Use historical IMF reports to assign income tax between individuals and firms in HA periods; interpolate 1984, 1991; use data from Fisunoglu et al. (2011) for social security contributions prior to OECD coverage; large and sustained drop in tax collection in late 1970s likely driven by political transition.	Limited existence of studies for historical comparison.
Ukraine	1993-2018	ICTD (1993-2018)		HA is available from early 2000s onward, but comparable in levels with ICTD and we prefer to minimize total number of sources; ICTD numbers for social security contributions are corroborated in government documents, and dip in late 1990s is observed in additional sources (UN-SNA and Fisunoglu et al., 2011).	Polackova (1996) tax/GDP estimate in 1993-1994 is comparable.
United Kingdom	1965-2018	OECD (1965-2018)			
United States of America	1965-2018	OECD (1965-2018)			
Uruguay	1972-2018	HA (1972-1989); OECD (1990-2018)	*Anuario Estadistico*	Earlier HA series exist, but implied levels of tax/GDP are not comparable to historical estimates (Lotz and Morss, 1967), we begin series in 1972 when there is stronger consistency with other studies; very limited collection of PIT in historical periods is confirmed in IMF (1992); CIAT corroborates income tax split between PIT and CIT in 1990s.	Historical estimates of tax/GDP, centered on year of overlap between OECD and HA, match with data reported in IMF (1989) and OECD (1990).
Uzbekistan	1993-2018	ICTD (1993-2018)		ICTD includes social security contributions and which match well in levels with Fisunoglu et al. (2011), including the rise in collection in 2000s; interpolate 2013-2014.	Our tax/GDP estimates are slightly higher than Grigorian and Davoodi (2007) for 1998-2004 period, and slightly lower than Mokhtari and Ashtari (2012) for 2005-2010 period.
Vietnam	1994-2018	HA (1994-2002); OECD (2003-2018)	*Annuaire Statistique, Monthly Bulletin of Statistics, Vietnam Statistical Data in the 20th Century, Statistical Yearbook*	HA sources rely on IMF data; prior to Doi Moi reforms in late 1980s, revenues were largely generated from non-tax sources; social security contributions likely in place during HA coverage, but no data in HA series [flagged]; Bhattarai, Nguyen and Nguyen (2018) corroborate split between PIT and CIT between 1994 and 2010.	Close tax/GDP estimates reported in Cottarelli (2011) for the period between 2001 and 2008.
Yemen	1990-2012	HA (1990-1997); ICTD (1998-2012)	*Statistical Yearbook*	No data past 2012, given political unrest and violence; good match on levels between HA and ICTD in overlapping years; use data from Fisunoglu et al. (2011) for social security contributions.	Tax/GDP estimates are comparable to those reported in IMF (2002) for late 1990s period, generally limited historical comparisons.
Zambia	1965-2018	HA (1965-2018)	*Financial Report, Financial Statistics of Government Sector*	Use social security contributions from Fisunoglu et al. (2011) as it exists during full period, corroborated with HA estimates in specific years; limited comparison with ICTD data as it appears to include resource revenues in CIT numbers in certain years [flagged]. Interpolate 1986, 1990-1991.	Comparable tax/GDP estimates in 1990-2004 with DiJohn (2010) and Weeks et al. (2006), lower tax/GDP estimate than in Colclough (1988) for 1975-1985 period though this may be due to our omission of category 'miscellaneous capital receipts' in HA which we do not count as tax revenue.
Zimbabwe	1980-2018	ICTD (1980-2018)		Data coverage begins after independence; interpolate 1998; social security contributions began in 1989 (SSA, 2017), data matches well between ICTD and Fisunoglu et al. (2011); increase in unallocable income taxes between 2010 and 2018, but limited additional information to clarify allocation between individuals and firms; collapse in tax collection in late 2000s driven by economic conditions.	

Table O2: Weights in Synthetic Control for Trade Liberalization Events

Treated Country	Event Year	Trade Openness	Weight	ETR_K	Weight	ETR_L	Weight	Reference
Argentina	1989	Bangladesh	97.3	Bangladesh	41.6 %	Chile	35.9 %	Goldberg and Pavcnik (2006)
		United States	2.7 %	Haiti	14.1 %	Togo	31.6 %	
		.	.	Bolivia	13.4 %	Jordan	16.8 %	
		.	.	...	...	...	...	
Brazil	1988	Bangladesh	59.8 %	Jordan	35.7 %	Panama	25.7 %	Goldberg and Pavcnik (2006), Dix-Carneiro and Kovak (2017)
		United States	32.2	Sudan	21.2 %	Guyana	21.7 %	
		Japan	6.1 %	Zimbabwe	12.7 %	Chile	14.5 %	
		...	...	...	...	...	...	
China	2001	United States	36.2 %	Congo	41.8 %	Kuwait	31.1 %	Brandt et al. (2017)
		Bangladesh	36.0 %	Nicaragua	26.3 %	Pakistan	22.9 %	
		Dominican Rep.	12.2 %	Gabon	14.2 %	Uganda	20.2 %	
		...	...	...	...	...	...	
Colombia	1985	Bangladesh	50.7 %	Kuwait	67.9 %	Paraguay	45.5 %	Goldberg and Pavcnik (2006; 2016)
		Iran	22.6 %	Gabon	14.6 %	Sudan	15.0 %	
		Guatemala	12.5 %	Sierra Leone	12.6 %	Cameroon	11.5 %	
		...	...	...	...	...	...	
India	1991	United States	76.4 %	Uganda	41.4 %	Lebanon	37.9 %	Goldberg and Pavcnik (2006, 2016); Topalova et al. (2009)
		Bangladesh	23.6 %	Bolivia	14.0 %	Oman	17.6 %	
		.	.	Haiti	4.6 %	Jordan	16.2 %	
		.	.	...	...	...	...	
Mexico	1985	Bangladesh	72.0 %	Sierra Leone	33.2 %	Tunisia	31.1	Feenstra and Hanson (1997); Goldberg and Pavcnik (2006, 2016)
		Uruguay	9.6 %	Bahrain	23.6 %	Zimbabwe	25.8 %	
		Spain	8.0 %	Bolivia	14.7 %	Uruguay	15.9 %	
		...	...	...	...	...	...	
Vietnam	2001	Thailand	42.4 %	Korea	45.8 %	Bangladesh	72.8 %	Goldberg and Pavcnik (2016), McCaig and Pavcnik (2018)
		Ghana	22.6 %	Luxembourg	19.2 %	Myanmar	22.6 %	
		Venezuela	21.7 %	Trinidad & Tob.	17.3 %	Haiti	4.6 %	
		...	...	...	...	.	.	

Notes: This table shows the seven treated countries and the three countries with the largest weight in the synthetic control group for each treated country and outcome (trade openness, ETR_K, ETR_L). For each outcome, the pool of possible donor countries consists of all non-treated countries with a balanced panel over all the pre-event periods that are used in the matching procedure. This table is discussed in Section 5.1.

Table O3: Heterogeneity Impacts of Trade on Taxation by Country Characteristics

Heterogeneity H_c :	Small population (1)	Capital openness (2)
Panel A: CIT rate (first-diff).		
Trade	-0.067***	-0.003
	(0.016)	(0.038)
Trade$*H_c$	-0.022	-0.110**
	(0.065)	(0.051)
Implied coef. for Trade in H_c	-0.089	-0.114***
	(0.057)	(0.027)
1^{st}-stage Kleibergen-Papp F-statistic	7.30	11.41
N	6489	5969
Panel B: ETR_K		
Trade	0.374*	0.511**
	(0.210)	(0.248)
Trade$*H_c$	-0.742	-0.374
	(0.540)	(0.340)
Implied coef. for Trade in H_c	-0.367	0.137
	(0.401)	(0.123)
1^{st}-stage Kleibergen-Papp F-statistic	7.01	11.41
N	6489	5969
Panel C: ETR_L		
Trade	0.190***	0.127
	(0.065)	(0.129)
Trade$*H_c$	0.067	0.187
	(0.242)	(0.209)
Implied coef. for Trade in H_c	0.258	0.315***
	(0.216)	(0.109)
1^{st}-stage Kleibergen-Papp F-statistic	7.01	11.41
N	6489	5969

Notes: This table presents results from estimating heterogeneous effects of trade on outcomes in the full sample of developed and developing countries. Trade is the sum of exports and imports divided by net domestic product. We estimate an IV similar to that in Table A7 but where the interaction term is an indicator for small population (column 1), or an indicator for capital openness (column 2). Small population takes a value of 1 if the country's population in 2018 was below 40 million. Capital openness takes a value of 1 if the country's average value of the Chinn-Ito index (Chinn & Ito, 2006) lies above the median value of all country-years. Both of these heterogeneity dimensions are therefore country-specific but time-invariant. The sample size is smaller in column (2) due to data-availability of the Chinn-Ito variable. The panels differ by outcome: panel a) is the first-differenced corporate income tax (CIT) rate; panel b) is the effective tax rate on capital, ETR_K; panel c) is the effective tax rate on labor, ETR_L. At the bottom of each column and panel, we report the implied coefficient and estimated standard error based on the linear combination of the $Trade$ and the $Trade$*H_c coefficients. We also report the 1^{st}-stage Kleibergen-Paap F-statistic. For more details on the IV, see Section 5.2. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

O.1 Event Study: Further Details

Sample construction Our sample is constructed by applying a synthetic matching procedure to every treated country, for each outcome of interest. The donor pool (the set of all control countries from which to choose the synthetic control group) has to be fully balanced in all pre-event periods. We then pool together all seven treated countries and their synthetic control units. Using this panel, we estimate the DiD:

$$Y_{ct} = \beta^{DiD} \cdot \mathbb{1}(e \geq 0)_t \cdot D_c + \theta_t + \kappa_c + \pi_{Year(t)} + \epsilon_{ct}$$

Here, β^{DiD} can be interpreted as an average treatment effect over the first 10 event-periods (e) after treatment. We run both regressions—the event study and the DiD regression—on the set of main outcomes, and cluster standard errors at the country level. Statistical inference based on small sample size should be approached with caution (Abadie, Diamond and Hainmueller, 2010): we also report standard errors from the wild bootstrap (Cameron, Gelbach, and Miller, 2008) in Table A1.

Moreover, we use the imputation method by Borusyak et al. (2021) to report average treatment effects comparable to β^{DiD} with a technique that deals with issues for two-way fixed effects and heterogeneous event timing. The approach provides a transparent alternative method to the difference-in-difference equation specified above. The average treatment effect (TE) is calculated in three steps:

1. We use untreated countries as well as treated countries in the years before treatment, to estimate unit and (relative) year fixed effects:

   $$Y_{ct} = \theta_t + \kappa_c + \pi_{Year(t)} + \epsilon_{ct}$$

   if $e < 0$ or $D_c = 0$. To bring us closer to the approach developed by Borusyak et al. (2021), we include year and relative time fixed effects.

2. With the fitted values $\hat{\theta}_t$ and $\hat{\kappa}_c$, we now impute unit specific treatment effects:

   $$\hat{TE}_{ct} = Y_{ct} - \hat{\theta}_t - \hat{\kappa}_c - \hat{\pi}_{Year(t)}$$

3. The final step is to average over those coefficients to produce a treatment effect. We report unweighted averages, but heterogeneity in treatment effects could be accounted for by specifying weights.

O.2 Instrumental Variables for Trade

In this appendix, we outline the construction of the two instrumental variables. Both instruments are drawn from Egger et al. (2019), who provide further details.

Instrument based on quantitative trade models The first instrument leverages the structure of gravity models in general equilibrium. These models permit calibration of country pair-year-specific trade costs from trade data, relying on three key assumptions: (i) producers are perfectly competitive and make zero profits or charge a constant markup; (ii) trade costs take the iceberg form; and (iii) aggregate expenditure and its allocation across products are separable. These assumptions imply that bilateral consumption shares towards country o by consumers in country c in year t, denoted π_{cot}, have multiplicative components that are exporter-year-specific (ψ_{ot}), importer-year-specific (ι_{ct}) and pair-year-specific (β_{cot}):

$$\pi_{cot} = \quad \psi_{ot} \quad \times \quad \iota_{ct} \quad \times \quad \beta_{cot}$$

The component ψ_{ot} is proportional to country o's supply potential and captures production costs and gross-of-tax factor income—and might be influenced by both capital and labor taxation. The component ι_{ct} depends on the consumer price index, which varies across years and countries.[67] β_{cot} captures trade frictions across country-pairs and time.[68] The product of the normalized shares gives the bilateral frictions of importing-exporting country-pairs at a point in time:

$$\frac{\pi_{cot}}{\pi_{cct}} \cdot \frac{\pi_{oct}}{\pi_{oot}} = \beta_{cot} \cdot \beta_{oct}$$

Finally, we use $\beta_{cot} \cdot \beta_{oct}$ to compute the average ct-specific costs of exporting and importing, which constitutes the instrument:

$$Z_{ct}^{gravity} = \sum_{o \neq c} [\beta_{cot} \cdot \beta_{oct}]$$

Note that all exporter-year and importer-year factors are removed from the instrument. This instrument is valid so long as the *distribution* of trade costs among country-pairs (not its level) is not influenced by the level of factor incomes or effective tax burdens. Constructing this instrument requires data on country-pair trade flows: we use UN COMTRADE data to construct a large sample of bilateral consumption shares.[69] First-stage regressions with $Z_{ct}^{gravity}$ are shown in Table A2.

[67]Both ψ_{ot} and ι_{ct} may capture country-year-specific trade costs, but the pair-specific component β_{cot} is free of such country-year specific influence.

[68]Egger et al. (2019) exploit the multiplicative model structure about π_{cot} to recover measures of β_{cot}. They assume that transaction costs between domestic sellers and customers are zero, such that $\beta_{cct} = 1$. Both the importer-year component and exporter-year component can then be eliminated by normalizing import and export trade shares by the importer and exporters' consumption from domestic sellers.

[69]We augment our raw data from COMTRADE with data from Harvard Growth Lab, who harmonized importer- and exporter-reported trade flows to expand the coverage and improve the precision of country-partner-year trade flow estimates.

Instrument based on global oil prices & transport distances The second instrument exploits spatial heterogeneity across countries in a way that interacts with oil price shocks. This instrument is based on global oil price changes over time and within-country transportation distances from cities to the nearest port.[70] The instrument is the variance of the product oil price p_t^{oil} $\times$ distance d_c^k across cities k in country c in year t:

$$Z_{ct}^{oil-dist} = \frac{1}{2} \sum_{k=1}^{3} [(p_t^{oil} d_c^k - p_t^{oil} \overline{d_c})^2]$$

where $\overline{d_c}$ is the average city-port distance in country c. This variance increases in countries whose main cities are far from the nearest port and far from each other, which implies a larger change to transportation costs following a global oil price shock in spread-out countries than in countries with concentrated populations. It is this transportation-cost shock that the instrument captures.[71]

This second instrument does not hinge on theoretical assumptions. Instead, this instrument is valid so long as the country-specific distribution of trade-costs, induced by the interaction between global oil price shocks and a country's fixed spatial concentration, is not correlated with contemporaneous changes in factor incomes and effective tax rates. First-stage results for $Z_{ct}^{oil-dist}$ are presented in Table A2.

[70] For the former, we retrieve the OPEC Reference Basket benchmark world price of crude oil. For the latter, we measure road distances from the three largest cities (according to UN population statistics) to their nearest port, using SeaRates international shipping logistics calculators.

[71] Alternatively, one could measure the variance in distance and then multiply it by the global oil price. The distribution of the variance instrument $Z_{ct}^{oil-dist}$ across country-years would not change; the only impact would be a level-shift by the price. We consider the main approach to more closely capture the sensitivity of transport costs to spatial concentration, but results based on this alternative variance measure are similar.

O.3 Comparison with Pre-existing ETR Series

We compare our methodology to pre-existing ETR series. The main differences are summarized in Table B3.

