Globalization and Factor Income Taxation*

Pierre Bachas[†], Matthew Fisher-Post[‡], Anders Jensen[§], Gabriel Zucman[¶]

November 2021

Abstract

How has globalization affected the relative taxation of labor and capital? We address this question by constructing and analyzing a database of effective macroeconomic tax rates covering 150 countries since 1965, obtained by combining national accounts data with government revenue statistics. We obtain 5 main findings. 1) The average effective labor and capital tax rates have converged globally, due to a 10 points increase in labor taxation and 5 points decline in capital taxation since the 1960s. 2) The decline in capital taxation is concentrated in high-income countries and driven by the collapse of corporate profits taxation. 3) By contrast, capital taxation has increased in developing countries since the 1990s, especially in large economies. 4) Using a variety of research designs, we find that the rise in effective capital tax rates in developing countries follows the rise of trade openness, while trade has a negative effect on capital taxation in developed countries. In both groups of countries, trade leads to an increase in effective labor taxation. 5) Countries with higher trade have a larger share of output produced in the corporate sector (where income is easier to observe and therefore to tax) and less self-employment. Globalization thus appears to have had asymmetric effects on tax structures, contributing to the decline in tax progressivity in rich countries but also to the growth of tax capacity in developing countries.

*We gratefully 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, Rafael Proenca, Roxanne Rahnama and Anton Reinicke for excellent research assistance. We thank Denis Cogneau, Gordon Hanson and Dani Rodrik for insightful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

†World Bank Research, pbachas@worldbank.org

‡Harvard Kennedy School and Paris School of Economics, mfp@psemail.eu

§Harvard Kennedy School and NBER, anders_jensen@hks.harvard.edu

¶UC Berkeley and NBER, zucman@berkeley.edu

1 Introduction

How has globalization affected the relative taxation of labor and capital? Has international economic integration eroded the amount of taxes effectively paid by capital owners, shifting tax burdens to workers and reducing the progressivity of tax systems? And if so, which countries have been most affected by this process and through which mechanisms? Answering these questions is critical to better understand the distributional effects and long-run social sustainability of globalization.

Yet addressing these questions has so far proven difficult, due to a lack of internationally comparable statistics on tax structures and tax progressivity. While a number of studies analyze changes in taxation in high-income countries (e.g., Mendoza, Razin, and Tesar, 1994; Egger, Nigai, and Strecker, 2019), much less is known about the evolution of taxation in emerging economies. This makes it hard to contrast the effect of globalization on tax structures in countries with different income levels, and limits our ability to test hypotheses on the causal effect of international economic integration.

To help fill this gap, this paper builds and analyzes a database of effective tax rates on labor and capital covering more than 150 countries and more than half a century (1965-2018). Constructed following a common methodology, these series allow us to take a global and comparative perspective on the evolution of tax structures. Our effective tax rates capture all taxes paid at all levels of government: corporate income taxes, individual income taxes, payroll taxes, property taxes, estate and inheritance taxes, etc. This makes it possible to estimate total tax wedges—for instance the gap between what it costs to employ a worker and what the worker receives—and how these wedges vary internationally. This approach also allows us to put specific taxes into a broader perspective. For instance we can ask whether the noted decline in corporate income taxes (on mobile capital) has been offset by a rise in property taxes (on immobile capital).

To maximize the time and geographical scope of this database, we conducted a large-scale digitization and harmonization of historical data published by national statistical offices, which we combined with existing (but limited in time and space) series published by the United Nations, the OECD, and the IMF. The construction of the database proceeds in three steps. Using national accounts data we first compute total labor and capital income in each country. Using government revenue statistics we then classify all sources of government revenue into either labor taxes, capital taxes, or indirect taxes. Combining these two inputs, we compute effective macroeconomic tax rates on each of the two factors of production—labor and capital—by dividing taxes paid by the corresponding income flow.

Using this database, we make two contributions. The first is to establish a set of new facts. Taking a global perspective, we find that average effective labor and capital tax rates have converged globally, due to a 10 points increase in labor taxation and 5 points decline in capital taxation since the 1960s. Since capital income is always more concentrated than labor income, this evolution implies a large decline in tax progressivity globally. The decline in capital taxation is driven by the collapse in the taxation of corporate profits, from close to 30% in the 1960s to less than 20% in recent years, a decline which has not been offset by increased taxation of other forms of capital income.

These global trends mask heterogeneity by development level. Most importantly, the decline in capital taxation is concentrated in high-income countries, where the effective capital tax rate fell from close to 40% in 1965 to about 30% in 2018. In the United States for instance, the average capital tax rate fell from from more than 40% in the 1960s to 25% in 2018. In developing countries by contrast, the effective capital tax rate has increased from about 10% to 20%, with all the increase happening after 1995 and primarily in large economies. Between 1995 and 2018, the effective capital tax rate rose from 5% to 10% in Mexico; 18% to 28% in Brazil; 7% to 10% in India, for example. Despite the recent increase, effective capital tax rates in developing countries remain typically way below those observed in high-income countries.

While the decline in capital taxation in high-income countries is not entirely surprising given the known decline in statutory corporate tax rates (e.g., Slemrod, 2004), the rise observed in low- and middle-income countries in the era of hyper-globalization had not been noted in the literature before. This rise appears to be robust. It holds when we exclude China and oil-rich countries; when we restrict the analysis to a balanced sample of countries; and under different weighting schemes. Our finding are also robust to alternative approaches to computing capital and labor income in non-corporate businesses (where factor shares are not directly observable) and to alternative ways of assigning different types of taxes to capital versus labor.

Why did effective tax rates on capital fall in high-income countries but rise in developing countries? Our second contribution is to quantify the effect of international trade in these trends. Beginning in the 1980s, a number of developing countries adopted pro-trade policies, opening their markets and reducing tariffs on imported goods. This policy revolution, combined with technological improvements such as the rise of containerized shipping, led to a boom in international trade and reshaped the economy of countries such as Mexico, India, and China. We show that for developing countries effective capital tax rates rose in tandem with trade integration: countries with relatively high trade openness before 1995 saw their effective capital tax rate rise before 1995; countries that liberalized trade after 1995 saw an increase in this rate after 1995. Through a variety of research designs, we find that this association is likely causal—the rise in capital taxation follows the rise in trade openness in developing countries. In contrast, trade has a slightly negative effect on the effective capital tax rate in rich countries. In both groups of countries, trade leads to an increase in the effective tax rate on labor.

We implement three different strategies to establish these results. First, we run simple non-parametric estimations of the 5-year relation between changes in effective tax rates and changes in trade openness. Second, we analyse policy-induced trade liberalization events which occurred in seven large developing countries. We focus on the six liberalization events that caused the largest and most sudden reduction in trade barriers cited in the review papers by Goldberg and Pavcnik (2007) and Goldberg and Pavcnik (2016), plus the often discussed WTO accession of China in 2001 (Brandt et al., 2017). Using synthetic control methods to create counterfactuals for each country-event, we present event study graphs and estimate dynamic effects in regression models. Last, we extend the two instruments for trade openness presented in Egger, Nigai, and Strecker (2019), which combined yield a strong first stage in the full set of countries. In each of our research designs, trade openness leads to a substantial rise of the effective tax rate on capital in developing countries, and a slightly smaller increase in the effective tax rate on labor. These results are robust to a wide range of sensitivity checks.

We then investigate heterogeneity and potential mechanisms. In both the time-series correlation and IV designs, the positive impact of trade on capital taxation, in addition to being concentrated in developing countries, is stronger in more populous countries and in countries with restrictions on capital flows. This is consistent with the notion that larger countries are less exposed to tax competition, and suggests that countries that manage their capital accounts may be less exposed to some of the forms of tax avoidance that are facilitated by global economic integration, such as corporate profit shifting. We also show that, in developing countries, trade liberalization leads to an increase in the corporate share of domestic product and to a transition from self-employment to salaried employment. Since employees and large corporations are more visible and generate information trails (Kleven, Kreiner, and Saez, 2016; Basri et al., 2019), this evolution may have allowed governments in developing countries to tax capital and labor more effectively. Finally, we find a negative effect of trade openness on the statutory corporate income tax rate in both developed and developing countries.