In McDaniel (2020), updated from McDaniel (2007), there are two main differences with our benchmark methodology. First, the author assigns the capital-share in mixed income based on the observed factor share in the rest of the economy, while our benchmark assigns a fixed share (25%). Second, the author assumes that labor and capital in PIT are taxed at the same rate. This is the same assumption as in Mendoza et al. (1994), and differs from our benchmark where we create an allocation of PIT to capital that varies by country and year, $(1 - \lambda_{PIT,ct})$. These methodological differences are reasonably captured in our robustness checks. For the first difference, note that this choice effectively amounts to using the observed capital-share in the corporate sector to assign the capital share in mixed income (see equation 3 in McDaniel (2007)). This corresponds directly to one of our robustness checks (Panel B in Figure 3).

We can relate the second difference to our robustness check where we vary the capital share of PIT from 0% to 30% (Panel A in Figure 3). If labor and capital face the same tax rate, then the capital share of PIT increases in the capital factor share and in the share of capital that is taxable. Using the empirical measures for taxable shares established in the US (Piketty et al., 2018), and assuming both factors face the same rate, the capital share of PIT that would result at the 99^{th} percentile of observed capital factor shares in our full sample is $1 - \lambda_{PIT} = 0.305$. In other words, our robustness check which implements $1 - \lambda_{PIT} = 0.30$ constitutes a meaningful upper bound on the capital share in PIT that would result from any observed factor shares in our sample and assuming capital and labor pay the same rate and have taxable shares as measured in Piketty et al. (2018). Of course, under the assumption that both factors are fully taxable (unrealistic given empirical findings in Jensen (2022)) and face the same rate, the capital share of PIT would be equal to the observed capital factor share. Our benchmark methodology takes a step towards trying to measure the taxable factor shares as they vary across countries and time, with a $1 - \lambda_{PIT,ct}$ at the 99^{th} percentile that equals 0.32. Future work could improve on this measurement, by combining additional information from national accounts and tax records.

The ETR series that would result from applying the methodology in McDaniel (2020) to our sample is therefore reasonably bounded by our robustness checks which assign capital's share of mixed income based on the corporate capital share, and which vary the capital share in PIT between 0% and 30%.

In addition to these main differences, McDaniel (2020) considers property taxes paid by households as consumption taxes and property taxes paid by businesses as capital taxes, while our series considers all property taxes to be capital taxes. This difference is unlikely to be quantitatively significant. Finally, McDaniel (2020) uses tax data from national accounts, while we rely on various public finance sources.

In Panel B of Figure B1, we use McDaniel (2020)'s specific sample. The trends are similar between series. When weights are applied, our benchmark series is on average 18.75% higher in levels than McDaniel (2020). This wedge arises from the methodological differences (which we can account for in our robustness checks), and possibly from the differences in tax data-sources.

In Kostarakos and Varthalitis (2020), which applies the methodology in Carey and Rabesona (2004) to data from more recent years, there is one main methodological difference with our benchmark. In particular, the authors estimate relative labor income of self-employed on the basis of observable characteristics and a comparison with the wage of employees, while our benchmark method assigns a fixed labor share to mixed income. However, this alternative method corresponds closely to the method in ILO (2019), which we implement as a robustness check. Thus, this robustness check (Panel B of Figure 3) meaningfully captures the ETR series that would result from applying the methodology in Kostarakos and Varthalitis (2020) to our sample.

One additional difference is that Kostarakos and Varthalitis (2020) assume social security contributions are deductible from the taxable income of households while our method follows national accounts convention and assumes they are not. We confirm that implementing this change in our series does not meaningfully alter the results (available upon request). Finally, Kostarakos and Varthalitis (2020) draw their tax revenue data from a different source (Eurostat) than us.

In Panel A of Figure B1, we use Kostarakos and Varthalitis (2020)'s specific sample. The trends are similar between series. When weights are applied, our benchmark series is on average 14.2% lower in levels than Kostarakos and Varthalitis (2020). This wedge arises from the methodological differences (which we can account for in our robustness checks), and possibly from the differential treatment of social security contributions and from the different data-source for tax revenue.

O.4 Discussion of $\overline{ETR}_C^K$

Our measure of the backward-looking average effective tax rate on corporate profits, $\overline{ETR}_C^K$, is related to, but also distinct from, the forward-looking measures of the statutory tax burden in the corporate sector in developing countries (Section 2). There are two main reasons why these measures differ.

The first reason is that the measure of corporate profits in $\overline{ETR}_C^K$ is based on national accounts, which differs both empirically and conceptually from how corporate profits may be measured using tax data. Empirically, the data-sources for national accounts include corporate tax returns but also non-tax sources such as industrial censuses and surveys. The measure of corporate profits based on national accounts may therefore include profits which are not reported in tax returns. Indeed, the national account guidelines explicitly try to account for mis-reported profits and corporate profits are usually found to be larger in value when measured from national accounts than from tax returns (Lequiller & Blades, 2014). For this same reason, constructing an appropriately-weighted backward-looking firm-level effective tax rate based on taxes paid and profits reported in tax returns may not give the same value as $\overline{ETR}_C^K$. This firm-level measure is analyzed in Section 6; see also Dyreng et al. (2017) and Egger et al. (2009) for firm-level estimates in large samples focused on developed countries.

There are also conceptual differences, which are discussed in detail in IMF (2014), Lequiller & Blades (2014) and Ueda (2018). Consumption of fixed capital in national accounts adjusts for inflation and is estimated according to physical and economic laws of depreciation, whereas companies sometimes measure depreciation without regard to inflation and may shorten or lengthen the time of amortization according to tax advantages. In addition, inventory appreciation (the net gain in inventory) is usually accounted for in company profits but not in national accounts. Moreover, expenditure on intellectual property is counted as investment in national accounts, but may be listed as intermediate consumption by companies on their tax returns. Finally, some sources of property income (e.g. investment valuation increases; resource rents paid vs. received) and capital gains (e.g. sale of subsidiaries or currency transactions) are counted in company profits but not in national accounts.

The second reason is that the corporate statutory tax burden varies across firms due to economic variables, including sector, size and profitability (R. Kumar & James, 2022, Devereux, Griffith and Klemm, 2004). Changes in these economic variables will be reflected in $\overline{ETR}_C^K$, but may not be fully captured in the statutory measures.

Relation between $\overline{ETR}_C^K$ and CIT efficiency Our measure of the average corporate effective tax rate, $\overline{ETR}_C^K$, is related to empirical work done by IMF (2014) to measure CIT-efficiency in developing countries. IMF defines CIT-efficiency as actual corporate income tax (CIT) revenues divided by the product of the standard CIT rate and the gross operating surplus of corporations from national accounts, OS_{CORP}. CIT-efficiency is therefore related to $\overline{ETR}_C^K$ as follows:

$$\text{CIT-efficiency} = \overline{ETR}_C^K \cdot \frac{1}{\text{CIT rate}}$$

We have information on the CIT rate, so we can compute the CIT-efficiency measure using our data. In turn, we can limit the comparison to the sample of developing countries covered in the IMF study ('Non-OECD' countries). Before comparing the two series, we note the remaining methodological differences:

1. Our measure of OS_{CORP} is net of capital depreciation, while the IMF measure is not. This likely leads our variable to be higher in levels, though it is not clear how it would affect trends over time (see Appendix IV in the IMF study).

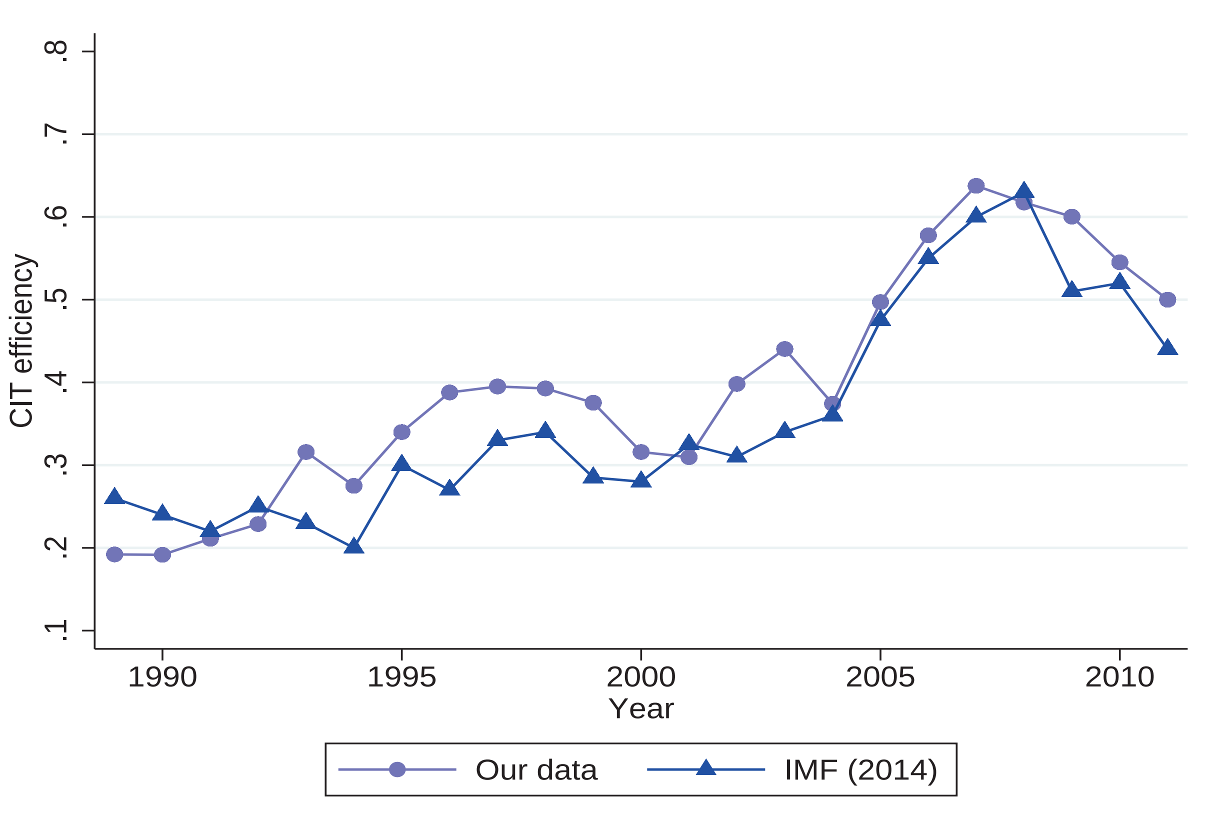
2. Our sample does not contain data for Malta, which is one of the developing countries ('Non-OECD') included in the IMF sample. Moreover, we can limit the comparison sample to the range of years indicated in the IMF study (1989-2011), but we cannot verify that our comparison sample contains exactly the same country-year observations.

3. In the IMF study, the data for both the CIT revenue collected and the CIT rate are taken from the IMF's Fiscal Affairs Department. Our sources for both CIT revenue and the CIT rate differ (see Section 3 and 6 in paper for details).

The average unweighted CIT-efficiency series that results from using our data but restricted to the IMF sample is compared to the IMF series in Figure O10. Our values are a little higher in levels on average, which may reflect the deduction of depreciation. Despite this, as well as differences in underlying data-sources, the trends match well between the two series.

In regression results not shown, we find that trade has a positive impact on the CIT-efficiency measure. This is consistent with the hypothesis in the main paper that trade improves effective corporate taxation in developing countries.

Figure O10: CIT efficiency – Comparison of Our Data with IMF (2014)



Notes: This figure compares the values of corporate income tax (CIT) efficiency between the series estimated using our data (light blue circle line) and the series estimated in IMF (2014) (dark blue triangle line). We restrict our data to the 'Non-OECD' sample of developing countries used in the IMF study, with the exception that Malta is included in the IMF study but does not exist in our data. The reported IMF (2014) series is approximate, as the numbers have been extracted from Appendix Figure 1 of the study based on visual inspection. CIT efficiency is defined as the ratio of CIT revenues collected divided by the product of corporations' operating surplus and the CIT rate. For more details, see Section O.4.

Effective Tax Rates Database: Case Studies Report

Pierre Bachas*, Matthew Fisher-Post†, Anders Jensen‡,
Xabier Moriana-Armendariz§, Gabriel Zucman¶

January 2024

Abstract

This report presents detailed country-by-country guidance on the creation of time series data from 1965 to 2020 on tax revenues, factor income shares and effective tax rates for the countries shown in the Atlas of the Offshore World (link). Originally based on the paper Bachas, Fisher-Post, Jensen, and Zucman (2022), the guideline details the data sources of historical tax revenue data and explains the choices and adjustments made. Beyond the well-known institutional datasets on government tax revenue (OECD, ICTD, IMF), we have digitized and harmonized thousands of archival public finance documents from developing countries, in order to construct the long-run time series. In the following pages, we discuss these sources, as well as all relevant methodological considerations and academic sources to help produce reliable time series estimates for each country. In doing so, we provide a transparent resource on the rationale for our data input, while welcoming suggestions for further improvement and expansion of the database.

- **Methodology and How to Read this Document:** click here
- **Country List:** click here

This document is work in progress. At the moment, we cover countries with more than 15 million inhabitants and some additional countries (67 in total). The sample of countries is currently being expanded.

*ESSEC-Business School and World Bank Research, pbachas@worldbank.org
†Paris School of Economics, mfp@psemail.eu
‡Harvard Kennedy School and NBER, anders_jensen@hks.harvard.edu
§Paris School of Economics, xabier.moriana@psemail.eu
¶Paris School of Economics, UC Berkeley and NBER, zucman@berkeley.edu

Methodology and How to Read this Document

Country List

Please follow the link to see the list of countries included in this document. We emphasize that this is a work in progress. At the moment, we cover 67 largest countries (>15 million inhabitants)[1]. Our team is currently working to complete the case studies for the remaining countries in the full sample.

(i) Introductory note

The present paper provides a country-by-country guideline on the construction of the effective tax rates on capital and labor incomes. In addition, it offers information on the effective tax rates on corporate profits, and on dis-aggregated tax revenue as a share of national domestic product.

Since national accounting statistics are compiled following internationally standard concepts and methods, effective tax rates (ETRs) are conceptually comparable over time and across countries. These series provide a picture of tax burdens on capital and labor in these countries. By considering the tax revenues actually collected (rather than statutory rates), ETRs show the net past effects of all tax rules—base reductions, exemptions, tax credits—and of tax avoidance and tax evasion. Low effective tax rates can result from tax avoidance or tax evasion practices, but may also result from policy choices.

(ii) The methodology in brief

The effective tax rate on labor (ETR_L) is the total amount of taxes actually collected on labor income, divided by total labor income in the economy; similarly, the effective tax rate for capital (ETR_K) is the total amount of taxes actually collected on capital income, divided by total capital income in the economy. Taxes and factor incomes are respectively allocated to the numerator and denominator as follows:

Tax Revenue Allocation:

- Corporate income taxes, wealth taxes, and property taxes are allocated to capital.
- Payroll taxes and social security payments are allocated to labor.
- Personal income taxes are allocated partly to labor and partly to capital (see Bachas et al. (2022) for details).

Factor Income Allocation

- Labor income equals compensation of employees plus a share of mixed income (operating surplus of private unincorporated enterprises) – see Bachas et al. (2022) for details[2]
- Capital income equals the remaining share of mixed income, plus corporate profits (i.e., operating surplus), plus rental income (i.e., operating surplus of households).

[1] Myanmar (54 million inhabitants) and Venezuela (28 million inhabitants) are not included at the moment due to the lack of quality data for the historical period.

[2] Note that factor incomes (both capital and labor) are based on a dis-aggregation of the net domestic product (NDP). The NDP subtracts the consumption of fixed capital (capital depreciation) from gross domestic product (GDP). NDP is thus lower than GDP by 10% on average.