Taken together, these findings suggest a nuanced view of how integration has impacted capital and labor taxation over the last half-century. By facilitating the transition towards formal employment in large firms, international integration can increase tax capacity; but by opening up new tax avoidance opportunities and cross-border competition, it can also reduce effective tax rates on mobile production factors (including through statutory tax rate reductions). In large developing countries, the tax capacity effect quantitatively dominates the tax avoidance and competition effects, leading to an increase in both effective labor and capital tax rates. In high-income countries, where tax capacity is already high, the tax avoidance and competition effects seem to dominate, leading to a shift away from capital and onto labor taxation.

The rest of the paper proceeds as follows. In Section 2, we relate our work to the existing literature. In Section 3, we describe the methodology and data collection. Section 4 presents novel facts on factor income taxation over the long-run. In Section 5, we present graphical results on the association between trade openness and factor income taxation.

Section 6 studies the impacts of major trade liberalization events and in Section 7 we use instruments for trade and investigate heterogeneity and mechanisms. Section 8 concludes.

2 Related Literature

2.1 Globalization and Taxation

Our paper relates to the literature on globalization and taxation. Since Adam Smith (1776), economists have conjectured that increased openness pushes governments to reduce the taxation on capital to prevent its flight and that the resulting revenue shortfalls can be recovered by increasing the taxation of labor—the less mobile factor (Bates, Da-Hsiang, and Lien, 1985; Rodrik, 1997). This hypothesis is often referred to as the "internationalization" or "efficiency" proposition. These considerations are thought to be most prevalent in rich countries with high international openness (Rodrik, 1997), in small countries (Wilson, 1999; Kanbur and Keen, 1993), and in the post-1995 period of hyper-globalization (Egger, Nigai, and Strecker, 2019).

Empirically, Egger, Nigai, and Strecker (2019) provide evidence that globalization led to an increase in labor taxation for the middle class and a decline in labor taxes paid by the top 1 percent in OECD countries since 1994. At the macro level, studies have produced mixed results (reviewed in Adam et al., 2013).[1] Due to data limitations, these studies are restricted to developed countries or to a single tax (e.g. corporate income tax). We build a new global dataset including developed and developing countries and all labor and all capital taxes; and we implement several research designs to identify the effect of trade on effective labor and capital tax rates.

Methodologically, our work builds on Mendoza, Razin, and Tesar (1994) who measure macroeconomic effective tax rates in a number of OECD countries, work that was extended by (Carey and Rabesona, 2004; McDaniel, 2007) to other high-income countries.

[1] Rodrik (1997) finds that trade openness is associated with a decline in effective capital tax rates and an increase in effective labor tax rates for 14 OECD countries with high levels of capital mobility, between 1965 and 1991. In a sample of 14 OECD countries between 1981 and 1995, Swank and Steinmo (2002) find that trade is not associated with changes to effective tax rates, neither on capital nor on labor. Slemrod (2004) similarly finds, in a panel of countries between 1980 and 1995, that trade is not associated with either the average of statutory corporate income tax rate, once country and year fixed effects are included.

2.2 Tax Capacity in Developing Countries

Our paper also relates to the growing literature on tax capacity in developing countries. This literature highlights a number of factors driving the rise of taxation in developing countries, including the rise of salaried employment (Jensen, 2021), growing capacity to observe income (Pomeranz, 2015), the threat of whistle-blowing in large firms (Kleven, Kreiner, and Saez, 2016), and investments in tax capacity (Besley and Persson, 2014).[2]

Through a variety of research designs, we highlight the role of international trade in the rise of tax capacity in developing countries since the mid-1990s. The increase in effective tax rates we document is consistent with previous studies showing that trade has a positive effect on growth (e.g., Goldberg and Pavcnik, 2016) and growth is associated with higher tax rates (e.g., Besley and Persson, 2014). Our approach goes further in two dimensions: first it quantifies the role of trade; second it documents some mechanisms through which trade liberalization affects tax rates by showing an increase in the corporate share of domestic product and a transition from self-employment to salaried employment.[3]

Our paper also complements studies on the tax revenue effects of trade liberalization (Baunsgaard and Keen, 2009; Cage and Gadenne, 2018; Buettner and Madzharova, 2018). Cage and Gadenne (2018) find that these events had negative effects on tax revenue in low-income countries before 1995. Our paper highlights positive revenue effects of trade liberalization but looking at different outcomes (effective capital and labor tax rates, as opposed to indirect tax revenues) and sample periods (the positive effects we obtain are concentrated after 1995). These two sets of results are complementary, as trade liberalization may both reduce trade tax revenue upon impact and eventually improve the capacity of developing countries to raise taxes on capital as activity becomes more concentrated in large firms, which are easier to tax.

[2]Cogneau, Dupraz, and Mesple-Somps (2021) and Cogneau, Dupraz, Knebelmann, et al. (2021) study the very long-term evolution of tax revenue in 18 former French colonies in Africa, including the colonial era. Interestingly, they find that independence did not lead to a drastic change to tax capacity.

[3]Our IV results suggest that trade integration causes a transition to formality, consistent with recent within-country evidence from McCaig and Pavcnik (2018). Other recent work on trade and informality include Dix-Carneiro and Kovak (2017) and Dix-Carneiro, Goldberg, et al. (2021).

3 Construction of Factor Shares and Effective Tax Rates

This section describes the construction of our database of effective tax rates on labor and capital. The data covers the 150 most populous countries from 1965 to present, with exceptions only for pre-independence, civil war, and command economy eras.

Section 3.1 describes our conceptual framework to measure factor shares and factor income taxation. Section 3.2 describes our data construction.

3.1 Conceptual Framework and Methodology

3.1.1 Factor Shares

Net domestic product (NDP) at factor prices can be expressed as:

$$NDP = CE + OS_{CORP} + OS_{HH} + OS_{PUE} + NIT \tag{1}$$

where CE refers to labor income, the 'compensation of employees' in wages and salaries; OS_{CORP} is capital income, the 'operating surplus of corporations' in the corporate sector (net of employee compensation); OS_{HH} is capital income in the household sector, from actual and imputed rental income; and OS_{PUE} is 'mixed income,' the return to self-employment and informal or private unincorporated enterprises. NIT represents indirect taxes net of subsidies. Thus, factor-price net domestic productis defined as $Y = NDP - NIT$.

We calculate the labor share as:

$$\theta_L = \frac{Y_L}{Y} = \frac{CE + \alpha \cdot OS_{PUE}}{CE + OS_{CORP} + OS_{HH} + OS_{PUE}} \tag{2}$$

$\frac{Y_L}{Y}$ is the labor share of net domestic product, comprising the compensation of employees and a share α of mixed income.

The capital share of net domestic product is then $\theta_K = 1 - \theta_L$, which can be derived from the same national accounting income concepts above, as:

$$\theta_K = \frac{Y_K}{Y} = \frac{OS_{CORP} + OS_{HH} + (1-\alpha) \cdot OS_{PUE}}{CE + OS_{CORP} + OS_{HH} + OS_{PUE}} \tag{3}$$

We clarify several methodological points from equations (2) and (3). First, we subtract indirect taxes (net of subsidies) from net domestic product, since we assume that they are borne by capital and labor in proportion to their factor shares (see Browning, 1978; Saez and Zucman, 2019). Second, we exclude the consumption of fixed capital (depreciation) from our concepts of capital income flows in the national accounts—this is the distinction between net domestic product and gross domestic product. In our view, the former concept is more economically meaningful than the latter (i.e., domestic product *net* rather than gross of depreciation), so we frame all of our estimates in terms of net domestic product. Third, net government surplus is zero by construction and hence does not need to be distributed to capital and labor (see Lequiller and Blades, 2014).[4] Finally, as for all GDP statistics, these estimates do not balance net foreign income.[5] We focus on the domestic net product as we focus on tax bases which primarily tax domestic sources. In practice, adjusting for net foreign income in the denominator does not alter our results.