In the construction of the ETRs, we face **two challenges**:

1. **Tax Revenue:** what share of personal income tax (PIT) revenues should be allocated to capital versus labor. Starting from a baseline where 15% of PIT revenues derive from capital (consistent with US measures in Piketty et al., 2018), two adjustments at the country-year level are performed. The proportion of capital revenues within PIT is increased in countries with a high PIT exemption threshold in the income distribution and the proportion is lowered in countries where dividends face lower tax rates than wages. We report the range of values for the capital share in each country. In the companion paper (Bachas et al., 2022), two bounding scenarios are constructed.
2. **Factor income:** what share of mixed income (unincorporated enterprises) should be allocated to capital versus labor income. For this case-study, a 75% labor vs. 25% capital split is assumed. In the companion paper (Bachas et al., 2022), two bounding scenarios are constructed.

We refer to the paper and to the Methodology Note available in the website for further methodological issues.

(iii) The database in brief

Tax Revenue: The tax revenue data dis-aggregates revenues by type, following the OECD Revenue Statistics classification of taxes. We rely on three sources:

1. When available, OECD Revenue Statistics data is the preferred source, as it covers all types of tax revenues and goes back to 1965 for OECD countries.
2. Data from ICTD is added, which includes most developing countries, and with coverage that starts in the 1980s.
3. To complement these sources, the team conducted an archival data collection. Within the concept of 'Historical Archive', we include data from:
   - Lamont Library at Harvard University (Historical public budgets and national statistical yearbooks)(website link).
   - Offline IMF Government Finance Statistics (1972-1989) (website link).
   - Annual Reports from the country's Ministry of Finance/National Central Bank for recent years.

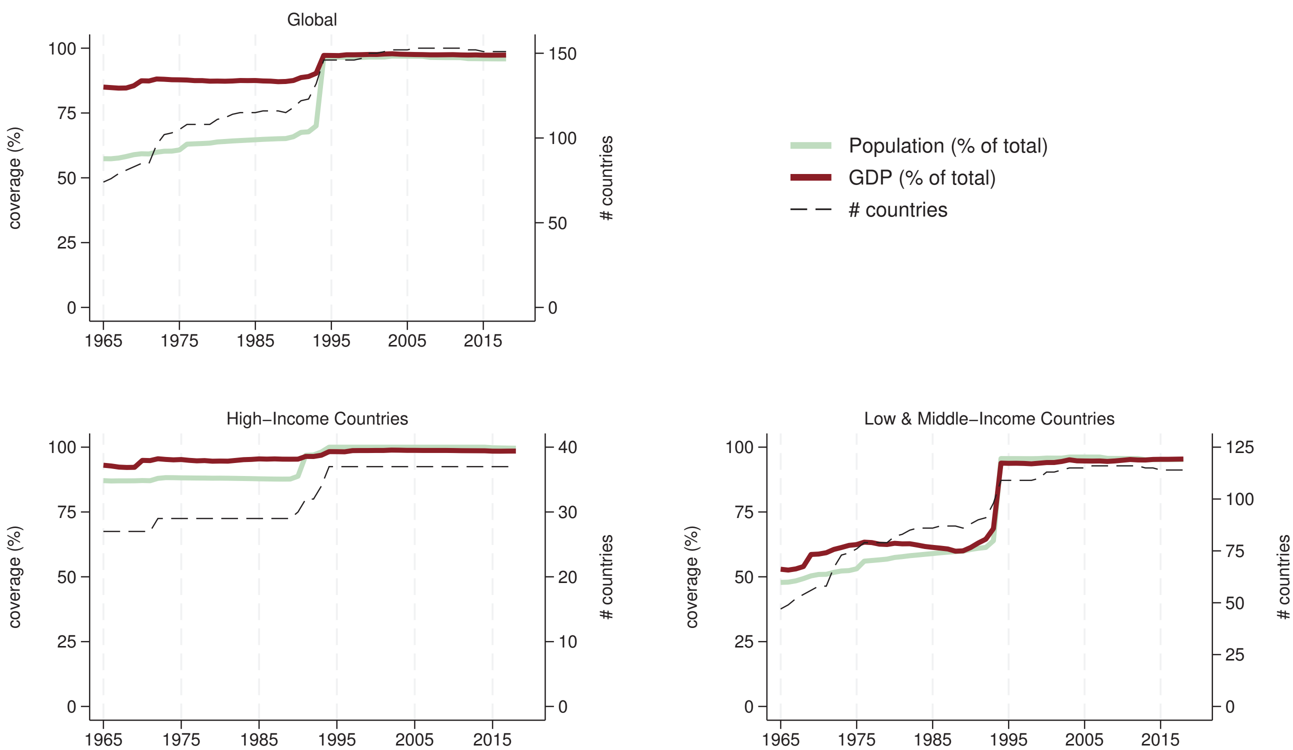
Factor income: The factor income dataset is based on the construction of a panel of national accounts. It comes from two sources:

1. The SNA2008-framework online repository, which has global coverage in recent decades.
2. The SNA1968-framework offline repository, which covers historical observations from the 1960s and 1970s for most countries.
3. If there is no data for a sub-component of factor incomes, we: 1) recover its value from the other SNA dataset and using national accounting identities; or, if not possible, 2) impute values for the component following the procedure from Blanchet, Chancel, Flores, and Morgan (2021).

The final sample contains 7070 country-year observations in 154 countries, over the period 1965-2020. Note that the companion paper (Bachas et al., 2022) considered the 1965-2018 range of years. The number of countries starts at 75 in 1965 and grows to 105 by 1975 (due to independence or country creation). The key jump in coverage—from 116 to 146 countries—occurs in 1994 and corresponds to the entry of ex-communist[3] countries, including China when it modernized its public finances (see Notes on page 6).

Figure 1 shows graphically how the dataset is effectively composed of two quasi-balanced panels: the first covers 1965-1993 and excludes communist regimes, accounting for 85-90% of world GDP; the second covers 1994-2020 and includes former communist countries, accounting for 98% of world GDP. Most of the ex-communist countries are low-middle income countries, making the jump bigger for this subgroup. The small drop in 2019-2020, coming from low and middle-income countries, is explained by the lack of recent data for Venezuela, Syria and Yemen.

Figure 1. Data Coverage



[3]We use ex-communist to refer to basically former soviet countries, China and Vietnam. We do not include in this group other communist/socialist regimes in other parts of the world, such as the African socialist countries in certain periods (Ethiopia, Mozambique or Angola).

Country List

Small Countries (< 40 millions)
*(Work in progress)

1 Afghanistan

Sources

Source	First year	Last year	Interpolation
ICTD	2003	2020	

Time series We start the data series for Afghanistan in 2003. We do have data for prior years: historical archive data for 1973-1978 and ICTD for 1982-1989. However, we exclude this period because these figures are either too short in time (historical archive) or they are not disaggregated while being quite volatile (ICTD).

Harmonization

The main adjustments for Afghanistan are made on income taxation, social contributions and property taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT with questionable values for 2006-2010 and not available after 2017. Use information in additional documents, and interpolation of proportional ratios for 2006-2010 using 2005 and 2011 as references and extrapolation of proportional ratios after 2017 • Small unallocable income tax revenue during the whole period • Capital share of PIT: mean = 30%; constant
Social Contributions	Minuscule but non-zero values during the whole period. Extrapolation of values after 2017
Property Taxation	Minuscule but non-zero values during the whole period. Extrapolation of values after 2017
Decentralized Revenues	Not determined

Notes

2 Algeria

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2020	1967, 1970-71, 1974

Time series We use archival data from 1965 through 2020. Years 1967, 1970-1971 and 1974 are interpolated.

Link to historical archive data: click here.

Harmonization

The major adjustments for Algeria are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT available all the way back to 1965 • All unallocable income tax revenue is considered as PIT, based on information in additional documents. • Capital share of PIT: mean = 20%; ↓ trend (25% to 18%)
Social Contributions	Minuscule but non-zero during the whole period. We rely on external sources (RPC data)
Property Taxation	Not included
Decentralized Revenues	Not determined

Notes

- Our archival data differs from IMF historical data. However, this difference is completely driven by the consideration of the tax "*Taxes sur le chiffre d'affaires*". The IMF considers it as an indirect tax, whereas we include it as a corporate income tax in our historical archive data, based on additional policy reports.

3 Argentina

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	
OECD	1990	2020	

Time series We use archival sources from 1965 until 1989, referring to OECD for the period since then.

Link to historical archive data: click here.

Harmonization

The major adjustments for Argentina are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for 1961-69 and 1974-1989. Interpolation of proportional ratios using prior and later years for which we have disaggregated data, as well as information from additional policy documents. • Small unallocable income tax revenue (<0.4%) in OECD data (1990-2020) • Capital share of PIT: mean = 18%; ↓ trend (20% to 15%)
Social Contributions	We rely on ICTD for social contributions for 1985-1989, as it matches with OECD in 1990-2016 (and our historical archive (Alvaredo, 2010) clearly underestimates social contributions in this period). Prior to 1985 we trust historical archive data as it matches for overlapping years (1980s) with ICTD
Property Taxation	Included
Decentralized Revenues	Not determined for historical archive data. Excluded local government tax revenues but included provincial revenues for OECD data

Notes

4 Bangladesh

Sources

Source	First year	Last year	Interpolation
Historical archive	1976	2000	1980-81
ICTD	2001	2020	

Time series Our data starts in 1976. The series follows historical archive data from 1976 through 2000, then uses ICTD data from 2001 through 2020. We interpolate years 1980 and 1981.

Link to historical archive data: click here.

Harmonization

The main adjustments for Bangladesh are on income taxation and on social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available in 2017-2020. Extrapolation of proportional ratios from 2016 • Unallocable tax in historical data prior to 2001. Part of it (from non-corporate income tax) included in PIT, based on additional information in policy documents. • Capital share of PIT: mean = 30%; constant pre-2010 (30%), ↓ trend post-2010 (30% to 26%)
Social Contributions	Minuscule but non-zero levels. Use of RPC data for the whole period, extrapolating years 2014-2020. SSA (2017) indicates that the social security policy began in 1971
Property Taxation	Minuscule but non-zero property tax revenue in archival data (1965-2001) while not present in ICTD (2001-2020). No wealth tax but real estate taxes. Although we do not have data for property tax revenue after 2001, we expect it to be very small
Decentralized Revenues	Our estimates are most likely central government only, as there is no mention of province-level revenues. However, external sources' estimates of the tax/GDP ratio also seem to exclude decentralized revenues

Notes

- Early literature shows similar values (Ghafur & Chowdhury, 1988) with the likeliest source of variation being the difference in the GDP denominator.
- The major reform and policy landmarks were in 1984 (an income tax ordinance under military rule), 1991 (instituting the VAT) and more recently in 2009 (to reform the VAT). However, tax/GDP ratio has remained in very low values until nowadays.

5 Brazil

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	1968-1975 (partial)
IDB-CIAT	1990	2020	

Time series We use IDB-CIAT data for detailed data on the tax structure from 1990-2018, and refer to historical archive data for the years prior. For years 2019 and 2020, we use information from the joint publication by the OECD and IDB-CIAT, available at OECD website.

Link to historical archive data: click here.

Harmonization

The main adjustments for Brazil are on income taxation, on social contributions and on property taxation.

Category	Adjustment
Income Taxation	• Interpolation of income tax revenue between 1968 and 1975 • Extrapolation of the PIT vs. CIT split back to the 1960s, from 1976 • Income taxes on non-residents considered as a corporate income tax • Unallocable income tax revenue 2019-2020. Extrapolation of proportional ratios of the split of CIT/PIT from 2018 • Capital share of PIT: mean = 25%; ↓ trend (30% to 22%)
Social Contributions	No data prior to 1980. We use RPC data (Real Political Capacity database, alias 'RPC') for social security pre-1980, 'backcasted' by its ratio in 1990 with IDB-CIAT
Property Taxation	• There is no net wealth tax on individuals in Brazil, but there is a municipal real estate tax and a federal tax on rural land • Financial transactions tax (IOF): considered here as property tax whereas in original source (Afonso, Araujo, and Vianna (2004), used for 1980-1989) was considered an indirect tax. Interpolation of the structure of this tax holding constant its 1980 value back to 1965 and adding it to the other property taxes (observed in Chelliah (1971)) while subtracting it from the indirect tax bill
Decentralized Revenues	IDB-CIAT accounts for both central government revenue and decentralized local revenues. On the contrary, the historical archive data only refers to the central government. However, income tax is only assessed at the central government level, so we can at least use the historical archive material for this purpose

Notes

6 China

Sources

Source	First year	Last year	Interpolation
Historical archive	1994	2007	
OECD	2007	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include China before its first year of modern tax system, 1994 (see note below). For years prior to 2007, our Chinese government time series comes primarily from historical archive data online. We have used digitized statistics from China's Statistical Yearbook (also online) and physical copies of the Compendium of Statistics from Harvard archives, as well as long-run public finance data available online from the National Bureau of Statistics (see link). After 2007 we use OECD data.

Link to historical archive data: click here.

Harmonization

The main adjustments for China are on social contributions before 2019.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1994 • Capital share of PIT: mean = 21%; ↓ trend (27% to 15%)
Social Contributions	OECD data is not available before 2019. We use historical archive data that covers until 2018
Property Taxation	Included
Decentralized Revenues	China's fiscal policy has been always under the overall authority and observation of the central government. While it would be possible to examine and disaggregate further the patterns of tax on a decentralized (province-by-province) level, there is little doubt that our estimates are capturing all public revenues

Notes

- **Chinese Modern Tax System:** In our benchmark setting, we only include formerly communist economies into our data starting in 1994. Given China's weight in the global economy, it is worth reviewing the reason for that choice. The tax revenue data for China covers most of our sample period although its quality improves markedly in the 1980s.

Prior to the 1980s, China had a command economy model of 'profit delivery,' in which the state directly received the revenues of profitable SOEs, and subsidized unprofitable ones. A corporate income tax first appears in China in 1983-84, but the majority of the base continues to be state-owned enterprises. In 1985, the tax system was further reformed into a 'fiscal contracting' system whereby firms negotiated a fixed lump-sum payment (regardless of economic outcomes), which cannot be split into labor versus capital taxes (nor into consumption taxes). We therefore exclude the 'pseudo'-CIT revenue dating from 1985 through 1993.

Rather, we consider that China's modern tax system began in 1994. Lou and Wang (2008) shows that, in 1994, China established for the first time a central tax administration; reformed the 'fiscal contracting' system; unified the PIT; created a VAT; and reduced 'extra budgetary' (non-tax) revenues. Thus from 1994 onward we can categorize tax revenue precisely by type, assign them to capital or labor, and estimate our ETR.

- **Prior Literature on China:** It is worth highlighting two important, external sources of revenue estimates: the ICTD and those of UCSD professor Barry Naughton (2007). ICTD patterns agree with ours on broad orders of magnitude, for the years in which ICTD revenue data for China is available, but these do differ slightly in their classifications of tax revenues.[4]

  Naughton, in turn, cited statistics from an OECD study (Bouin, Coricelli, & Lemoine, 1998). He estimates total revenue in 1978 (not included here due to the reason above) at 34 percent of GDP, of which 22 percentage points were profits and taxes from state-owned enterprises and 11 from general sales taxes. Our different gloss on the breakdown of public revenues at that time is actually from two causes: (i) ours is a higher estimate of GDP (verified from the World Bank and the World Inequality Database); and (ii) a different treatment of state-owned enterprises' corporate 'profits.' Therefore we do not treat as entirely *tax* revenue, and we apply the government's own nomenclature of revenue categories to describe tax revenue as roughly 40 percent of overall budgetary revenues (and even less when one includes extrabudgetary revenues). Our total (budgetary) revenue numbers, in raw levels, are always within 10 percent of the OECD estimates from 1978-94. By 1994 the discrepancy in their treatment of public revenue disappears; Bouin et al. (1998)'s estimate of tax (and total budgetary) revenue drops below 15 percent of GDP—in accordance with our own estimates.

[4]It seems likely that ICTD has classified the 'unified business tax' as a indirect tax series revenue (specifically, a type of consumption tax), while we have instead categorized this as a corporate income tax. Also, ICTD does include some personal income tax revenue in the late 1990s, more than 1 percent of GDP. It is unclear from where this number is retrieved for the Article IV report, as our data from the national statistical bureau lists minuscule amounts until after the 1999 reform.

7 Colombia

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	
OECD	1990	2020	

Time series We use OECD data for the period since 1990, but refer to archive sources for the historical period from 1965.

Link to historical archive data: click here.

Harmonization

The main adjustments for Colombia are on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available before 1990. Extrapolate the split in 1990 (first OECD year) back through the historical period, while using the total income revenue levels of the historical data, and based on additional information in policy documents. • Unallocable income tax revenue after 1990. We include all of it in corporate income tax revenue following the evidence shown in the W. Bank (2014) • Capital share of PIT: mean = 14%; ↓ trend (15% to 12%)
Social Contributions	Included
Property Taxation	Marginal importance before year 2000, shown in both data sources
Decentralized Revenues	High level of certainty that the data includes both central and local government revenues

Notes

- For the historical data period, our estimation matches previous literature. We show a similar raise in tax revenue for the 1970s as García-García and Guterman-Bromberg (1988). For the 1980s, our estimations are below McLure Jr (1992) and above Junguito, Rincón, et al. (2004). This discrepancy is explained by different considerations of social security and indirect taxes, supporting the idea that decentralized taxation is either already included implicitly, or not significant prior to 1980.

8 Democratic Republic of the Congo

Sources

Source	First year	Last year	Interpolation
Historical archive	1968	1990	1973
ICTD	1991	2020	1992-95

Time series We use historical archive data from 1968 to 1990, then ICTD data from 1991 to 2018. We refer to ICTD for the overall level of tax/GDP in the 1980s but refer to archive data for its disaggregation ('within' component shares relative to that total). ICTD data improves in quality only after 1996 (when they refer to IMF Article IV reports), so we interpolate relative shares of taxes (within the overall tax/GDP level) for the years 1992-95. We also interpolate year 1973.

Link to historical archive data: click here.

Harmonization

The main adjustments for the Democratic Republic of the Congo are on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for years 1981 and 1982. Interpolation of the proportions of PIT vs. CIT within overall income tax revenue using 1980 and 1983 as reference • Unallocable income tax in 1977-1990, and especially relevant during the first part (1977-1982). It includes a direct tax category called 'divers' that we decide not to allocate exclusively to CIT or PIT, but instead assign shares based on additional policy documents. • Capital share of PIT: mean = 24%; ↑ pre-2000 (19% to 30%) and ↓ post-2000 (30% to 26%)
Social Contributions	Minuscule value in ICTD, missing in historical archive. The very low values in ICTD indicates that its absence in our archival data does not represent a significant gap in the Time series
Property Taxation	Data available only for 2010-2020. According to IBFD, there is no wealth tax in DRC, but there is a property tax (for all but agricultural or non-profit uses). However, its value according to ICTD is minuscule for recent years. Without having the data to confirm that, we expect that property tax do not play a significant role for total revenue collection in historical periods
Decentralized Revenues	Not determined

Notes

- The early 1990s drop in all revenue categories is genuine, as this was a period of hyper-inflation. The massive decline of (unindexed) tax revenue collection during hyperinflation is known as the Tanzi-Olivera effect. This is the common explanation for the 1990s for the DRC suggested in prior literature (De Herdt, 2002; Nachega, 2005).

- The volatility of both direct and indirect taxation (and within proportions of each) are likely genuine, as DRC was a resource-dependent economy and subject to shocks in the world price of copper (cf. De Herdt (2002)).

9 Egypt

Source

Source	First year	Last year	Interpolation
Historical archive	1965	1989	1965-67, 1970-71, 1973-74
ICTD	1990	2001	
OECD	2002	2020	

Time series We use archival data for the period 1965-89 (including IMF historical data 1975-89), then ICTD from 1990 until 2001, with OECD data from 2002 until 2020. We interpolate years 1965-67, 1970-71 and 1973-74.

Link to historical archive data: click here.

Harmonization

The main adjustments for Egypt are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for 1998. Interpolation of proportional ratios using 1997 and 1999 • Unallocable income tax for OECD (2002-2008). It is entirely due to a tax called 'Tax on movable capital revenues from C.B.E' [Central Bank of Egypt] which we allocate to capital based on additional policy documents. • Capital share of PIT: mean = 20%; ↓ (25% to 15%)
Social Contributions	RPC database for social security revenues in the periods 1960-1974 and 1990-2001, when OECD data begins. For the period 1975-89 we rely on IMF historical data (historical archive)
Property Taxation	Available in the three sources. There is both a real estate tax and an agricultural land tax
Decentralized Revenues	We do not observe any decentralized revenues in the OECD data, nor in our archival (nor historical IMF) nor ICTD source

Notes

- In the overlap periods (1987-89 and 2001-2002) the series match almost perfectly across sources. In addition, our series match the World Bank online database for recent decades (after 1975) as well as they match prior literature for the late 1960s and 1970s (Nyrop, 1976; Smith, 1970).
- The sharp drop in tax revenues in the 1980s after the previous increase in the late 1970s is genuine, per IMF historical data.

10 Ethiopia

Source

Source	First year	Last year	Interpolation
Historical archive	1965	1992 (2020*)	1989
ICTD	1993	2019	2005

Time series We use historical archive data from 1965 through 1992, then turn to ICTD (based on IMF Article IV) for 1993 to 2019 and we rely again on historical archive (based on Ethiopian Central Bank's Annual report) for the last year 2020. We interpolate years 1989 and 2005.

Link to historical archive data: click here.

Harmonization

The main adjustments for Ethiopia are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available prior to 1975 and between 2008 and 2019. Extrapolation of proportional ratios back from 1975 and interpolation between 2008 and 2019 using 2007 and 2020 as a reference, and corroborate based on additional policy documents. • Small unallocable income tax revenue ($\approx$ 1% of total revenue) after 2001 • Capital share of PIT: mean = 30%; constant
Social Contributions	RPC database for the whole period, extrapolated after 2014. Minuscule but non-zero values. According to SSA (2017), there has been a law in place since 1963. Thus, the little data suggests we are not missing a major component of government's tax collection
Property Taxation	Available for both historical archive and ICTD. Per IBFD, there is no wealth tax in Ethiopia, nor a real estate tax at the federal level. The fact that we observe property tax revenue (the land use fee) seems to indicate that we are observing state and local taxes
Decentralized Revenues	Evidence supporting the inclusion of state and local taxes (see Property taxation, above)

Notes

- We show a very good match on overall levels for the whole period compared with Mascagni (2016).
- 1974 marked a coup, after the world oil crisis, and beginning a socialist era which lasted through 1991 (not considered ex-communist, see note in Preface, iii). We do not observe any marked difference in our data in 1974, but the 1991-92 transition corresponds to a considerable decline in tax revenue from businesses (corporate income tax), although it soon recovered

11 France

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to 2020.

Harmonization

Data from OECD on France present detailed information on all the tax categories shown below. Notwithstanding, the considerably high values of the capital ETR are currently under revision.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 13%; stable trend (14% to 12%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

12 Germany

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to present. The data prior to 1990 is only for West Germany.

Harmonization

Data from OECD on Germany present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 12%; stable trend
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Local taxes are present from 1973, representing around 9% of total tax revenue in every year. In most cases more than 70% of these are from local income taxes, with the remainder from property and several types of sales taxes

Notes

13 India

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2020	

Time series Our primary source for Indian government revenue statistics is historical archive data. We collect data from two official data sources: prior to 2018 we rely on Indiastat (website link) and for 2019-2020 on the Reserve Bank of India (website link).

Link to historical archive data: click here.

Harmonization

The main adjustments for India are on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available after 1991. – Two main income taxes listed in its public finance accounting: the 'corporation tax' and the 'income tax other than corporation tax.' The latter is not just income taxes on individuals. However, prior to 1991, public financial records were disaggregated within this category. For years after 1991, we extrapolate the proportion of CIT/PIT in the revenue collected from 'income tax other than corporation tax' observed in years prior to 1991, and corroborate with additional information from policy documents. – For 2019 and 2020 we only have information on CIT/PIT split for central revenues. We compute the "*ratio of ratios*" for 2018: the ratio of the CIT/PIT split for central revenues over the CIT/PIT split for total revenues (central + state and local). Assuming a parallel trend from 2018 to 2020, we extrapolate this"*ratio of ratios*" for years 2019 and 2020 • Capital share of PIT: mean = 28%; ↓ trend (30% to 22%)
Social Contributions	The data sources used do not observe data on social contributions for India. However, according to the SSA (2018), India has had a social security policy in place since 1952, with a wide-ranging set of contributions to public social insurance scheme. Nevertheless, this revenue category is unobserved in official government as well as UN data, and listed as zero in ICTD-IMF data until a very small amount from 1998 onward (less than 0.05% of GDP). We suspect it is likely included within the income tax category
Property Taxation	Included but very small prior to 2018. Not included in 2019-2020
Decentralized Revenues	Our dataset includes comprehensively all sources of tax and non-tax revenue from central and state governments prior to 2019. See Income Taxation (above) for the adjustment for 2019-2020

Notes

14 Indonesia

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1996	1968-1971, 1994
OECD	1997	2020	

Time series Our main data source is historical archive data, through 1996, before OECD data begins in 1997. We interpolate years 1968-71 and 1994.
Link to historical archive data: click here.

Harmonization

The main adjustments for Indonesia are on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available from 1985 to 2001 (all unallocable taxes). Interpolation of proportional ratios using information from years prior to 1985 and later to 2001, and corroborated with additional information in policy documents • Capital share of PIT: mean = 29%; ↓ trend (30% to 24%)
Social Contributions	We do not observe social contributions in Indonesia (nor did Amir, Asafu-Adjaye, and Ducpham (2013)). According to SSA (2017), the first social security law was implemented in 1977. However, this was a very small proportion of government revenue, as the OECD data does not begin to include any social security revenue until 2015, and it is less than 0.1% of NDP
Property Taxation	Property taxation is minuscule according to the reported data
Decentralized Revenues	Without explicit mention on it, our data source seems to be federal. Either local taxes were unimportant or they were not accounted for. However, OECD data does not include them until 2000, and they do not make up more than 1% of the total revenue until 2013

Notes

- It is worth highlighting how natural resource revenues have been important and significant in Indonesia. Gillis (1985) estimated that oil and gas tax revenues rose from 1.2% in 1968 to 16.0% of GDP by 1981, while Ribeiro, Villafuerte, Baunsgaard, and Richmond (2012) estimate current (public) resource revenue at 4.5% of GDP. Resource revenue explains most of the differences observed between our estimations and prior literature (such as Gillis (1985) and Prasetyo (2018)). For our purpose, we attempt to stitch together a series on the strictly tax revenues with Indonesia's public finance (agreeing also with estimations from Amir et al. (2013)).
- The VAT was introduced in 1984 (replacing sales tax and turnover tax). After that, recent reforms have largely been aimed at increasing revenue and compliance as Indonesia was seen as a low-tax-effort country.

15 Italy

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to present.

Harmonization

Data from OECD on Italy present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Unallocable income tax in some years from 1980 to 2020 • Capital share of PIT: mean = 17%; ↓ trend (27% to 12%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Local government revenue is present in the data since 1973, but represents an increasing (and more than 3%) share of total tax revenue only since the mid-1990s, and has grown to 15% of overall tax revenue. The increase is largely due to the presence of local sales taxes, although there is also a miscellaneous local business tax

Notes

16 Japan

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to present.

Harmonization

Data from OECD on Japan present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 11%; ↑ trend pre-2000 (11% to 14%) and ↓ trend post-2000 (14% to 9%)
Social Contributions	Included
Property Taxation	included
Decentralized Revenues	Local government revenue is present in the data since 1973, and represents a steady 23-28% of total tax revenue. Income taxes represent a decreasing share of local tax revenue, but still more than half (down from 60% from 1973-92). Property taxes have gained in importance what income taxes have lost, up from 20 to 30% of total local tax revenue since the 1970s

Notes

17 Kenya

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2000	
OECD	2001	2020	

Time series We use historical archive data for the years 1965-2000, and then OECD data for the period from 2001 to 2020.

Link to historical archive data: click here.

Harmonization

The main adjustments for Kenya are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split of CIT/PIT available pre-1974 reform (reviewed historical information on PAYE vs. employer-surtax regulations) and post-2002 (OECD). Interpolation of proportional ratios for years 1973-2001, and use additional information to verify interpolated shares. • Unallocable income tax after 2001. From 2001 to 2014 allocated totally to PIT (withholding tax), whereas from 2015 to 2020 allocated to CIT/PIT according to OECD sheets (part withholding tax to PIT and part capital gains to CIT) • Capital share of PIT: mean = 28%; ↓ trend (30% to 26%)
Social Contributions	Really small but non-zero during the whole period. We rely on external sources for the period prior to OECD data (2000): RPC data pre-1996 and UN for 1996-2000.
Property Taxation	Per IBFD there is not a capital tax, but there is a local government tax on land value. We do not observe these taxes in our archive material
Decentralized Revenues	We are only able to collect revenues of the central government, with no information at the sub-national level

Notes

- Our estimation matches with prior literature, both in levels and trends (Karingi & Wanjala, 2005; Macha, Lado, & Nyansera, 2018). Little divergences, especially at income taxation for the earliest period (Maina, 2014) are likely explained by the value of the denominator (GDP).
- The structure of the kenyan tax system has been stable over time (Wawire, 1991) with some tax reforms expanding production taxes (sales and VAT), especially during the 1980s (Karingi & Wanjala, 2005).

18 Korea (Republic of)

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1971	1968-1971
OECD	1972	2020	

Time series We use historical archive data for the period 1965-71, then OECD from its debut in 1972. We interpolate from 1968 to 1971, as there were several missing elements in the available archival data for that period.

Link to historical archive data: click here.

Harmonization

For Korea, there are not major adjustments specifically earmarked for a particular tax category. The main adjustment comes from the interpolated period 1968-1971 (see Notes).

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 12%; ↓ trend pre-2000 (17% to 11%) and ↑ trend post-2000 (11% to 14%)
Social Contributions	Only present after 1985. According to SSA (2017), social security programs did not begin until 1973 and did not include a national pension (beyond a 'welfare' pension) until 1986 – a history matching our OECD revenue data
Property Taxation	Included
Decentralized Revenues	Decentralized revenue is not included in the archival data. However, the 1967 levels (final year of this data source) are not far off from the 1972 levels (first year of OECD). In addition, previous literature also show very low level of local revenue (less than 2% of revenue) (Kwack & Lee, 1992).

Notes

- The low levels of revenue in the 1960s are genuine (and not an artifact of, e.g., incomplete data), as the government's overall development strategy favored the inflow of foreign capital, implementing major tax reforms during this period (Yoo, 2000). Only after the 1967 tax reform when a 'global income tax' was introduced for the first time is when tax revenue increased in Korea (ibid.). This slight increase in personal and corporate income tax revenue matches with the interpolated period.

19 Mexico

Sources

Source	First year	Last year	Interpolation
Historical Archive	1965	1979	
OECD	1980	2020	

Time series We use OECD data from 1980, and refer to data from archives before 1980.

Link to historical archive data: click here.