The labor share of mixed income In the data we retrieve (discussed in Section 3.2 below), we directly observe all components of equations (2) and (3), except for the labor share of mixed income, α. The difficulty of observing the labor intensity of self-employment and unincorporated enterprises has been well discussed in Gollin (2002) and Karabarbounis and Neiman (2014), and our main results follow the literature which assumes that $\alpha = 70\%$ (e.g., see Alvaredo et al., 2020).

However, we test the sensitivity of our results to this assumption, following the innovative method developed in Guerriero (2019) and ILO (2019). This method, which we implement as a robustness check, attributes to self-employed workers a similar wage to the one they would have earned in an employer-employee environment, matching on a range of observable characteristics. We extend the estimates in ILO (2019) to more country-years

[4]Large public-sector enterprises are usually included in the corporate sector in national accounts, but even when they are large, they rarely account for a significant proportion of the economy's overall corporate operating surplus.

[5]For example, many citizens in Lesotho earn labor income in South Africa, and many international corporations book profits in Ireland. These income flows are reported domestically in South Africa and Ireland, respectively, whereas a complete accounting for 'net foreign income' would increase the labor income within Lesotho's accounts, and would decrease capital income within Ireland's accounts.

using ILOSTAT (2021) data on self-employment shares of the workforce for all countries since 1991, and impute values to extend the series further back.[6]

3.1.2 Effective tax rates on capital and on labor

Having defined the respective tax bases for labor and capital, we allocate all sources of tax revenue to labor, capital, or a mix of the two. We make the following assignment: (1) Corporate income taxes, as well as wealth and property taxes, are allocated to capital income. (2) Payroll taxes and social security payments are entirely allocated to labor income. (3) Personal income taxes (PIT) are allocated between labor and capital, reflecting the fact that personal income is composed of salaries, capital income and gains, and mixed income.

Personal income tax allocation to factors of production In countries where the PIT includes both labor and capital income and where the tax base covers a large share of the population, policy-studies have found that approximately 15% of PIT collection is from capital income. We adjust this PIT capital share assumption in two ways, to account for countries with narrower PIT tax bases and/or dual tax schedules (one tax schedule on labor income, another on capital income). First, countries with a high PIT exemption threshold, and thus a large proportion of exempted low-income taxpayers, have a larger share of capital income in their PIT. This is because richer taxpayers derive a larger share of their income from capital. We use data from Jensen (2021), which systematically documents the location of the PIT threshold across countries, to adjust for the share of capital income in the

[6] For each of the three classes of self-employed workers—self-employed employers, 'own-account' workers, and 'contributing family members'—we estimate a coefficient that is specific to each country-year. This coefficient is the ratio of the wage of the average employee to that of the given class of self-employed workers. Total labor income for a given country-year, then, can be retrieved as follows:

$$Y'_L = CE + \sum \frac{CE}{s_{emp}} \cdot \gamma_i s_i$$

where CE is the total compensation of employees in the national accounts; s_{emp} represents the share of employees in the workforce (whose collective earnings equal CE), such that $\frac{CE}{s_{emp}}$ is the average employee wage; and s_i denotes the share in the workforce of each type of self-employed laborers. Each self-employment category i corresponds to a specific earnings coefficient γ_i relative to the average employee compensation. Capital income Y'_K, then, is $Y - Y'_L$, and factor shares are calculated as before.

PIT at the country-year level.[7] Second, some countries offer more favorable tax treatment to dividends than to labor income, thus lowering the share of capital taxed through the PIT. Using data from the OECD on whether dividends receive a lower tax rate, we compute the ratio of the statutory rate on dividends to that of the top rate on labor income, and adjust the capital share of PIT revenue down when this ratio is below one.

Together, these adjustments imply that the assumed capital share of PIT revenue depends on the personal income tax exemption threshold location and the extent to which capital income and labor income are taxed along differing schedules within the personal income tax code. The share of the PIT allocated to capital income varies between 7% and 35%, depending on countries and years. This share falls over time, from a weighted global average of 19% in 1965 to 14% in 2018, due to both a reduction in PIT exemption thresholds, and to the adoption of lower tax rates on dividends in selected countries.[8]

Overall, our allocation of taxes to capital or to labor is simple and transparent. We discuss robustness to alternative allocation rules throughout the text.

Total factor taxation The total tax revenue assigned to labor and capital is:

$$T_L = \sum[\lambda_{ic} \cdot \tau_i] \quad \text{and} \quad T_K = \sum[(1 - \lambda_{ic}) \cdot \tau_i] \tag{4}$$

where λ_{ic} is the allocation to labor of each type of tax τ_i in country c (see Table 1).

The effective tax rates on labor and on capital, ETR_L and ETR_K, are computed by dividing tax revenue collected by the size of the respective labor and capital tax bases:

$$ETR_L = \frac{T_L}{Y_L} \qquad \text{and} \qquad ETR_K = \frac{T_K}{Y_K} \tag{5}$$

[7]We predict the threshold level for country-years where we do not have data, based on GDP per capita. Jensen (2021) finds that the threshold level is tightly associated with GDP per capita, both across and within countries over time.

[8]We use net domestic product to weight the global average. The unweighted global average drops from 23% in 1965 to 20% in 2018. Smaller, lower-income countries tend to have (i) a higher PIT exemption threshold, and (ii) an integrated PIT system that does not favor capital relative to labor income.

These measures of macroeconomic effective taxation capture realized tax revenue, reflecting countries' historical tax rules and investment in capital and labor.[9] While we build on the work of Mendoza, Razin, and Tesar (1994) and Carey and Rabesona (2004), our revised methodology provides cohesion with national accounts, considers all income sources in domestic product, integrates mixed income, and vastly expands the global coverage of ETR estimates.

3.2 Data

Table 2 illustrates the coverage and sources of our panel dataset, described below.

3.2.1 Components of National Income

Our novel factor shares database expands coverage beyond existing data—in both time and geography—with comprehensive estimates of disaggregated national income components. To estimate factor share statistics for 156 countries since the 1960s, we create a harmonized panel of national accounts which combines data from the UN System of National Accounts, the World Inequality Database, and complementary sources. We briefly describe these sources here.

From the United Nations System of National Accounts (UN SNA, 2008), we retrieve online national accounts statistics that cover almost 4,000 country-years using the production and income accounts of the 'Main Aggregates and Detailed Tables.' These variables include gross value added; compensation of employees; gross operating surplus of corporations; gross operating surplus of private unincorporated enterprises; consumption of fixed capital; and indirect taxes net of subsidies. Beyond the online data, the System of National Accounts Office (at the UN Statistics Division) provided us with electronic archival data on the components of GDP, with over 2,000 country-year observations from the 1960s and 1970s. These observations are framed in the SNA1968 format (UN SNA, 1968).

[9] The measures of ETRs computed in (5) are called 'backward-looking' ETRs in the literature. A separate literature tries to model all statutory features of the tax system at a point in time in order to measure 'forward-looking' ETRs (see, e.g., King and Fullerton, 1984; Devereux, 2004). They model incentives at the margin and predict taxes on new investments, but are less adapted to study long-term global trends.

To our knowledge, this is the first factor shares dataset that harmonizes SNA2008 with SNA1968 data. The two series match well on levels and trends during their periods of overlap.[10] The main difference in the raw SNA1968 historical data compared to the SNA2008 online data, is that the historical data does not disaggregate 'mixed income' and 'operating surplus of the household sector' (imputed rental income of owner-occupied housing), but subsumes the two under 'operating surplus.' Therefore, we impute the split of mixed income vs. household operating surplus in SNA1968 data according to the split that occurs in SNA2008 data at the year of stitching.[11] We also follow the United Nations guidelines to stitch these two series together (UN, 2018). When national income components are occasionally missing, we make simple assumptions and interpolations.[12]

We complement the UN SNA data with additional information on national income components, depreciation, net foreign income, and indirect taxes. For missing data on indirect taxes, we draw from our tax revenue data (discussed below). From the World Inequality Database (WID, 2020), we draw aggregate variables on total gross domestic product and net national income, including depreciation (consumption of fixed capital) and net foreign income.[13] Table 2 shows the final coverage and sources of our national income components data.