Harmonization

The main adjustments for Mexico are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• No CIT/PIT split from 1980 to 2001. Interpolation of ratios using 1979 archive data and 2002 OECD data as references, and adjusted based on information in additional policy documents. • Capital share of PIT: mean = 15%, ↑ trend pre-1985 (23% to 35%), ↓ trend 1985-2013 (35% to 8%) and ↑ trend post-2013 (8% to 16%)
Social Contributions	Lack of social contributions from 1970 to 1979. Interpolation from 1969 until 1980 (when OECD beggins)
Property Taxation	Negligible value. There is a property tax in Mexico (*"impuesto predial"*) but accounts for only 0.02% of GDP. This estimation of property tax revenues matches Madrigal-Delgado (2021)
Decentralized Revenues	State and local tax revenues collected in OECD. Our historical archive data (1965-1979) matches the stitch year (1980), but we cannot corroborate the inclusion of all decentralized revenues before 1980

Notes

- There is a considerable rise in indirect tax revenues since 1980, the same year that we stitch across data sources. However, the jump is genuine and not an artifact of stitching. This increase is the result of the creation of the VAT in 1980, leading to higher indirect tax compliance (Burgess & Stern, 1993).

20 Nigeria

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1990	
ICTD	1991	2009	2008-2009
OECD	2010	2020	

Time series We use OECD data since 2010, and historical archive data starting from 1965. We use ICTD data for the years 1991 through 2009. We interpolate years 2008 and 2009.

Link to historical archive data: click here.

Harmonization

The main adjustments for Nigeria are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available 1987-1991. Interpolation of proportional ratios relying on 1986 and 1992 and overall income tax revenue observed for each year, and information in additional policy documents. • Unallocable income tax revenue between 1970 and 1986. Allocated to CIT based on historical sources. • Capital share of PIT: mean = 27%; ↓ pre-1980 (30% to 26%), ↑ 1980-1985 (26% to 30%) and ↓ post-1985 (30% to 24%)
Social Contributions	RPC database prior to 2010 and OECD database later to 2010. Minuscule but non-zero (RPC and OECD periods) and zero (ICTD period). According to SSA (2017), there has been a social security (pension) law in place since 1961, although the most recent version dates to 2014. It is possible that the levels are too low to appear in ICTD data and for the government to track in centralized budget and revenue statistics
Property Taxation	There is no tax on net worth in Nigeria, and we do not observe any property taxes in any of our data sources
Decentralized Revenues	The revenues in our archival data source are federal revenues, of the central government. Per IBFD (2019), some states charge a real estate tax, but we find no record of their measurement and magnitude. There are no income taxes at the subnational level

Notes

- Competing sources of data do not always precisely match on levels, but they do match on trends.
- Our estimates present differences from Ekpo and Ndebbio (1996). It might be due to the treatment of volatile *non-tax* revenue, and difficulty classifying public revenue streams.

- Oil revenues are the major story in Nigeria, especially as a source of capital (Ribeiro et al., 2012). Indeed, we observe the spike in capital income and revenue during the 1970s oil price shocks and its subsequent drop in the 1980s coincident with the collapse in oil prices (Pinto, 1987).

21 Pakistan

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2020	

Time series Among archival sources, we use central government sources for federal statistics, with particular attention to parallel publications for provincial statistics. Between 2015 and 2020 we collect the share of federal tax revenue that is collected at the provincial level. As this revenue is only available in aggregated terms, we assume that the share of each tax category is equal across the federal revenue and the share of it that is collected at the provincial level. We compliment that with "pure provincial" tax revenue.

Link to historical archive data: click here.

Harmonization

The main adjustments for Pakistan are on income taxation and decentralized revenues.

Category	Adjustment
Income Taxation	• Split CIT/PIT only available for 1965-1980 (from historical archive) and 1994-2004 (from ICTD-IMF): i) interpolation of CIT/PIT split ratios from 1980 to 1994 and ii) extrapolation of CIT/PIT split ratios from 2004 to present, both verified with additional policy documents. • Capital share of PIT: mean = 29%; ↓ trend (30% to 26%)
Social Contributions	Miniscule but non-zero levels. According to SSA (2018), the first law was passed in 1972, but first implemented in 1976, which corroborates our archival data that begins in 1973 (also shown in Syeda (2015))
Property Taxation	Taxes on assets are minuscule but non-zero , and larger in the early period (especially at provincial level). According to the IBFD, there is a tax (since 2013) on 'net movable wealth', a *zakat* wealth tax on Muslims and an *ushr* on agricultural land. The 10% property tax as well as a land tax on farm holdings greater than five irrigated acres (or ten unirrigated) are assessed at the province level
Decentralized Revenues	• Provincial tax revenue data for 1947-2014, benchmarking every four years and every time a significant type of tax comes into or out of the long-run data series. Interpolation of the intervening years • For 2015-2020 we follow the process explained in 'Time series'

Notes

- Hasan, Kemal, and Naseem (1997) is 0.2-0.3 percentage points higher than our estimations.

22 Phillipines

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1993	
OECD	1994	2020	

Time series We use historical archive data from 1965 through 1993, then turn to OECD data for the period since then.

Link to historical archive data: click here.

Harmonization

The major adjustments in Phillipines are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available from 1978 to 1998. Interpolation of proportional ratios using 1977 and 1999 as reference, and corroborated with additional historical sources. • Some unallocable income tax during the early 2000s • Capital share of PIT: mean = 25%; ↓ trend (29% to 20%)
Social Contributions	We use external data prior to 1994 (UN SNA) and rely on OECD for the period since then
Property Taxation	OECD data includes a central government wealth tax (on stock transactions) and a much more significant local government property tax estimate (0.5% of GDP). We do not have this estimate in our archival data
Decentralized Revenues	We do not have sub-national revenues for the period we use archival data (pre-1994). However, the smooth evolution across sources suggests that only small revenue comes from the state and local level (less than 5% of the total tax revenue for OECD period)

Notes

23 Poland

Sources

Source	First year	Last year	Interpolation
OECD	1994	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include Poland before 1994. We use OECD data from 1994 to 2020.

Harmonization

Data from OECD on Poland present detailed information after 1994 on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1994 • Capital share of PIT: mean = 11%; ↓ trend pre-2005 (12% to 10%) and ↑ trend post-2005 (10% to 12%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

24 Russia

Sources

Source	First year	Last year	Interpolation
Historical archive	1994	1999	1994-97
ICTD	2000	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we neither include Soviet Russia nor any Soviet states (in the era of the USSR) in our calculations. For the pre-2000 period, we use estimates from Ivanova, Keen, and Klemm (2005). For the post-2000 period we rely on ICTD data.

Link to historical archive data: click here.

Harmonization

The main adjustments for Russia are on all tax categories for the first years observed (1994-1998) and on property taxation for the whole sample.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1994 • Capital share of PIT: mean = 15%; constant trend
Social Contributions	Included
Property Taxation	Little information on property tax in Russia from our sources, as it is frequently lumped with 'other taxes' (case of ICTD). Preobragenskaya and McGee (2003) indicate that it does not exceed 2% of GDP and Owen and Robinson (2003) indicate that property tax was always between approximately 0.5 and 1.5% of GDP from 1994-99. For simplicity we set it equal to 1 percent, in the absence of better (and more recent) data
Decentralized Revenues	Dethier (2000) notes that the 1990s were a period of growing fiscal decentralization. Comparing our estimates to theirs supports the idea that we are including decentralized revenues in our estimates of Russia's public finance time series

Notes

- The IMF data for 1994-2003 matches well the ICTD data that begins in 2000.

25 South Africa

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	
OECD	1990	2020	

Time series We use historical archive data for 1965-89, then OECD data for the period since then.

Link to historical archive data: click here.

Harmonization

The main adjustments for South Africa are on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT before 1990 only available at five-year intervals (1966, 1971, 1976, 1981, 1986). For the rest of the years before 1990: interpolation of the within-ratios using the five-year intervals as reference. • Capital share of PIT: mean = 21%; ↓ trend (22% to 19%)
Social Contributions	We do not observe social contributions before 2000 even though in principle the social security policy was established in 1928 (SSA 2017). However, neither alternative sources (OECD, ICTD, IMF, historical archive) nor other primary sources (S. A. R. Bank, 2016) include any social security revenue prior to 2001
Property Taxation	Small (especially pre-2000) but non-zero during the whole period. Years 1967-1972 are interpolated
Decentralized Revenues	OECD observes zero state and local taxes until 2002. The pre-1990 archive data is drawn solely from (and regarding) central government finance records. We assume local and state taxes were also zero before 1990

Notes

26 Spain

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to present.

Harmonization

Data from OECD on Spain present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Negligible unallocable income tax prior to 1994 derived from local tax • Capital share of PIT: mean = 16%; ↓ trend (23% to 12%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Local government revenue (mainly sales and business tax) is present in the data since 1973, but represents an increasing (and more than 3%) share of total tax revenue only since the mid-1990s, and has grown to 15% of overall tax revenue

Notes

- The sharp decline in corporate income tax revenues coincides with the beginning of the Great Recession.

27 Sudan

Sources

Source	First year	Last year	Interpolation
Historical archive	1972	1980	
ICTD	1981	2020	

Time series The data series for Sudan starts in 1972. We use data from the IMF historical source from 1972 to 1980, then ICTD data from 1981.

Link to historical archive data: click here.

Harmonization

The major adjustments for Sudan are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Total income taxation revenue not available from 1981 to 1990. Interpolation of the share of total income taxation over total tax revenue using 1980 and 1991 as reference, and confirmed in additional sources. • Split CIT/PIT not available before 1994 and after 2006 with some missing years between 1994 and 2006. For 1980-1991 interpolation of proportional ratios using 1980 and 1991 as reference, and corroborated with historical sources. For missing years in ICTD period pre-2006, interpolation of CIT-to-total-PIT ratios. For period pre-1994 and post-2006, extrapolation of constant CIT-to-total-PIT ratios. • All unallocable income tax revenue is allocated to PIT, based on additional policy documents. • Capital share of PIT: mean = 29%; ↓ trend pre-2005 (30% to 28%) and ↑ trend post-2005 (28% to 30%)
Social Contributions	Minuscule but non-zero during the whole period. We rely on external sources (RPC data) Assumed as constant after 2014
Property Taxation	Only available in historical archive data (1970s), but negligible in quantity
Decentralized Revenues	Not determined

Notes

28 Tanzania

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2020	1972, 1977, 1993-95

Time series We use historical archive data for a unified series from 1965 to 2020, interpolating several years and stitching across two data sources, where the second data source came into effect in 1996.

Link to historical archive data: click here.

Harmonization

The main adjustments for Tanzania are on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available before 1974. Extrapolation back from 1974 to 1965 of the proportional ratios of 1974 with the total income revenue collected for each year, verified with historical sources. • Capital share of PIT: mean = 30%; constant pre-2010 (30%) and ↓ trend post-2010 (30% to 28%)
Social Contributions	Minuscule but non-zero values over the whole period. According to SSA, there has been a law in place since 1964, while a plethora of regulations govern the current funds from social contributions (IBFD)
Property Taxation	Minuscule but non-zero value before 1990, and missing after 1990. According to IBFD, there is no wealth tax, but there is a real estate tax since 1974. However, alternative sources (ICTD, IMF) show how this tax does not form a large part of ICTD or IMF estimates
Decentralized Revenues	Not determined

Notes

- Our estimations are significantly below prior literature for Tanzania (Fjeldstad, 1995; Nord et al., 2009; Osoro, 1993). However, we find correspondence in the *within*-weight of types of taxes, as a proportion of total tax revenue. A potential explanation points to differences in GDP's values (check for instance World Bank (2019)). We suspect that scholars were using the same public revenue numbers as us but a smaller GDP denominator. As a result, with same revenue sources they would have found a higher tax-to-GDP ratio.
- The 1990s revenue trend appears genuine (not an artifact of our data construction), after comparing notes across these sources. It corresponds with a period of notable reforms: a notable civil service reform in the early 1990s, the institution of the Tanzania Revenue Authority (TRA) in the mid-1990s, and other reforms in the late 1990s (including imposition of the VAT in 1998).

29 Thailand

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1999	1972, 980, 1985, 1987, 1994, 1998
OECD	2000	2020	

Time series We use historical archive data for the period through 1999, then OECD data from 2000 to 2020. Our archival data for Thailand is less exhaustive than for other countries and features more interpolation than in other countries.

Link to historical archive data: click here.

Harmonization

The major adjustments in Thailand are mainly due to the large interpolation during the whole period. In addition, some modifications are made on social contributions and property tax.

Category	Adjustment
Income Taxation	• Split CIT/PIT available for the years for which we have data • Very small unallocable income tax during the 1980s • Capital share of PIT: mean = 22%; ↓ trend (30% to 15%)
Social Contributions	Low but different from zero after 1980. We use external data prior to 2000 (UN SNA database, interpolating 1997-1999) and OECD for the period since then
Property Taxation	Low but different from zero during the whole period. For period prior to 2000 we rely on IMF data. There is no wealth tax but there is a local property tax. We capture the latter in our OECD data but it is not available in our archival data source
Decentralized Revenues	It is likely that the OECD data contains local government tax revenues, while our archival data does not. However, the smooth evolution across sources suggests a sub-national level of revenue that is small by orders of magnitude

Notes

- Our estimation matches previous literature both in levels and trends (Bernardi, Fumagalli, & Gandullia, 2005; Jansen & Khannabha, 2012).

30 Turkey

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to 2020.

Harmonization

Data from OECD on Turkey present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Unallocable income tax in the 1970s • Capital share of PIT: mean = 16%; ↓ trend (22% to 11%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Local government data is present since 1980, and represents usually between near 9 or 10% of total tax revenue per OECD, spiking at 16% in 1998

Notes

31 Uganda

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1991	1984, 1991
OECD	1992	2020	

Time series We use archive data for the historical period from 1965, then OECD data from 1992. Years 1984 and 1991 are interpolated.

Link to historical archive data: click here.

Harmonization

The major adjustments for Uganda are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available prior to 1990 for the historical archive data. We use IMF historical data for 1972-1990. We interpolate proportional ratios prior to 1972 and for years 1987 and 1989 • Very small unallocable income tax revenue since 1998 • Capital share of PIT: mean = 27%; ↑ trend pre-1990 (27% to 30%) and ↓ trend post-1990 (30% to 23%)
Social Contributions	We draw on RPC data for social contributions. The value is relatively small, and does not appear in OECD data. We extrapolate for the period from 2014 to present as RPC data for social contributions is missing for these years
Property Taxation	Not included. There are no net wealth/worth taxes in Uganda. Property taxes are administered by the local authorities annually. They are based on the value of the property as assessed by the local authorities
Decentralized Revenues	Not determined. Possibly not included (at least the part derived from property taxes)

Notes

- The match between archive data sources and OECD data is very close.

32 United Kingdom

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to 2020.

Harmonization

Data from OECD on United Kingdom present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 14%; ↓ trend pre-1990 (16% to 11%) and ↑ trend post-1990 (11% to 14%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

33 United States of America

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to 2020.

Harmonization

Data from OECD on United States present detailed information on all the tax categories shown below. There are no major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 16%; ↓ trend (22% to 12%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

34 Vietnam

Sources

Source	First year	Last year	Interpolation
IMF	1994	2002	
OECD	2003	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include Vietnam before 1994. We refer to IMF Time series until 2002 and to OECD revenues series from 2003 to 2020.

Link to historical archive data: click here.

Harmonization

The major adjustments in Vietnam are on social contributions. Social contributions are not included here even though we are aware that they represent a considerable share (at least after 2010).

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1994 • Capital share of PIT: mean = 26%; ↓ trend (30% to 19%)
Social Contributions	We do not have data on social contributions until 2010 (the IMF does not include it and OECD only available after 2010). However, according to SSA (2018), Vietnam has a longstanding social security policy (1961, pre-dating the war). For 2010-2020, social contributions accounts for a very significant part of the total tax revenue (4-6% of GDP). As we do not have data prior to 2010, we decide to temporarily exclude social contributions here. This decision significantly understates the estimations on revenue/ETR on labor income, so we urge the reader to keep that in mind when interpreting the graphs
Property Taxation	Very small (especially in OECD data) but non-zero during the whole period
Decentralized Revenues	Included

Notes

- The drop in indirect taxation in 2020 is genuine, as it is confirmed by the source, the OECD revenue statistics.