The integration of these new data sources explains the expanded coverage we obtain, compared to reference database developed in Karabarbounis and Neiman (2014).[14] We also extend the scope of our measure beyond theirs: whereas the Karabarbounis and Neiman (2014) paper focused primarily on factor shares in the corporate sector, we capture the

[10]The overlap usually occurs in the 1970s, as countries transitioned from the historical SNA framework to the modern one. The SNA1968 data contains the 'compensation of employees' and 'operating surplus of corporations,' so we have reason to be confident in minimal mismeasurement across eras.

[11]In this estimate we also include a time-varying global coefficient which estimates the share of each component according to a country's income level. See Fisher-Post (2020) for more on this.

[12]For example, one or more components of national income can be missing in the raw data, but we retrieve its value using national income accounting identities. We refer to these values as implicitly available in the raw data. Beyond that, we make simple imputations and interpolations for minor components missing from the national income decomposition, whose values are not critical.

[13]For data on the labor share of net foreign income, we refer to the IMF Balance of Payments data (Blanchet and Chancel, 2016). We use current accounts' "primary income, compensation of employees abroad, net (*credit - debit*)" as the labor share of net foreign income, under the hypothesis that self-employment income accounts for a small proportion of income from abroad.

[14]Karabarbounis and Neiman (2014) restrict their sample to online data, and to countries with at least 15 consecutive years of complete-case data. We improve on their coverage thanks to the offline UN SNA data.

entire economy. One advantage of this new approach is that we can include net foreign income (see Zucman, 2015) and imputed rental income of owner-occupied housing.[15] Karabarbounis and Neiman (2014) also note the importance of measuring factor share trends net of capital depreciation—even when this is only retrieved with difficulty from SNA data. Indeed, we extend their work on capital depreciation by measuring this for all countries where it exists, and imputing the remainder (see also Blanchet and Chancel, 2016). In Appendix A we discuss these points in further detail.

3.2.2 Tax Revenue Data

After expanding global coverage on factor shares, our new tax revenue dataset aims to achieve two similar objectives. First, to obtain disaggregated tax revenue data by type of tax (separating personal income taxes from corporate income taxes, and measuring payroll and property taxes). Second, to achieve a global coverage, with a focus on integrating previously unused historical data from developing countries.

We first stitched together existing high-quality data using online sources from OECD (2020) and ICTD/UNU-WIDER (2020) for recent years, and computerized (but offline) historical data from the IMF GFS (2005) for older years. Second, we retrieved thousands country-year observations of historical revenue data from the Harvard University Library archives,[16] as well as online sources from national statistical offices, finance ministries, and scholarly publications. Third, we assigned to each disaggregated tax revenue source a tax label, following the OECD's taxonomy of tax classification (see OECD, 2020).

Table 2 details the data sources used. When available, OECD tax revenue data is our preferred source. This is because it covers all types of revenue—including payroll tax and decentralized tax revenues—back to 1965 for OECD countries, and it already classifies taxes according to the useful OECD taxonomy. OECD data accounts for 41% of the country-year observations in our dataset. Its drawback is its very partial coverage of non-OECD countries: it only covers 93 countries and only over the past two decades.

[15] While capital income from housing can confound the study of within-industry trends, or the elasticity of substitution between capital and labor, or gross vs. net returns to capital, we are less concerned with these aspects than with the overall proportion of capital income in net domestic product.

[16] Lamont Library, Government Documents section.

To increase our coverage we turn to two other frequently used tax revenue datasets: the ICTD/UNU-WIDER (2020) (17% of observations) and the IMF GFS (2005) (10% of observations).[17] These datasets achieve near worldwide coverage but face limitations. They often only start in the 1980s; they do not follow the tax revenue classification of the OECD; they do not always separate personal versus corporate income taxes; and they often miss data on payroll taxes and decentralized revenues. Thus, we complement those data with additional data collection, drawing historical public finance data from government documents manually digitized at the Harvard Library archive, which eventually accounts for 30% of our country-year observations.

In stitching together country-by-country, long-run time series of disaggregated tax revenues, we follow three guiding principles. First, we aim to only rely on two data sources by country: the OECD when it exists, and the alternative source with the best coverage of time and types of taxes.[18] Our data hierarchy choice also depends on which source best matches the OECD data over their shared time frame. Second, for countries' series with gaps, we interpolate data, but only for a maximum of four years between two reliable data points, with triangulation. Finally, we check country-specific scholarship and policy reports to triangulate across data sources and to identify events (tax reforms, economic or political crisis) which can explain discordance across sources or large changes in tax revenue across years.

Tax revenues are disaggregated as finely as possible by source, according to the OECD tax classification scheme discussed above (OECD, 2020). Given our objective to measure taxes on capital and labor, we focus on three dimensions. First, we separate income taxes into personal and corporate income. Second, we ensure that payroll taxes are included: this required us to the date of creation for social security schemes in every country, and to add new data sources focused exclusively on payroll taxes. We digitized these from the UN SNA and obtained further data from Fisunoglu et al. (2011). Third, we study taxes on

[17]The ICTD/UNU-WIDER data draws heavily from the IMF GFS data that is available online, for recent years. Thus our use of the IMF GFS database is restricted to the *historical* dataset, not available online. This source provides data from 1972-89 and fills gaps from the OECD and historical archives data. The ICTD does not report pre-1980 data. Beyond the IMF GFS source, the ICTD refers to IMF Article IV reports, which we also use when necessary.

[18]Historical archive data is our second in priority since it often contains the most detailed revenue disaggregation and often goes back to the 1960s or 70s.

property (as a principal source of decentralized tax revenues), with attention to heavily decentralized countries within a federal tax system, as they are likely to have substantial sub-national government revenues.

Finally, we made use of the additional datasets described in the following sub-section.

3.2.3 Other datasets

Tax revenue robustness: On subnational/decentralized tax revenues from the OECD-UCLG (2019). For oil and natural resource production, we combine Haggarty and Shirley (1997); Guriev, Kolotilin, and Sonin (2011); Estrin and Pelletier (2018); Ross and Mahdavi (2015); Mahdavi (2020); and Esberg and Perlman (2020).

International trade: The principal source of our data for trade analysis is UN COMTRADE (2020). We will discuss the connection with trade in the following sections.

4 Global Trends in Tax Revenues, Factor Shares and Effective Tax Rates

Our new dataset achieves a substantial improvement in coverage of factor shares and factor taxation. Appendix Figure A1 shows that our data coverage fluctuates from 86% of World GDP in 1965 to 98% in 2015, as the number of countries grows from 78 to 156.[19] The main entrance event into the panel corresponds to the early 1990s transformation to market economies of the former communist countries.[20] This includes Russia and, importantly, China, for which a modern market-based tax system arguably appears in 1994.[21]

Thus 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 98% of World GDP. At their time of entry

[19]Even in the most recent years we do not cover 100% of world GDP, as we did not try to digitize data from countries with under 1 million inhabitants when these were not available through one of the online sources.

[20]Late decolonization and end of civil wars are other reasons to enter the panel later than 1965.

[21]See Appendix A for a case study of China. We argue that 1994 can be considered the most relevant start year for China's time series to enter our global dataset.

into the dataset, ex-communist countries account for 8% of World GDP (4.5% for China and 3.25% for Russia). We discuss how their entry impact the taxation patterns described below.

With this new dataset, we document the global evolution from 1965 to 2018, of tax revenues, factor shares, and effective taxation on capital and labor. Our objective is to show time series for each outcome which can be interpreted as the global value, worldwide, in each year. For example, the global effective tax rate on capital equals worldwide capital tax revenue divided by worldwide capital income in the same year. Its interpretation is the expected tax rate that a unit of capital chosen at random across the world would face at that point in time.[22] For each outcome we show the global levels and trends, and then show the trends separately for high vs low and middle-income countries.

Figure 1 shows the time series of tax revenue as a share of net domestic product (NDP), separated into its main components: corporate income taxes, property and asset taxes, personal income taxes, payroll taxes, and indirect taxes (VAT and tariffs). Globally, tax revenue as a share of NDP increased from 26% to 32% between 1965 and 2015. This is driven by an increase in payroll and personal income taxes, which went from 11% to 16%. Indirect taxes slightly rose over the past 50 years from 8% to 9%, while revenues from taxes on capital (corporate and property) stagnated at around 6%.

We observe two striking differences in tax revenue patterns between high-income versus low- and middle-income countries. First, tax revenue as a share of NDP is currently much higher in rich countries than in developing countries (37% vs. 23%), as is already well documented. Second, in developing countries, all types of taxes increase their revenue collection over time (particularly from 1990 onward), including those on capital. By contrast, corporate income tax revenue decreased over time in high-income countries (as a share of domestic product), and revenues from property taxes stagnated. Higher tax revenue in rich countries came primarily from the expansion of payroll taxes between 1965 and 1985.

[22]One should keep in mind that global figures depend on countries' changing shares of world GDP: Figure A2 shows how the weight of different countries (or group of countries) evolved over time, highlighting the notable growth of China's weight in the past 20 years, and to a lesser extent the rest of developing countries. The weight of China, in world GDP, was far less in the pre-1994 era during which it is excluded. Refer to Appendix A for further discussion.

Figure 2 shows the capital share of net domestic product over time (solid line) and the capital share within the corporate sector (dotted line). Globally, the capital share of world income increased from 20% to 26%. The striking global trend is due both to a rise in the capital share of rich countries, from 25% to 28%, and to a steeper rise in developing countries, from 34% in 1965 to 41% by 2008, before a recent decline to 37% in 2015. The capital share within the corporate sector followed the same evolution as that of the aggregate capital share; it increased from approximately 18% in 1965 to 23% in 2015. Again, this is due to both a modest increase in the capital share of the corporate sector in rich countries, and a steep increase in developing countries.

Figure 3 shows the evolution of the effective tax rates on labor (red) and capital (blue); and, within capital income, the evolution of the effective tax rate on corporate profits (dashed blue).[23] Globally, the ETRs on capital and labor largely converged between 1965 and 2018. This is due to a large increase in labor taxation and a mild decrease in capital taxation. The global ETR on labor increased from 15% to approximately 25%, while over that same period, the ETR on capital decreased from an average of 33% in the mid-1960s to an average of 27% in the late 2010s. Within the corporate sector, the global ETR on corporate profits saw a more pronounced decline, from 28% in 1965 to 18% in 2018.[24]

These global trends mask important heterogeneity by development level. The decline in the ETR on capital is entirely concentrated in high-income countries. In contrast, the ETR on capital doubled in developing countries from 10% to 20% over the long-run, following a steep acceleration after 1995. Over the past 50 years, the ETR on labor increased by more in developed than in developing countries, even though the base was already much larger in rich countries (18%) than in poorer countries (6%) at the start of our data in 1965.

[23]The ETR on corporate profits is computed as the ratio of the revenue from the corporate income tax over the operating surplus of the corporate sector.

[24]Figure A4 shows the ETR series in a fully balanced panel of all countries since 1965, by way of imputation, to control for the changing sample composition over time (most importantly China and Russia pre-1994). Globally, the entry of Russia and China hardly changes the time-series of ETRs. This is because, on a global scale, these countries only represent 8% of GDP in 1993, and Russia's ETRs are close to the global average at its entry in 1994. When we focus on developing countries only—where these two countries now represent a substantial share of GDP—imputing missing country-years raises both the ETR on labor and on capital pre-1994 by close to 2 percentage points. The resulting series do not have an economic interpretation (since ETRs are arguably undefined in the imputed years), but the imputations shown in Figure A4 account for composition effects due to countries entering the sample.

One of the most notable pattern of our data is the sizable increase in effective capital taxation within low- and middle-income countries. This ETR_K increase is a surprising and novel finding, but it is robust to a variety of sensitivity analyses, and a battery of data quality checks. Figure 4 shows the evolution of ETRs in developing countries for several subsamples of the developing countries. First, given the importance of China and its extraordinary growth over the last 30 years, we show the trends in ETRs excluding China and other ex-communist countries from the sample. The rise in the ETR on capital over the period is reduced, highlighting that China is an important contributor to the observed pattern; however the rise in ETR_K remains, from approximately 10% to 13% from 1965 to 2015. As a second robustness check, we exclude oil-rich countries—defined as deriving 7% or more of their GDP from oil—since these countries' corporate tax revenue may often heavily depend on oil. The sample of non-oil rich developing countries shows an even more striking increase in effective capital taxation than the complete sample (from 10% to 24% over the past half-century). Lastly, we separate the sample between the 18 largest (non-oil rich) developing countries whose 2018 population exceeds 40 million, and the 55 countries with a population under 40 million. We find that the increase in effective capital taxation is much more pronounced in large countries, where it increases from approximately 10% to 25%; while the increase is more modest in the smaller countries (from 8% to 12%). Similarly, the effective tax rate on corporate profits saw a much larger increase in populous countries.

5 Correlation in Trade and Factor Income Taxation

5.1 Motivation

What are the determinants of the long-run trends in factor taxation? A large literature in economics and political science focuses on the role of increased international market integration. The prevailing theory is that cross-border factor mobility limits the scope of taxation. This is consistent with the long-run decline of the ETR on capital observed in developed countries. On the contrary, since the mid-1990s, many developing countries saw both a rise in the ETR on capital and increased trade integration. Focusing on developing countries only, we observe that the positive association between trade and taxation runs deeper: when we separate countries based on their initial level of trade in the pre-1995 period, early globalized countries saw trade and the ETR on capital rise in tandem prior to the 1990s, and stagnate thereafter (Appendix Figure A11). Developing countries which participated in the 'second wave' of globalization (after the early-1990s proliferation of trade agreements) saw an increase in their trade and capital taxation in the past 25 years. These trends motivate our systematic analysis of the impact of globalization on factor taxation, which we undertake in the remainder of the paper.

5.2 Setup

In our first empirical strategy, we simply present the within-country association between trade and our outcomes of interest: factor shares and effective tax rates on labor and capital. We measure trade as the share of imports and exports relative to GDP. We create 5-year growth rates within-country in both the trade measure and the outcomes. By studying associations over the medium-run of five years, we smooth out confounding short-run variations in national income components driven by business cycles (Dao, Koczan, and Lian, 2017). To visualize these associations, we plot binned scatters of each outcome against trade, after residualizing all variables against year fixed effects. Each dot in the figure corresponds to a ventile (20 equal-sized bins) of the residualized trade openness distribution; we add back the (non-residualized) mean of each variable to ease interpretation. The resulting figure provides a non-parametric estimate of the medium-run within-country association, conditional on global time trends. Since we do not add any other controls or weights, this empirical design is flexible and transparent. We also show the best-fit linear estimate based on the underlying country-year panel of growth rates.

5.3 Results

The results are displayed in Figure 5. We observe a large positive association between the within-country growth in trade openness and effective taxation of capital; the association is well-approximated by the linear estimator. Trade openness is also associated with an increase in ETR_L, although the slope is less pronounced than for capital. In Appendix Figure A6, we show that these associations are robust to varying the time-horizon between 2 and 10 years; the association grows in magnitude over longer time-horizons for ETRs, but remains constant for factor shares. We also show in Appendix Figure A7 that the relationship is robust to alternative measures of ETRs.

Previous studies on ETRs and globalization focus on rich countries (Section 2). In Figure 6, we find that the associations between trade and effective tax rates differ significantly between high-income versus low- and middle-income countries. Importantly, there is a mild *negative* slope between trade openness and ETR_K in high-income countries, compared to a strong positive association in developing countries. While the negative slope in high-income countries is consistent with the cross-border mobility hypothesis discussed above (see Section 2), the opposite result in developing countries suggests that the mediating factors which determine how globalization impacts taxation differs across development levels. We return to this point in Section 7.

Figure 5 also shows a positive association between trade openness and the capital share of income; this association is almost twice as large for the corporate capital share than for national product capital share. Trade may thus positively impact capital taxation in the economy, both through increasing capital's share of aggregate income and by raising capital's effective tax rate.