35 Australia

Sources

Source	First year	Last year	Interpolation
OECD	1965	1998	

Time series We use OECD data from 1965 to 2020.

Harmonization

Data from OECD on Australia present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 13%; ↓ trend pre-1985 (15% to 8%) and ↑ trend post-1985 (8% to 15%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

36 Cameroon

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1992	1969-1970, 1989
OECD	1993	2020	

Time series We refer to historical archive data for the initial period 1965-1992, and we use OECD for the period since then. We highlight that for the historical archive period, use information from Amin (1998) for 1969-1992 to help assign taxes to tax-types in specific years.

Link to historical archive data: click here.

Harmonization

There are adjustments on Cameroon for income taxation, social contributions and property taxes.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for period 1969-1992. Interpolation of proportional ratios using 1968 and 1993 as reference, and corroborated based on historical sources. • Small unallocable income tax revenue during the OECD period (1993-2020) • Capital share of PIT: mean = 29%; fluctuating between 27% and 30%
Social Contributions	Use of external source (RPC data) from 1965 to 1993. Per SSA (2017) social security policy dates to 1969, while RPC has data from 1960. However, it is a small level (0.5% of GDP) and only rises above 1% of GDP in 1969. We adjust for the period 1989-1993 by interpolating levels instead of relying on RPC (which is lower than OECD during the period of their overlap)
Property Taxation	Available only intermittently for the period 1969-1992, though limited decentralized revenues suggest it may not be a major source of tax collection (but worthy of further investigation).
Decentralized Revenues	Not determined in historical archive period (1965-1992), not included in OECD period (post-1993)

Notes

- The 'taxe unique' is a type of indirect tax, replacing others, as a preferential tax status (Gauthier, Soloaga, & Tybout, 2002), even though it is not the only source of revenue for Cameroon (Gauthier & Gersovitz, 1997). Trade taxes accounted for up to 50% of total tax revenue in the mid-1960s, but it fell to 10% by the mid-1970s and down to less than 5% by the 1990s (shown in Amin (1998)).

37 Canada

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We use OECD data from 1965 to 2020.

Harmonization

Data from OECD on Canada present detailed information on all the tax categories shown below. Notwithstanding, the extremely high values of the capital ETR are currently under revision. See Methodology Note in the Atlas website for further details on this issue.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 13%; ↓ trend pre-2008 (14% to 11%) and ↑ trend post-2008 (11% to 13%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

38 Cote d'Ivoire

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	1977-1979, 1987-1989
OECD	1990	2020	

Time series We refer to historical archive data before 1990 and we use OECD for the period since then. Years 1977-1979 and 1987-1989 are interpolated.

Link to historical archive data: click here.

Harmonization

The main adjustments for Cote d'Ivoire are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available before 1982. Interpolation of proportional ratios based on 1963 and 1986-1989 (relying for that on ICTD data), and corroborated in additional policy sources. In addition, split CIT/PIT from OECD has been corrected: allocation to PIT of the consistent 'wedge' between the social contributions series according to OECD vs according to ICTD • Unallocable income tax revenue in historical archive data (all is tax on 'bénéfices', allocated as explained above), very small in magnitude, exists also in OECD data • Capital share of PIT: mean = 13%; ↓ trend pre-2008 (14% to 11%) and ↑ trend post-2008 (11% to 13%)
Social Contributions	Rely on ICTD values for 1980-2010. For 2010-2020, OECD values divided by the ratio OECD/ICTD from 2010. For missing years prior to 1980, interpolation based on available observations
Property Taxation	Included
Decentralized Revenues	Not determined

Notes

39 Ghana

Sources

Source	First year	Last year	Interpolation
Historical archive	1967	1999	
OECD	2000	2020	

Time series We refer to historical archive data for the period 1967-1999 and we use OECD data for the period since then.

Link to historical archive data: click here.

Harmonization

The main adjustments for Ghana are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available prior to 1972. Backward extrapolation from 1972 to 1965 using 1972 as reference • Small but non-zero unallocable income tax revenue after 1989. Prior to 2000, historical archive has very small revenue from 'tax on interest and dividends' that is not included as such in OECD data. Due to its small magnitude, we do not implement any further adjustment here. • Capital share of PIT: mean = 28%; ↑ trend pre-1980s (28% to 30%) and ↓ trend post-1990 (30% to 24%)
Social Contributions	Rely in external sources (RPC data) for the historical archive period prior to 2000
Property Taxation	Local tax, very small but non-zero in some years of historical archive. Not available in OECD data
Decentralized Revenues	Not fully determined before 2000 (only available a local property tax for some years). Not included after 2000 (OECD data)

Notes

- Our estimations match prior literature (Osei & Telli, 2017). In addition, the same literature corroborates the strange patterns of the 1980s including the two consecutive periods of fiscal management under IMF recovery and structural adjustments programs during the middle and end of the decade, into the 1990s.

40 Madagascar

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	1969-1971, 1974-1976, 1981-1983
ICTD	1990	1998	
OECD	1999	2020	

Time series We refer to historical archive data for years prior to 1990, move to ICTD for the 1990s and use OECD for the period since 1999. Our historical archive database for Madagascar is less exhaustive than for other countries. Indeed, we interpolate a significant number of years.

Link to historical archive data: click here.

Harmonization

The main adjustments for Madagascar are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT available for the years for which we have data (not for interpolated years) • Unallocable income tax for 1965-1968 included in PIT, based on additional policy documents • Capital share of PIT: mean = 30%; stable pre-1975 (29%-30%) and constant post-1975 (30%)
Social Contributions	Rely on external sources (RPC data) for the whole period. Extrapolation of constant values for years post-2014
Property Taxation	Very small but non-zero during the whole period
Decentralized Revenues	Not determined before 1999 (ICTD data), and not included after 1999 (OECD data)

Notes

41 Malaysia

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	1979-1980, 1988
OECD	1990	2020	

Time series We refer to historical archive from 1965 to 1989 and we use OECD for the period since then. We interpolate years 1979, 1980, and 1988.

Link to historical archive data: click here.

Harmonization

The main adjustments for Malaysia are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available before 1978. We use IMF historical data to assign to PIT and CIT. Extrapolated backwards for years prior to 1975 using 1976 as a reference, and corroborated with IMF information. • Small but non-zero unallocable income tax revenue since 1988 • Capital share of PIT: mean = 20%; ↓ trend (28% to 15%)
Social Contributions	Rely on external data (RPC data) before 1971. Extrapolate values for 1972-1999 using 1971 (from RPC data) and 2000 (from first year of OECD that includes social contributions) as reference
Property Taxation	Not included in historical archive (pre-1990) and included but barely zero in OECD data (post-1990)
Decentralized Revenues	Not determined

Notes

- There is a gap in the stitching year 1990, between historical archive and OECD, of 0.7 percentage points of GDP for specific tax types. This shortfall is an artifact of what is included in 'indirect taxes' and in 'other taxes'. It seems like historical archive is including stamp duties in unallocable income taxes whereas OECD allocates them to other taxes. On the other hand, the OECD includes a motor vehicle tax not collected in historical archive.

- The late 1980s drop in revenue is corroborated in ICTD data (referring to IMF Article IV).

42 Morocco

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1999	1966 1971-1973 1995
OECD	2000	2020	

Time series We use historical archive data from 1965 to 1999 and refer to OECD for the period since then. We interpolate years 1966, 1971-1973 and 1995.

Link to historical archive data: click here.

Harmonization

The major adjustments for Morocco are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available from 1988 to 1996. Interpolation of proportional ratios using 1987 and 1997 as reference, and using information from historical documents • Unallocable income tax revenue before 1999. Allocated completely to PIT, based on additional policy documents. • Capital share of PIT: mean = 28%; constant pre-1986 (30%) and ↓ trend post-1986 (30% to 22%)
Social Contributions	External data prior to 2000 (RPC data)
Property Taxation	We observe wealth tax data at a federal level all the way back to 1965. However, we include local property tax data only after 2000 (when OECD becomes available)
Decentralized Revenues	We observe federal and local taxes in the OECD period but we do not have data on sub-national taxes in our archival data

Notes

- The lack of sub-national revenue prior to 2000 could explain the changes observed in 1999-2000: i) the bump in indirect taxes (likely revenue that was included in non-tax revenue in our archival data) and ii) the increase in property tax revenue (local property tax included after 1999).
- The spike in CIT in 1975 as well as the dip in indirect taxes in 1986 are genuine.

43 Mozambique

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2014	1991, 2001
ICTD	2015	2020	

Time series We start the series in 1975 (first post-independence year with consistent data). We refer to historical archive from 1965 to 2014 and we use ICTD for the period since then. We interpolate years 1991 and 2001.

Link to historical archive data: click here.

Harmonization

The main adjustments for Mozambique are on direct taxation and social contributions. Additionally, the extremely high values of the capital ETR are currently under revision. See Methodology Note in the Atlas website for further details on this issue.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available between 1994 and 2014. We use Castro, Junquera-Varela, Schenone, and Teixeira (2009) for the period 1994-2007 and extrapolate proportional ratios from 2007 to 2014 • Capital share of PIT: mean = 30%; constant trend pre-2010 (30%) and ↓ trend post-2010 (30% to 28%)
Social Contributions	Rely on external sources (RPC database), constant value for the whole period (0.001%). Extrapolation of that constant value for period post-2014
Property Taxation	Not included in historical archive and close to zero in ICTD data
Decentralized Revenues	Not determined

Notes

- 1980s dip in revenue is genuine, per ICTD corroboration.
- The big increase in revenue after 2010 is supported by the overlapping period between historical archive and ICTD (2014-2015).

44 Nepal

Sources

Source	First year	Last year	Interpolation
Historical archive	1976	2005	
ICTD	2006	2020	2018-2020

Time series We start the series in 1976. We refer to historical archive from 1976 to 2005 and we use ICTD for the period since then (year 2005 relies on ICTD for indirect taxation). We interpolate years 2018, 2019 and 2020.

Link to historical archive data: click here.

Harmonization

The main adjustments for Nepal are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for pre-1987, 2006-2009 and post-2017. For pre-1987: extrapolation back using 1987 as reference, and corroborated using historical sources. For 2006-2009, interpolation using 2005 and 2010 as references. For post-2017, extrapolation using 2017 as reference • Capital share of PIT: mean = 29%; constant trend pre-2005 (30%) and ↓ trend post-2005 (30% to 26%)
Social Contributions	Rely on external sources (RPC data) for the historical archive period, pre-2006 (and years 2010 and 2011)
Property Taxation	It evolves from small (0.2% of GDP) in historical archive (2005) to zero in ICTD (post-2005). We highlight this issue, although it is not significant for our purpose due to the limited size of property taxes collected
Decentralized Revenues	Not determined

Notes

45 Niger

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1999	1969-1972
OECD	2000	2020	

Time series We refer to historical archive from 1965 to 1999 and we use OECD for the period since then. We interpolate the period 1969-1972.

Link to historical archive data: click here.

Harmonization

The main adjustments for Niger are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available before 1975 and in 1999. Extrapolation back from 1975 to 1965, confirmed in additional documents, and interpolation of year 1999 using OECD period • Unallocable income tax revenue allocated to PIT for the period 1975-1998 ('sur un rôle' and 'général' are assigned to PIT instead of remaining unallocable, based on policy documents) • Capital share of PIT: mean = 29%; constant trend pre-2005 (30%) and ↓ trend post-2005 (30% to 26%)
Social Contributions	Rely on external source (RPC data) through 1998, interpolation of 1999 (outlier value). According to SSA (2017) the first social security law was in 1967, so we do not include data from RPC before 1967 (RPC estimates it at 0.2% of GDP)
Property Taxation	Very small but non-zero during the whole period (with some discrepancy between historical archive and OECD estimates of property tax in 1999)
Decentralized Revenues	Not determined for historical archive data (1965-1999), not included for the OECD period (2000-2020)

Notes

46 Peru

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1989	1965-1967
OECD	1990	2020	

Time series We refer to historical archive from 1965 to 1989 and we use OECD for the period since then. Values for years 1965 and 1967 are extrapolated backwards. We calculate the ratios of each tax category's (from historical archive data) share of total revenue (from RPC data) in 1968, and extrapolate them back to 1965 using aggregated values from RPC data for 1965 to 1967.

Link to historical archive data: click here.

Harmonization

The main adjustments for Peru are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available from 1983 to 1989. Interpolation using 1982 and 1990 as reference • Unallocable income tax small but non-zero during the whole period. Allocated all of it to PIT for the interpolated period (1983-1989), based on historical sources • Capital share of PIT: mean = 27%; ↓ trend (30% to 22%)
Social Contributions	External data (RPC data) for the historical archive period (pre-1990)
Property Taxation	OECD accounts for three types of wealth taxes (individual, corporate and financial and capital transactions), whereas historical archive data only accounts for one heading 'impuesto al patrimonio'. There is a gap between historical archive and OECD that might be explained by the fact that some of the latter taxes collected in the OECD were local-level
Decentralized Revenues	OECD includes information on sub-national level taxes (explicitely after 2000, but the smooth evolution suggest they are included in total revenue in 1990-2000). On the contrary, the evidence from property taxes (above) indicates a lack of information regarding decentralized revenues in historical archive data. However, the matching for the rest of the categories in the overlapping period (1990-1994) suggests that other sub-national-level taxes are not significant

Notes

- There is a small jump in 1990 from the evolution of property taxes explained above. However, we are confident that the rest of the evolution is genuine, as the rest of tax categories match perfectly between historical archive data and OECD data for the overlapping period (1990-1994)

47 Saudi Arabia

Sources

Source	First year	Last year	Interpolation
ICTD	1994	2020	2006-2008

Time series We start the data series for Saudi Arabia in 1994. We refer to ICTD for the whole period. We interpolate from year 2006 to 2008.

Harmonization

The main adjustments for Saudi Arabia are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Assume equal split for unallocated income tax revenues between PIT and CIT, hard to corroborate due to limited historical sources, but note that small in magnitude • Capital share of PIT: mean = 15%; constant
Social Contributions	Minuscule but non-zero during the whole period. We rely on external data (RPC data). Extrapolation after 2014
Property Taxation	Not included during the whole period
Decentralized Revenues	Not determined

Notes

- Direct non-oil tax revenue is almost inexistent. However, we highlight that non-tax revenue is always more than 25% and sometimes as much as 50% of GDP.

48 Ukraine

Sources

Source	First year	Last year	Interpolation
ICTD	1994	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include Ukraine in our calculations prior to 1994. We refer to ICTD for the period since then.

Harmonization

Data post-1994 from ICTD on Ukraine present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1994 • Very small unallocable income tax in the 2000s • Capital share of PIT: mean = 25%; ↑ trend pre-2000 (25% to 29%) and ↓ trend post-2000 (29% to 23%)
Social Contributions	Included
Property Taxation	Very small but non-zero during the whole period
Decentralized Revenues	Not determined

Notes

49 Uzbekistan

Sources

Source	First year	Last year	Interpolation
ICTD	1994	2020	2013-2014

Time series In keeping with our rule to not include communist countries prior to their transition, the data series for Uzbekistan starts 1994. We refer to ICTD for the whole period. Years 2013 and 2014 are interpolated

Harmonization

The main adjustments for Uzbekistan are on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available since 2003. Extrapolation of proportional ratios using 2002 as reference, and corroborated with recent policy documents • Unallocable income tax revenue (especially relevant since 2003), all allocated to PIT to based on policy documents • Capital share of PIT: mean = 27%; ↓ trend (30%-22%)
Social Contributions	Small values from ICTD in the early 1990s. Not available from 1995-1998: interpolation using 1994 and 1999 as reference. Other sources (IMF, WB) exclude it from the definition of tax revenue for Uzbekistan
Property Taxation	Included
Decentralized Revenues	Included

Notes

- Our estimation gives very close values to prior literature on Uzbekistan. This is the case for the whole period analysed: for the end of the 1990s (slightly below Davoodi and Grigorian (2007) and der Hoek (2008)), for mid 2010s (Mokhtari & Ashtari, 2012) and for recent years (check World Bank (2020) or IMF (2019)).

50 Venezuela

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1979	1976-1978
ICTD	1980	1989	
OECD	1990	2017	

Time series We refer to historical archive from 1965 to 1979. We rely on ICTD from the subsequent decade (1980-1989) and use OECD for the period since 1990. We interpolate years 1976, 1977 and 1978 (we have overall tax revenue for these years). Data on government revenue is not available beyond 2017.

Link to historical archive data: click here.

Harmonization

The main adjustments for Venezuela are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available prior to 1971, extrapolate the share of CIT of total income tax revenue back from 1971 to 1965 and verify using historical sources • Unallocable income tax revenue category, use information from (Barreix, Benítez, & Pecho, 2017; McLure Jr, 1992; Zolt & Bird, 2005), to assign all of it to PIT • Capital share of PIT: mean = 23%; no clear trend (fluctuating between 22% and 26%)
Social Contributions	Rely on external data prior to OECD first year of observation (RPC database). No values post-2012: assumed 2013-2017 values equal to the level of social contributions in 2012
Property Taxation	We do not collect any wealth or property tax prior to OECD data
Decentralized Revenues	The OECD, does include any local taxes

Notes

- There is a jump in stitching year 1980 (historical archive to ICTD). This gap is within the normal range of Venezuela year-to-year patterns, and consistent with the pattern from 1980-81.
- In the early period, pre-1970, a significant proportion of revenue was coming from non-tax sources, nationalized mining and extractive industry. While some revenue from extractive industry remained public and non-tax in the later eras, perhaps the strictly *tax* study in Venezuela is not an entirely revealing picture of their public finances. This may be the case in many petroleum-exporting states.

51 Yemen

Sources

Source	First year	Last year	Interpolation
Historical archive	1990	1997	
OECD	1998	2012	

Time series Our first year observation is 1990 (year of unification and creation of the modern Republic of Yemen), and stops in 2012 (since the Yemeni Crisis began). We refer to historical archive data for the period 1990-1997 and we use ICTD data for the period 1998-2012.

Link to historical archive data: click here.

Harmonization

The main adjustments for Yemen are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not reliable before 1998 and not available after 2003 (and in year 2001). Pre-1998: extrapolation backward of proportional ratios using 1998 as reference, confirmed with historical source. Post-2003: extrapolation forward of proportional ratios using 2002 as reference, and corroborated with recent policy documents • Capital share of PIT: mean = 30%; constant
Social Contributions	Rely on external sources (RPC database). It is constant at a very small magnitude (0.1% of GDP)
Property Taxation	ICTD includes a extremely small amount of property tax during the years of its highest precision (1998-2002), but we have this category missing before and after
Decentralized Revenues	Not determined

Notes

- It is worth keeping in mind that non-tax revenue (presumably always oil/resource revenues) is always important in Yemen—from 8% of GDP in 1990 (historical archive) to a high of 30% (ICTD) by 2008.

52 Taiwan

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2020	

Time series We refer to historical archive based on the reports by the Ministry of Finance for the whole period from 1965 to 2020.

Link to historical archive data: click here.

Harmonization

Historical archive data on Taiwan presents detailed information on all the tax categories shown below. There are no major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 16%; ↓ trend pre-1980 (23 to 15%) and constant post-1980 (15%)
Social Contributions	Not included • The Ministry of Finance website lists its tax collections by OECD classification, and leaves social contribution series blank. However, this policy source includes social contribution revenue from 2002 (and we could calculate it until 2000) • Per 2004 Statistical Data Yearbook [see chapter 16] it would seem that there are several decentralized social insurance mechanisms, perhaps enforced and cataloged if not collected at a national level. Unfortunately, there seems to be no centralized catalog of historical statistics on the size of these contributions, nor a clear demarcation between the public and private parts of these contributions. • If the social security system is organized privately in both contributions and expenditures, then it could be rightly excluded from revenues. However, this would blur the line between what is a tax and what is a required (private) expenditure
Property Taxation	Included
Decentralized Revenues	Determined

Notes

53 Mali

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1979	
ICTD	1980	1999	
OECD	2000	2020	

Time series We refer to historical archive from 1965 until 1979, we use ICTD for the two subsequent decades (1980-1999) and rely on OECD for the period since 2000.

Link to historical archive data: click here.

Harmonization

The main adjustments for Mali are made on income taxation, social contributions and property taxes.

Category	Adjustment
Income Taxation	• Split CIT/PIT not reliable before 1980. Extrapolation back of proportional ratios from 1980 to 1965, verified using additional historical sources • Split CIT/PIT not available in years 1982-1983 and 2000-2001. Interpolation of proportional ratios using years for which we have data (before and after missing observations) • Small but non-zero unallocable income tax revenue during the whole period • Capital share of PIT: mean = 30%; constant
Social Contributions	Rely on external sources (RPC database) before 2000 (first year of OECD)
Property Taxation	Small but non-zero during the whole period. Interpolation of years 1975-1983
Decentralized Revenues	Not determined before 1999, not included for OECD data

Notes

- There are limited historical documents on Mali's historical public finance. Founou-Tchuigoua (1989) who discussed a purported long-run public finance crisis, did not attempt to measure tax/GDP. However, the disaggregated presentation of indirect tax revenue as more than twice the amount of direct tax revenue matches well our own findings.

54 Burkina Faso

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1999	1974, 1980-1983, 1987
OECD	2000	2020	

Time Series We refer to historical archive from 1965 to 1999 and we use OECD for the period since then. Several years interpolated, coincide with periods of political unrest and violence. For the period 1993-99, we use ICTD for overall tax/GDP level[5] and interpolate compositional 'within-tax' ratios from 1992 (historical archive) to 2000 (OECD).

Link to historical archive data: click here.

Harmonization

The main adjustments for Burkina Faso are made on social contributions and property taxes.

Category	Adjustment
Income Taxation	• Split CIT/PIT available for the non-interpolated years • Small but non-zero unallocable income tax during the OECD period (post-2000) • Capital share of PIT: mean = 30%; constant
Social Contributions	Rely on external source (RPC database) for the period pre-OECD (prior to 2000)
Property Taxation	Small but non-zero property taxes during the whole period. Period 1982-1999 is fully interpolated
Decentralized Revenues	Not determined for historical archive period, and does not include tax revenues collected by local authorities but includes revenues collected by central government on behalf of local authorities (in all periods)

Notes

[5] We also interpolate forward the comparison ratio of historical archive tax level to ICTD tax level in 1992.

55 Sri Lanka

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2015	1998-1999
ICTD	2016	2020	

Time series We refer to historical archive data from 1965 to 2015, and we use ICTD for the period since then. We scale the historical archive data by ICTD for 2015-2020 using the 2014 ratios between the two sources as reference. Years 1998-1999 are interpolated.

Link to historical archive data: click here.

Harmonization

The main adjustments for Sri Lanka are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT missing or not reliable prior to 2001 for historical archive data. We rely on split CIT/PIT according to ICTD available from 1980, and extrapolate back to 1965 corroborated with historical sources • Small unallocable income tax after 2002 • Capital share of PIT: mean = 25.5%; ↓ trend (30 to 16%)
Social Contributions	Rely on external source (RPC database) for the whole period. Extrapolation forward to 2020 using the ratio RPC to ICTD from 2013 (last year available for RPC)
Property Taxation	Very small but non-zero values available from 1982 to 1997. Not available for the rest of the period
Decentralized Revenues	Not determined

Notes

- Ravinthirakumaran (2011) corroborates our archival long-run public finance series for the years 1977-2009, which also helps us trust the continuous time series from the earlier era.

56 Malawi

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	2004	
OECD	2005	2020	

Time series We refer to historical archive data from 1965 to 2004, and we use OECD for the period since then. We scale the historical archive data by ICTD for 1990-1993 using the 1989 and 1994 ratios between the two sources as reference.

Link to historical archive data: click here.

Harmonization

The main adjustments for Malawi are made on social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • Capital share of PIT: mean = 30%, constant
Social Contributions	Rely on external source (RPC database). SSA (2019) lists 2011 as the first year for the pension policy.
Property Taxation	Not included, neither in archival data nor in OECD
Decentralized Revenues	Not determined in historical archive period, not available in OECD period

Notes

- Our estimate in the historical period is slightly lower than that of, e.g., Shalizi and Thirsk (1990) and Chipeta (1998), both sourced from Malawi government reports, but we match this series on trends. It is possible that the GDP denominator from Malawi's government at that time is overestimated, vis-a-vis the WID figure we use. Our historical archive levels do match those of the ICTD in the 1980s, as well.
- The spike in personal income tax revenue in 2000-01 is genuine, per raw data, and represented a sharp upturn in the 'actual' vs. budgeted (previously predicted) revenues for that year.

57 Chile

Sources

Source	First year	Last year	Interpolation
Historical archive	1965	1979	1978, 1979
ICTD	1980	1989	
IDB-CIAT	1990	2020	

Time series We refer to historical archive data from 1965 to 1979, then use ICTD for the 1980s and finally rely on IDB-CIAT for the period since 1990. Years 1978 and 1979 are interpolated. We refine the historical archive using World Bank (1980) for central government revenues for 1960-77.

Link to historical archive data: click here.

Harmonization

The main adjustments for Chile are made on income taxation, social contributions and property taxes.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for years 1987-1989, interpolate using end-points as references and draw on additional historical sources • Capital share of PIT: mean = 21.6%, ↓ trend (26 to 14%)
Social Contributions	Rely on external sources for period prior to IDB-CIAT data (1990). For 1970-85 we use Corbo (1989) and for 1965-1975 we complement it with official government publications (consistent with evidence from Cerda (2005))
Property Taxation	Rely on external sources for the period after 1990 (UN data). Interpolation of the period 1980-1989 using 1979 and 1990 as reference
Decentralized Revenues	Included

Notes

58 Kazakhstan

Sources

Source	First year	Last year	Interpolation
ICTD	1994	1998	
OECD	1999	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include Kazakhstan before 1994. We refer to ICTD data from 1994 to 1998 and use OECD for the period since then.

Harmonization

The main adjustments for Kazakhstan are made on income taxation, social contributions and property taxes.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available from 1994 to 1998. Extrapolation of proportional ratios back using 1999 as reference • Capital share of PIT: mean = 16.9%, ↓ trend (22 to 15%)
Social Contributions	Available from 1995. Extrapolation back for 1994 (consistent with SSA (2019))
Property Taxation	Rely on historical archive for the period prior to OECD (pre-1999)
Decentralized Revenues	Included

Notes

- Oil revenue is enormously important for Kazakhstan, but much of it is non-tax revenue

59 Zambia

Sources

Source	First year	Last year	Interpolation
Historical Archive	1986, 1990-1991		
ICTD	2010	2020	

Time series We refer to historical archive data from 1965 to 2009. We use the evolving ratio of historical archive to ICTD total tax take (level) to interpolate the overall level of tax/GDP in missing years, and then we interpolate the values of 'within' ratios for the component taxes. We rely on the total tax revenue from ICTD for the period since 2010.

Link to historical archive data: click here.

Harmonization

The main adjustments for Zambia are made on income taxation, social contributions and property taxes.

Category	Adjustment
Income Taxation	• Split CIT/PIT available for all the non-interpolated years • Very small but non-zero unallocable income tax revenue since 1972 • Capital share of PIT: mean = 16.9%, ↓ trend (22 to 15%)
Social Contributions	Rely on external source (RPC database). Extrapolation of constant value from 2014 to 2020
Property Taxation	Very small but non-zero during the whole period
Decentralized Revenues	Not determined

Notes

- DiJohn (2010), citing Weeks and McKinley (2009), corroborate the general trends and levels of our data in 1990-2004, although we have to interpolate because we are missing many years.
- Colclough (1988) has a higher estimate than do our HA data for the period 1975-85, and matches the ICTD data from the 1980 onward. However, examination of our raw data shows that we have 'miscellaneous capital receipts' in the raw data, and these are not only highly variant from year to year in our data, but we also do not want to classify them as tax revenues. It seems likely that the gap between historical archive and ICTD-IMF in 1975 could also be what is driving the gap between historical archive and RPC total tax takes.

60 Romania

Sources

Source	First year	Last year	Interpolation
ICTD	1994	2020	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include Romania before 1994. We refer to ICTD data for the whole period since then.

Harmonization

Data from ICTD on Romania present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT available for all periods since 1994 • Very small but non-zero unallocable income tax revenue • Capital share of PIT: mean = 21.2%, ↓ trend (25 to 17%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

61 Senegal

Sources

Source	First year	Last year	Interpolation
Historical Archive	1965	1984	
ICTD	1985	1998	
OECD	1999	2020	

Time series We refer to historical archive data from 1965 to 1984. We use ICTD from 1985 to 1998, interpolating the relative share of taxes until 1990 using 1984 and 1991 as reference. We rely on OECD for the period since 1999.

Link to historical archive data: click here.

Harmonization

The main adjustments for Senegal are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT extrapolated back from 1972 to 1965 and interpolated for 1991-1992, corroborate using historical sources • Small but non-zero unallocable income tax revenue in multiple years • Capital share of PIT: mean = 29.5%, fluctuation (29-30%)
Social Contributions	Rely on external source (RPC database) for the period pre-OECD (pre-1998)
Property Taxation	Small but non-zero during the whole period
Decentralized Revenues	Not determined before OECD, not included for OECD period

Notes

- Our estimation for the historical archive period matches (with slight differences) previous literature (Boye, 1990; Chelliah, 1971)

62 Netherlands

Sources

Source	First year	Last year	Interpolation
OECD	1965	2020	

Time series We refer to OECD all the way back to 1965.

Harmonization

Data from OECD on Netherlands present detailed information on all the tax categories shown below. There are not major adjustments made.

Category	Adjustment
Income Taxation	• Split CIT/PIT available for all periods • Capital share of PIT: mean = 13.6%, ↓ trend (15 to 11%)
Social Contributions	Included
Property Taxation	Included
Decentralized Revenues	Included

Notes

63 Guatemala

Sources

Source	First year	Last year	Interpolation
HA	1965	1989	1984-1989*
OECD	1990	2020	

Time series We refer to historical archive data from 1965 to 1989 and we use OECD for the period since then. We interpolate the years 1986-1989 using ICTD's overall levels but relying on the share of each tax categories using 1983 and 1990 as reference.

Link to historical archive data: click here.

Harmonization

The main adjustments for Guatemala are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1965 • The unallocable income tax revenue in OECD since 1995 is always a 'solidarity tax', which is apparently assessed on corporations (per PwC 2020), so we re-assign it as a CIT • Capital share of PIT: mean = 27.8%, fluctuation (30-25%)
Social Contributions	Lack of data for the period pre-1990 (and probably starts in 1969 or 1977, per SSA 2017). We use OECD's level from 1990 extrapolated back to 1978 without re-scale to match the RPC number, as it does not appear that there is a fixed proportional gap between the two
Property Taxation	Really small but non-zero during the whole period
Decentralized Revenues	Not determined before OECD, included for OECD period

Notes

- Our estimation matches previous literature for the early historical archive period (Newlyn, 1985).

64 Chad

Sources

Source	First year	Last year	Interpolation
HA	1965	1982	
ICTD	1983	2009	
OECD	2010	2020	

Time series We refer to historical archive data from 1965 to 1982, we use ICTD for the period 1983-2009 and rely on OECD for the period since then. We interpolate the share of each tax category for periods 1978-1980 and 2006-2009.