The finding that trade is positively associated with the capital share is consistent with prior empirical work (Section 2). However, it stands at odds with classical trade models, such as Heckscher-Ohlin, which predict that trade openness leads to an increase in the country's abundant factor (i.e., usually labor in developing countries). Rather, this finding is consistent with bargaining models, in which increased opportunities to relocate abroad leads capital to command a stronger bargaining position and capture a larger share of the economic rents from globalization (Rodrik, 1998; Harrison, 2005; Rodriguez and Ortega, 2006). The positive association is also consistent with the global value chains theory. A large share of cross-border trade is in intermediate goods rather than in final goods, and takes place within narrow product categories (Feenstra and Hanson, 2001). In this theory, high-income countries focus on the capital-intensive portions in the global value chain and outsource their labor-intensive processes to developing countries. But these outsourced processes may be relatively more capital-intensive in developing countries. Trade integration then benefits capital in both groups, even if it is the scarce resource in developing countries.

6 Event Studies Around Large Trade Liberalization Events

6.1 Research Design

In this section, we implement an event study analysis of *policy-induced* trade liberalization events. To graphically discern sharp breaks from trends in the macro-aggregate outcomes, we search for liberalization events which caused large trade barriers reductions. Consequently, we focus on the six liberalization events studied in the review papers by Goldberg and Pavcnik (2007) and Goldberg and Pavcnik (2016) on the effects of trade policy, and add the often discussed China WTO accession event of 2001 (Brandt et al., 2017). The resulting set of events is Colombia in 1985; Mexico in 1985; Brazil in 1988; Argentina in 1989; India in 1991; Vietnam in 2001; and, China in 2001.[25] These events present two compelling features. First, they are characterized by large reductions in measurable, policy-induced trade barriers: For instance, Brazil reduced average tariff rates from 59% to 15% percent, India from 80% to 39%, and China from 48% to 20%. Second, these events have been extensively studied, with in-depth narrative analyses of the conditions surrounding the timing of the reforms. We reproduce these arguments and discuss endogeneity concerns in Appendix C.2.

For each event and outcome, we construct a synthetic control country following the methodology described in Abadie, Diamond, and Hainmueller (2010). The use of synthetic controls has recently been used in the analysis of policy reforms, both at the national level (Jaeger, Noy, and Schoefer, 2021; Gruber, Jensen, and Kleven, 2021; Smith, 2015) and

[25] 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. A detailed description of all seven trade liberalization events is provided in Appendix C.

sub-national level (Havnes and Mogstad, 2015; Billmeier and Nannicini, 2013; Akcigit et al., 2021). The synthetic control is created as a weighted average over the donor pool of all countries in our full sample that did not experience an event at any point in time. To construct the weights, we match on the level of the outcome of interest in the 10 years prior to the event. The weights are chosen to minimize the mean squared prediction error between the event-country and the synthetic control countries in the ten years pre-reform. Appendix Table A2 details the matching for each treated country and each outcome.[26] We create event-study graphs in levels as the unweighted average of the outcome variable, separately over all treated countries and all synthetic control units, and separately by relative time to the liberalization event.

We are interested in the impact of trade liberalization events on several outcomes (trade, factor shares, factor taxation). In our baseline approach, we create a separate synthetic control for each event and each outcome. The upside of this approach is that it increases the likelihood of creating a synthetic control with strongly similar pre-trends (Akcigit et al., 2021); the downside is that, for a given country-event, the countries used in the synthetic control possibly differ across outcomes. We therefore implement a parallel strategy, where we simultaneously match on all outcomes of interest for each country-event (similar to Jaeger, Noy, and Schoefer, 2021). This leads to a small deterioration in the pre-trends, as expected, but to very similar point-estimates (see Appendix C.2).

We also implement a regression version of the event-study, where we include country and calendar year fixed effects. Concretely, we estimate the regression equation:

$$Y_{it} = \sum_{j=-10, j\neq -1}^{10} \mu_j * \mathbb{1}(j=t)_t * D_i + \theta_t + \kappa_i + \pi_{Year(it)} + \epsilon_{it}$$

using the seven treated and seven synthetic control countries in the 10 years before and after the events. In the regression, θ_t are event-time fixed effects, κ_i are country-reform (unit) fixed effects, and $\pi_{Year(it)}$ are calendar year fixed effects. The calendar-year fixed effects control for common shocks to factor shares and taxation which may be correlated

[26] The table shows that, most often, the weights are concentrated amongst two to three control countries; this helps to avoid over-fitting, whereby the synthetic match is artificially created by combining idiosyncratic variations in a large sample of control countries.

with clusters of reforms that happen in the same year. D_i is a dummy equal to one if observation i is a treated country. Hence, μ_j captures the difference between the treated countries and the group of synthetic controls in event time j, relative to the immediate pre-reform year $j = -1$ (omitted time-period). Statistical inference based on a small sample size as ours should be approached with caution (Abadie, Diamond, and Hainmueller, 2010). Therefore, we plot 95% confidence bounds based on the wild bootstrap method (Cameron, Gelbach, and Miller, 2008), clustered at the country-event level.

Finally, in addition to plotting the full set of dynamic treatment effects, we estimate the difference-in-differences (DiD) effect from the following model

$$Y_{it} = \mu^{DiD} * \mathbb{1}(j \geq 0)_t * D_i + \theta_t + \kappa_i + \pi_{Year(it)} + \epsilon_{it}$$

Here, μ^{DiD} can be interpreted as the average treatment effect over the first 10 years after the liberalization event. We report classic as well as wild bootstrap standard errors, clustered at the country-event level. We also report the average post-reform coefficients based on the imputation methodology of Borusyak, Jaravel, and Spiess (2021), which overcomes estimation problems arising from two-way fixed effects and heterogeneous event-times, like our context. The full set of DiD coefficients μ^{DiD} are reported in Appendix Table A3, and the full methodological details in Appendix C.2.

6.2 Results

The results for the five main outcomes are shown in Figure 7. The left-hand panels display the event-studies in levels, while the right-hand panels display the regression-based event-studies. Starting with trade openness, we observe that the synthetic control matches the average treated country closely during the 10 years prior to the event. Trade openness increases in the year of the event and the trend in trade breaks in post-reform years, compared to the stable pre-trends. The absence of a dip in the immediate pre-reform years limits concerns about intertemporal substitution, although some liberalization events were predictable (especially in China and Mexico where the event is WTO accession).

Turning to the factor share outcomes, the trade liberalization events coincide with a positive break from trend in the capital share of domestic product.[27] This effect is even sharper for the capital share of corporate income, where the figure shows that the increase materializes precisely in the reform-year. Apart from a small dip in the very early post-reform years, the synthetic control replicates well the mildly positive trend observed in the treated countries pre-reform (even in the long-run).[28]

Finally, the impacts on factor taxation are displayed in the two bottom panels of Figure 7. Similarly to the immediate increase in trade openness and the capital share of corporate income, the effective tax rate on capital sharply increase following the liberalization event. Both ETR on capital and labor break from the stable pre-trend at the time of liberalization, but the capital effect is much larger than the labor effect. Despite the small sample size, the dynamic post-treatment effects for all outcomes are most often significant at 5%; the p-value for the joint significance of all post-dummies is below 0.05 for all outcomes (except for the capital share of corporate income, which equals 0.057). Based on the difference-in-differences model, the liberalization events led to a 10 percentage point average increase in trade openness over 10 years and a 4.3 (1.8) percentage point increase in the effective tax rate on capital (labor) (Appendix Table A3).

6.3 Discussion

There are two important comments in this event-study design. First, the timing of the liberalization event may coincide with unobservable changes in determinants of factor shares and taxation. The relatively stable trends in treated countries all the way until the liberalization event imply that these confounding changes would have to sharply coincide with the event itself. But the small sample of treated countries makes it a plausible concern that the average effect is driven by a concurrent unobservable shock in one of the treatment countries. The narrative analyses of the reform contexts, reproduced from pre-existing studies in Appendix C), do not highlight domestic confounding shocks. These narratives

[27] Wacziarg and Welch (2008) similarly found that liberalization events in developing countries, based on the Sachs-Warner openness index, led to an increase in physical capital investment as a share of GDP

[28] This also helps to alleviate concerns that the true counterfactual level would be overstated if trade flows and returns to capital were diverted away from countries in the synthetic control.

notwithstanding, in any case in Appendix Figure A10 we remove one treated country at a time and apply the synthetic control method to the remaining treatment-set. The resulting dynamic treatment-effects for all sub-sets of treated countries is similar to the average treatment-path. This robustness checks suggests that the sharp changes in the immediate post-liberalization periods are not driven by a country-specific unobservable shock. Remaining confounding variation would have to coincide more systematically with liberalization events that are scattered both across time and regions of the world.

Second, the interpretation of the event-impacts depends on whether other reforms or confounding economic shocks occurred in the post-reform years (Rodriguez and Rodrik, 2001). In Appendix C, we discuss events which occurred in the post-liberalization years. Further trade reforms were implemented after the liberalization events in some countries (Mexico joined NAFTA in 1994; Argentina and Brazil joined MERCOSUR in 1991). Some countries also liberalized their cross-border capital flows (Mexico removed capital inflow restrictions in 1989; India liberalized foreign direct investment rules in 1993). Mostly, these external reforms occurred in the medium-run (5 to 10 years after liberalization), while we observe graphically sharp effects in the short-run (first 5 years). The short-run results are more likely to be directly related to the policy-effects of trade liberalization.[29]

In summary, our preferred interpretation is that the short-run breaks from stable pre-trend reflect impacts of policy-induced trade liberalization. The medium to long-run effects may reflect deepened cross-border integration of both trade and non-trade factors, longer-run equilibrium impacts (Goldberg and Pavcnik, 2007), or slow-moving confounding factors. This discussion highlights that the graphical results in the synthetic events provide compelling evidence for a causal impact of trade on factor shares and taxation, but we caution against attaching too much importance to the specific magnitudes uncovered.

Finally, we note that our results are based on a selected set of liberalization events characterized by very large reductions in trade-costs. It is possible that the finding of

[29] Wacziarg and Wallack (2004) study if trade liberalization events in developing countries coincide with domestic reforms. Out of our seven events, only Mexico has a confounding domestic privatization reform within the first five years of our event-year; Brazil (privatization) and Colombia (broad market-oriented reforms) had confounding reforms between 5 and 10 years after liberalization; and, the remaining four countries had no confounding reforms. The results are robust to excluding Mexico, the single country with a confounding reform in the short-run (Figure A10).

sharp effects in the immediate post-reform years may not carry over to other studies of liberalization reforms. Moreover, we note that the events occur in developing countries which are large in population size and have relatively more legal constraints on capital mobility (Chinn and Ito, 2006). The impacts of trade liberalization events are most likely to carry over to other countries with similar characteristics (see also Section 7.3).

7 Regressions with Instrumental Variables for Trade

In this section, we study the impact of trade in a regression setting. We use instruments to alleviate endogeneity concerns and to investigate heterogeneities and mechanisms.

7.1 Research design

We are interested in estimating the following regression model:

$$y_{ct} = \mu * trade_{ct} + \Theta * X_{ct} + \beta_c + \pi_t + \epsilon_{ct} \tag{6}$$

where y_{ct} is the outcome of interest in country c in year t, $trade_{ct}$ is the share of import and exports in total net domestic product and μ_c and π_t are country and year fixed effects. We cluster the error term, ϵ_{ct}, at the country level. We also estimate models which include, in X_{ct}, proxies for confounding determinants of factor shares and factor taxation (Rodrik, 1997; Harrison, 2005): the exchange rate, gross capital formation, log of population, log of GDP per capita, and capital openness.

We weigh observations by their yearly national domestic product to match the global trends procedure in Section 4. As such, the coefficient μ represents the effect of trade in the representative country across all levels of development and over 55 years (1965-2020). OLS estimation may, however, be biased due to reverse causality and unobservable confounding factors which correlate with changes in trade. Since we are interested in uncovering the causal effect in the representative country, the challenge becomes to find exogenous variation which impacts trade similarly across development levels and over time. This leads us to focus on the two instruments used in Egger, Nigai, and Strecker (2019).

The first instrument relies on the general structure of quantitative general equilibrium models of trade (Eaton and Kortum, 2002; Arkolakis, Costinot, and Rodriguez-Clare, 2012). Under the assumptions in the standard gravity model, this instrument uses the average bilateral trade frictions between exporting and importing countries as the source of variation (aggregated to the country-year level).[30] In our context, this instrument is valid if the distribution (not the level) of trade costs among individual country-trading pairs is not influenced by the level of factor shares or factor taxation in the import or export country. The second instrument exploits the time-series variation in global oil prices interacted with a country-specific measure of access to international markets. Specifically, access at the country-level is captured by the variance of distance from the three most populated cities to the closest maritime port. Intuitively, this time-invariant measure captures the internal geography of a country which is an important component of transportation costs. Following a global shock to oil prices, the transportation costs will be larger in countries with less concentrated access to maritime ports, leading to a larger drop in imports and exports.[31] Conceptually, both instruments capture identifying variation in trade costs driven by economic forces. Both instruments are explained in detail in Appendix B.

We extend the data coverage of these instruments to our full set of countries and time periods.[32] Since the IV estimate of equation (6) recovers a local average treatment effect (LATE), it becomes important to understand the relevance of each instrument across our full sample. This is investigated in Appendix Figure A12. We find that each instrument is relevant in different subsamples: the oil-distance instrument has a strong first stage in later time periods and in higher-income settings; in contrast, the gravity instrument has a stronger first-stage in earlier time periods and in lower-income settings.[33] Importantly, this exercise reveals that an IV estimate based on either of the individual instruments will be

[30]Other studies which leverage the structure of the gravity model to create instruments for trade include Frankel and D. Romer (1999), Wacziarg (2001), and Anderson and Wincoop (2003).

[31]Storeygard (2016) shows that oil price hikes lead to an increase in the cost of fuel at the pump, suggesting that oil price fluctuations do impact trade costs. In the transport logistics literature, oil prices are an important determinant of transportation costs within networks (Gross, Hayden, and Butz, 2012).

[32]Egger, Nigai, and Strecker (2019) focuses primarily on OECD countries between 1980 and 2007.

[33]Conceptually, the oil-distance instrument may be stronger in high-income countries if economic development is associated with improvements in domestic road networks (holding the physical distances from cities to maritime ports constant). We measure transportation networks in the latest year available; this introduces possible measurement error which weakens the instrument's relevance in earlier periods.

driven by first-stage compliers which have different characteristics than the representative country in the full sample. Alternatively, restricting the analysis to the subsamples where an individual instrument has a strong first stage in general generates invalid IV estimates which are upward-biased (Abadie, Gu, and Shen, 2019). Instead, we combine the two instruments to improve statistical power in the full sample (Mogstad, Torgovitsky, and Walter, 2020), and estimate a LATE that is representative across development levels and time periods.[34] The LATE identified using multiple instruments retains an intuitive interpretation: it is a weighted combination of the instrument-specific LATEs using the instruments one at a time.[35]

7.2 Main results

The OLS and IV estimation of equation (6) for our four core outcomes are presented in Table 3. Panel A presents the OLS results, while Panels B through D present different IV specifications. The OLS and IV coefficients have the same sign, but the IV coefficients are systematically larger. In Panel B, we estimate the IV model without controls X_{ct}. The 1st stage is strong, with an F-statistic of 26.07. The IV estimation shows a positive impact of trade on capital's share, both in national income and within the corporate sector.[36] Turning to the taxation of factor shares, the IV-results indicate that trade leads to statistically significant increases in the effective tax rate of both capital and labor, but the effect on capital (0.375) is more than twice as large as the effect on labor (0.163). The IV-coefficient on ETR_L appears more precisely estimated (p-value = 0.003) than the coefficient on

[34]We modify the weighting scheme in Section 4 and measure constant net domestic product using purchasing power parity rather than market exchange rates. Using PPP gives more weight to lower-income countries which, due to the variation in instrument strength across development levels (Figure A12), significantly improves the first-stage statistical power. Using market exchange rate yields qualitatively similar results that are only less precisely estimated (results available upon request).

[35]The weights are a function of the strength of each instrument in the first-stage regression in the full sample (Angrist and Imbens, 1995). In our setting, a stronger weight is placed on the oil-distance instrument (see first-stage regressions in Appendix Table A5). Appendix Table A11 reports the IV-results based on each individual instrument.