Link to historical archive data: click here.

Harmonization

The main adjustments for Chad are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for 1978-1993. Interpolation of proportional ratios using 1977 and 1994, and corroborate with historical sources • Very small but non-zero unallocable income tax revenue during the whole period • Capital share of PIT: mean = 30%, constant
Social Contributions	Rely on external source (RPC database) for the period 1976 to 2010
Property Taxation	Very small but non-zero after 1994
Decentralized Revenues	Not determined before OECD, not included for OECD period

Notes

- The volatility in recent years is notable but also genuine (corroborated both in our raw data and in ICTD).

65 Cambodia

Sources

Source	First year	Last year	Interpolation
ICTD	1994	2009	

Time series In keeping with our rule to not include communist countries prior to their transition, we do not include Cambodia before 1994. We refer to ICTD for the whole period since 1994.

Harmonization

The main adjustments for Cambodia are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for the years 1994, 2004-05, and 2015. Interpolation of proportional ratios using years before and after • Very small but non-zero unallocable income tax revenue • Capital share of PIT: mean = 29%, ↓ trend (30 to 26%)
Social Contributions	Rely on external source (RPC database). Per SSA (2018), the first social security program began in 1997 (our data starts in 1998)
Property Taxation	Very small and in most periods zero
Decentralized Revenues	Not determined

Notes

- The World Bank (2020) does show a similar pattern to the one we observe, at least in the modern period (their data begins in 2002).

66 Ecuador

Sources

Source	First year	Last year	Interpolation
Historical archive	1973	1989	
OECD	1990	1992	
IDB-CIAT	1993	2020	

Time series We refer to historical archive from 1973 (first year of observation) until 1989. We use OECD data for the period 1990-1992, and we use IDB-CIAT data for the period since 1993.

Link to historical archive data: click here.

Harmonization

The main adjustments for Ecuador are made on income taxation.

Category	Adjustment
Income Taxation	• Split CIT/PIT not available for 1990-1992. Interpolation of proportional ratios using 1989 and 1993 as references • Unallocable income tax during IDB-CIAT period. Use additional information from SRI (2020), Ecuador's tax collection authority, to assign portion to PIT. • Unallocable income tax in 2019-2020 allocated proportionally based on the observed split CIT/PIT in 2018 • Capital share of PIT: mean = 28%, ↓ trend (30 to 27%)
Social Contributions	Included
Property Taxation	Small but non-zero during the whole period
Decentralized Revenues	Included

Notes

- Oil revenue is the biggest story in Ecuador. It rose from zero to 20% of public sector revenue from 1970-74, and to 48% by 1982, according to Bocco (2016). Meanwhile non-oil tax revenue decreased from 10.1% of GDP in 1972 to 5.4% in 1983 (ibid.). However, different data sources categorized this prominent source of revenue according to varying guidelines as well as the share of public revenue toggled between non-tax expropriated and privatized CIT revenue.

- Prior to 1996, it is not clear what part of public revenue is from oil and what is not. We therefore use IMF historical data and arrive at a series from 1973-89 that nearly agrees with the OECD by 1990 (when OECD begins). However, the historical archive may have erroneously classified oil revenue as indirect taxation instead of CIT. On the other hand, we corroborate OECD and oil vs. non-oil statistics by reference to the Ministry of Finance (MFE).
- Even though oil revenue makes the comparison hard, our estimations matches considerably well with previous literature. First, CEPAL (1991) corroborates our levels for pre-OECD period (including the 1985 spike). Second, for the period 1989-1992, García and Uquillas (1992) agree with the values from OECD. For more recent decades, different sources (Bucheli (2014), Fretes-Cibils, Shankar, and Currie (2008), Ministerio de Economia y Finanzas (2016)) match our estimation.

67 Zimbabwe

Sources

Source	First year	Last year	Interpolation
ICTD	1980	2020	1998

Time series We refer to ICTD for the whole period since 1980 (year of independence). Year 1998 is interpolated.

Harmonization

The main adjustments for Zimbabwe are made on income taxation and social contributions.

Category	Adjustment
Income Taxation	• Split CIT/PIT all the way back to 1980 • Unallocable income tax category during the whole period, most relevant between 2010-2015 (around 1% of GDP), use additional sources to allocate to firms versus workers • Capital share of PIT: mean = 29.8%, small fluctuations (30 to 29%)
Social Contributions	Rely on external source (RPC database) only since 1989, following SSA (2017)
Property Taxation	Very small during the whole period, and zero after 1998
Decentralized Revenues	Not determined

Notes

- The drop in revenues during the period 2005-2009 corresponds with the peak of the hyperinflation crisis in the country.

References

Afonso, J. R., Araujo, E. A., & Vianna, S. T. W. (2004). Carga tributária indireta no brasil: análise da incidência efetiva sobre as famílias. *Inter-American Development Bank*.

Alvaredo, F. (2010). The rich in argentina over the twentieth century, 1932–2004. *Top incomes: A global perspective*, *2*.

Amin, A. A. (1998). Cameroon's fiscal policy and economic growth. *The African Economic Research Consortium*.

Amir, H., Asafu-Adjaye, J., & Ducpham, T. (2013). The impact of the indonesian income tax reform: A cge analysis. *Economic Modelling*, *31*, 492–501.

Bachas, P., Fisher-Post, M. H., Jensen, A., & Zucman, G. (2022). Globalization and factor income taxation. *National Bureau of Economic Research*. (No. w29819)

Bank, S. A. R. (2016). *Tax chronology of south africa: 1979-2015.* Supplement to the South African Reserve Bank Quarterly Bulletin.

Bank, W. (2014). *Colombia's 2012 tax reform: Poverty and social impact analysis.* Author.

Barreix, A., Benítez, J. C., & Pecho, M. (2017). Revisiting personal income tax in latin america: Evolution and impact.

Bernardi, L., Fumagalli, L., & Gandullia, L. (2005). Tax systems and tax reforms in south and east asia: Overview of the tax systems and main policy tax issues. *MPRA Paper No. 18214*.

Blanchet, T., Chancel, L., Flores, I., & Morgan, M. (2021). *Distributional national accounts guidelines: Methods and concepts used in the World Inequality Database.*

Bocco, A. M. (2016). Ecuador: The state, public policy and capital accumulation, 1970–83. In *The state and capital accumulation in latin america: Argentina, bolivia, colombia, ecuador, peru, uruguay, venezuela* (pp. 88–105). Springer.

Bouin, O., Coricelli, F., & Lemoine, F. (1998). *Different paths to a market economy.* Organization for Economic Co-Operation and Development (OECD).

Boye, F. (1990). *A retrospective analysis of the senegalese economy.* Center for Economic Research on Africa, Department of Economics, School of . . . .

Burgess, R., & Stern, N. (1993). Taxation and development. *Journal of economic literature*, *31*(2), 762–830.

Castro, P., Junquera-Varela, R., Schenone, O., & Teixeira, A. (2009). Evaluation of reforms in tax policy and administration in mozambique and related ta–1994-2007. *IMF/IFAD. September*.

Cerda, R. A. (2005). Does social security affect retirement and labor supply? evidence from chile. *The Developing Economies*, *43*(2), 235–264.

Chelliah, R. J. (1971). Trends in taxation in developing countries. *IMF Staff Papers*, *1971*(002), A002. Retrieved from https://www.elibrary.imf.org/view/journals/024/1971/002/article-A002-en.xml doi: 10.5089/9781451947335.024.A002

Chipeta, C. (1998). Tax reform and tax yield in malawi.

Colclough, C. (1988). Zambian adjustment strategy—with and without the imf.

Corbo, V. (1989). *Public finance, trade, and development: the chilean experience* (Vol. 218). World Bank Publications.

Davoodi, H. R., & Grigorian, D. (2007). Tax potential vs. tax effort: a cross-country analysis of armenia's stubbornly low tax collection.

De Herdt, T. (2002). Democracy & the money machine in zaire. *Review of African Political Economy*, *29*(93-94), 445–462.

der Hoek, M. P. v. (2008). Enlarging the european union: Taxation and corruption in the new member states. In *Taxation and public finance in transition and developing economies* (pp. 11–23). Springer.

Dethier, J.-J. (2000). *Governance, decentralization and reform in china, india and russia.* Springer Science & Business Media.

DiJohn, J. (2010). State resilience against the odds: an analytical narrative on the construction and maintenance of political order in zambia since 1960.

Ekpo, A. H., & Ndebbio, J. E. (1996). Fiscal operations in a depressed economy: Nigeria, 1960-90. *African Economic Research Consortium*.

Fisunoglu, A., Kang, K., Arbetman-Rabinowitz, M., & Kugler, J. (2011). *Relative Political Capacity Dataset (Version 2.4).* Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/NRR7MB (Last access: September 2023) doi: 10.7910/DVN/NRR7MB

Fjeldstad, O.-H. (1995). Taxation and tax reforms in tanzania: A survey. *Chr. Michelsen Institute*.

Founou-Tchuigoua, B. (1989). La crise des finances publiques et la dénationalisation de l'etat: le cas du mali. *Africa Development/Afrique et Développement*, 19–41.

Fretes-Cibils, V., Shankar, R., & Currie, E. (2008). Fiscal sustainability and debt management in ecuador. *Ecuador's*, 43.

García, R., & Uquillas, E. (1992). Evolución y perspectivas de la economía ecuatoriana en 1992.

García-García, J., & Guterman-Bromberg, L. (1988). Medición del déficit del sector público colombiano y su financiación: 1950-1986. *Revista Ensayos Sobre Política Económica; Vol. 7. No. 14. Diciembre, 1988. Pág.: 115-133.*.

Gauthier, B., & Gersovitz, M. (1997). Revenue erosion through exemption and evasion in cameroon, 1993. *Journal of Public economics*, *64*(3), 407–424.

Gauthier, B., Soloaga, I., & Tybout, J. (2002). A firm's-eye view of commercial policy and fiscal reforms in cameroon. *The World Bank Economic Review*, *16*(3), 449–472.

Ghafur, A., & Chowdhury, O. H. (1988). *Bangladesh.* Economics Office, Asian Development Bank.

Gillis, M. (1985). Micro and macroeconomics of tax reform: Indonesia. *Journal of Development Economics*, *19*(3), 221–254.

Hasan, P., Kemal, A., & Naseem, S. (1997). Learning from the past: A fifty-year perspective on pakistan's development [with comments]. *The Pakistan Development Review*, *36*(4), 355–402.

Interamerican Development Bank (IDB) & Inter-American Center of Tax Administration (CIAT). (2018). *Revenue Collection database.* Data retrieved from https://www.ciat.org/idb-ciat-revenue-collection-database/?lang=en. (Last access: September 2023)

International Centre for Tax and Development. (2020). *UNU-WIDER Government Revenue Dataset. Version 2023.* Data retrieved from https://doi.org/10.35188/UNU-WIDER/GRD-2023. (Last access: September 2023)

Ivanova, A., Keen, M., & Klemm, A. (2005). The russian 'flat tax'reform. *Economic policy*, *20*(43), 398–444.

Jansen, K., & Khannabha, C. (2012). The fiscal space of thailand: An historical analysis. In *Fiscal space* (pp. 347–420). Routledge.

Junguito, R., Rincón, H., et al. (2004). La política fiscal en el siglo xx en colombia. *Borradores de economía*, *318*.

Karingi, S. N., & Wanjala, B. (2005). *The tax reform experience of kenya* (No. 2005/67). WIDER Research Paper.

Kwack, T., & Lee, K.-S. (1992). Tax reform in korea. In *The political economy of tax reform* (pp. 117–136). University of Chicago Press.

Lou, J., & Wang, S. (2008). *Public finance in china: reform and growth for a harmonious society.* World Bank Publications.

Macha, R. R., Lado, E. P. Z., & Nyansera, O. C. (2018). An empirical analysis of tax ratios and tax efforts for kenya and malawi. *African Journal of Economic Review*, *6*(2), 152–171.

Madrigal-Delgado, G. d. J. (2021). Recaudación del impuesto predial en méxico: desafío del federalismo fiscal. *Investigación administrativa*, *50*(127).

Maina, A. W. (2014). *Income taxes and economic performance in kenya* (Unpublished doctoral dissertation).

Mascagni, G. (2016). A fiscal history of ethiopia: Taxation and aid dependence 1960-2010. *ICTD Working Paper 49*.

McLure Jr, C. E. (1992). Income tax reform in colombia and venezuela: A comparative history. *World Development*, *20*(3), 351–367.

Mokhtari, M., & Ashtari, M. (2012). Understanding tax reform in the central asian republics. *Journal of Asian Economics*, *23*(2), 168–178.

Nachega, J.-C. (2005). Fiscal dominance and inflation in the democratic republic of the congo. *IMF working paper*.

Naughton, B. (2007). *The chinese economy.* MIT press Cambridge, MA.

Newlyn, W. T. (1985). Measuring tax effort in developing countries. *The Journal of Development Studies*, *21*(3), 390–405.

Nord, R., Sobolev, Y., Dunn, D. G., Hajdenberg, A., Hobdari, N., Maziad, S., & Roudet, S. (2009). Tanzania: the story of an african transition. *International Monetary Fund (IMF).*

Nyrop, R. F. (1976). *Area handbook for egypt.* US Department of Defense, Department of the Army.

Organization for Economic Co-Operation and Development. (2020). *Revenue Statistics.* Data retrieved from OECD.stats, https://stats.oecd.org/index.aspx?DataSetCode=REV. (Last access: September 2023)

Osei, R. D., & Telli, H. (2017). Sixty years of fiscal policy in ghana. *The Economy of Ghana Sixty Years after Independence*, *66*.

Osoro, N. E. (1993). *Revenue productivity implications of tax reform in tanzania.* Centre for the Study of African Economies, University of Oxford, Oxford, GB.

Owen, M. D. E. W., & Robinson, M. D. O. (2003). *Russia rebounds.* International Monetary Fund.

Pinto, B. (1987). Nigeria during and after the oil boom: A policy comparison with indonesia. *The World Bank Economic Review*, *1*(3), 419–445.

Prasetyo, K. A. (2018). Tax administration reform and the society in indonesia: Some lesson learnt. *Working Paper*.

Preobragenskaya, G., & McGee, R. W. (2003). Taxation and public finance in a transition economy: a case study of russia. *Available at SSRN 480862*.

Ravinthirakumaran, K. (2011). The relationship between government revenue and expenditure in sri lanka. In *International conference on business and information.*

Ribeiro, M. P., Villafuerte, M., Baunsgaard, T., & Richmond, C. J. (2012). Fiscal frameworks for resource rich developing countries. *Staff Discussion Notes*, *2012*(004).

Shalizi, Z., & Thirsk, W. R. (1990). *Tax reform in malawi* (Vol. 493). World Bank Publications.

Smith, H. H. (1970). *Area handbook for the united arab republic (egypt)* (Vol. 43). US Government Printing Office.

Syeda, M. H. (2015). Making an impact analysis of social protection programs in pakistan. *Journal of the Research Society of Pakistan*, *52*(1).

Wawire, N. H. (1991). *An empirical assessment of tax performance in kenya: 1958 to 1989* (Unpublished doctoral dissertation). Kenyatta University.

Weeks, J., & McKinley, T. (2009). Does debt relief increase fiscal space in zambia? the mdg implications. In *Economic alternatives for growth, employment and poverty reduction: Progressive policy recommendations for developing countries* (pp. 180–208). Springer.

Yoo, I. (2000). Experience with tax reform in the republic of korea. *Asia Pacific Development Journal*, *7*(2), 75–104.

Zolt, E. M., & Bird, R. M. (2005). Redistribution via taxation: The limited role of the personal income tax in developing countries. *UCLA Law Review*, *52*, 05–22.