[36]The re-allocation towards capital inside the corporate sector implies that our results are not simply confounded by a positive impact of trade on the corporate share of national income. Moreover, most previous studies of global trends in factor incomes focus on shares within the corporate sector; our results can thus relate to those trends directly.

ETR_K (p-value=0.081), but we will see that the latter masks important heterogeneities in sub-samples with opposite signs.

The μ^{IV} estimates in Panel B are globally representative, by way of including weights based on countries' yearly national domestic product. Panel C shows that the results are robust to removal of weights: while the 1st stage strength is reduced (F-statistic=8.415), the results are broadly similar.[37] In particular, trade's impact on the effective taxation of capital (0.250, p-value=0.018) continues to be positive and larger in magnitude than the impact on labor taxation (0.133, p-value=0.013).

In Panel C, we include the set of country-year varying controls contained in X_{ct}. The inclusion of controls may help improve the precision of the estimates as well as increase the likelihood that the exclusion restriction holds. The inclusion of controls leads to a general reduction in the p-values; trade continues to have a positive impact on the capital factor share and a larger positive impact on capital factor taxation than on labor taxation.

These results are robust to a battery of additional robustness checks that are reported in the Appendix. First, the results hold with different measures of trade intensity (Appendix Table A6): winsorizing trade openness at the 95th percentile; focusing on openness of trade in goods; and using the growth rate in the value of trade (in constant USD). Second, since one of the instruments relies on oil price variation, we confront the concern that our estimating variation is correlated with trends in factor shares and effective taxation specific to oil-producing countries by allowing this group to be on a separate non-parametric time path (Table A8). Third, the results are broadly similar when we use different measures of factor shares and effective taxation. Trade's impact on the capital share is larger when using the ILO-adjusted method; in turn this leads to a smaller trade impact on the ETR on capital, which is now comparable in magnitude to the impact on the ETR on labor (Table A9). The results remain unchanged when we use the alternative assignment of taxes to capital versus labor proposed by Mendoza, Razin, and Tesar (1994) and when we change the mixed income assignment parameter [α in equation (2)] to extreme values of 0% or 30% (Appendix Table A10).

[37] The results are also robust to using population size, rather than total domestic product, as weights - see Table A7).

7.3 Mechanisms

In this subsection, we investigate two potential mechanisms: the 'internationalization' hypothesis and the 'tax capacity' hypothesis. In the internationalization hypothesis, increased openness pushes governments to reduce the taxation of capital to prevent its flight while the resulting revenue loss can be recovered by increasing the taxation of labor, the less mobile factor (Bates, Da-Hsiang, and Lien, 1985). To test this hypothesis, we study the impacts of trade on the statutory rate of the corporate income tax (CIT). Admittedly, this is a coarse measure of CIT incentives since non-rate dimensions of the CIT base are equally important to determine the tax liability (Kawano and Slemrod, 2016). Nonetheless, the CIT rate is a salient policy instrument that is easier to measure across space and time. The CIT rate saw a dramatic reduction over the past 55 year, both in low and high-income countries (Figure A13), which could reflect pressures induced by globalization.

In the tax capacity hypothesis, increased trade openness changes labor markets and firm structure which relax tax enforceability constraints. Previous work, convincingly shows that tax enforcement is difficult in the absence of third-party information trails (Pomeranz, 2015; Naritomi, 2019). For example, the coverage of self-employed' economic activity is limited, leading to high tax evasion rates (Kleven, Knudsen, et al., 2011), and the movement from self-employment to formal wage employment over development is associated with growth in tax enforcement capacity (Jensen, 2021). At the same time, prior work in international trade shows positive effects on the transition from informal to formal employment (McCaig and Pavcnik, 2018)) and firm-size – which could raise the tax-enforceable share of both capital and labor income.

We first study test these hypotheses for the full sample of countries. In Panel A of Table 4, we find that increased trade leads to a *reduction* in the corporate income tax rate. This result is robust to alternative specifications.[38] At the same time, Panels B and C show that trade causes a transition out of self-employment and an increase in the corporate share of GDP. Panel D shows a positive effect of trade on the effective tax rate on capital inside

[38]Consistent with past studies on the determinants of tax policies (C. Romer and D. Romer, 2010; Zidar, 2019), the outcome variable is the first-difference of the CIT rate: $\Delta CIT_{t,t-1}$. Results are robust to alternative outcomes, including: the level of the CIT rate while controlling for the lagged CIT rate; and a reform-tracker which changes value when the CIT rate changes. Results available upon request.

the corporate sector. Thus, several countervailing forces appear to be at play: growth in employee-employment and in the corporate sector increase the tax enforceability of both capital and labor, and hence their effective taxation, while active government policies in the form of rate reductions lower the effective tax burden on capital.[39]

Heterogeneity by development level Motivated by the contrasting long-run trends of ETR_K in high versus low and middle-income countries (Figure 3), we then investigate whether trade has different impacts on taxation across development levels. We estimate heterogeneous IV effects by interacting the trade variable with a dummy for high income countries, $\mathbb{1}(HighIncome)$:

$$y_{ct} = \mu * trade_{ct} + \kappa * trade_{ct} * \mathbb{1}(HighIncome) + \Theta * X_{ct} + \beta_c + \pi_t + \epsilon_{ct} \tag{7}$$

The two instruments leverage very distinct sources of exogenous variation. As such, the interpretation of heterogeneous IV-coefficients is challenged by the confounding possibility that each instrument captures different LATEs Mogstad2020. We alleviate this concern in Table A11[40]

The results are presented in Table 5. In Column (1), we find that trade has a strong positive effect on ETR_K in developing countries, but the effect is completely overturned in developed countries. At the same time, Column (2) shows that trade increases ETR_L in both samples, but the effect is strongest in high-income countries. Column (3) shows that trade decreases the CIT rate in both samples, although the magnitude is again strongest in rich countries. In stark contrast, Columns (4)-(6) show that the positive effects of trade on the capital share, the employee-share and the corporate share are entirely concentrated in developing countries, with null effects in high income countries. These results point to heterogeneous mechanisms depending on countries income levels: globalization contributed to the decline in capital taxation in developed countries by

[39]In Appendix Table A12, we confirm that increases in the statutory CIT rate are associated with increases in both ETR_K and CIT revenue as a share of GDP. These associations are estimated within country over time, and are slightly larger in magnitude in rich versus in middle and low-income countries.

[40]The IV-estimates based on using each instrument separately are broadly similar and in the range of the estimates based on using both instruments simultaneously. Notwithstanding, we approach with caution the heterogeneity interpretation of the IV estimates.

reducing the statutory CIT rate; in developing countries, although this policy effect is still at play, it is more than counteracted by the increased tax enforceability of firms' activity, which leads on net to a rise in effective capital taxation.

Additional results on the internationalization hypothesis Beyond the split by development level, we can estimate additional sources of heterogeneity to corroborate the 'internationalization hypothesis'. The literature argues that efficiency concerns related to capital flight are more pronounced in countries with less legal restrictions on capital mobility (Rodrik, 1997; Chinn and Ito, 2006); in low population countries (Wilson, 1999); and, in the post-1995 era of hyper-globalization (Egger, Nigai, and Strecker, 2019). We test each of these mechanisms, by looking at heterogeneous treatment effects for the statutory rax rate and the effective tax rate on capital using equation (7).

The results in Table 6 are qualitatively consistent with the internationalization hypothesis. In Panel A, we find that increased trade openness leads to a reduction in the CIT rate which is larger in small countries, in countries with less capital restrictions, and in the post-1995 era. In panel B, we find that the size of the trade-effect on ETR_K is reduced in small versus large countries, in countries with more versus less legal capital flow restrictions, and in the post-1995 versus pre-1995 era, thus mirroring the policy effects.

An additional prediction of the internationalization hypothesis is that trade should lead to an increased tax burden on the relatively less mobile factor, labor, to balance the revenue loss from decreased capital taxation. In Panel C, we indeed find that the increase in the effective taxation of labor is qualitatively larger in smaller countries, in countries with less capital restrictions and in the post-1995 era. These results are consistent with a shift in the tax burden towards labor whenever capital is relatively more mobile.[41]

Additional results on the tax capacity hypothesis The internationalization hypothesis is derived in the context of a model which assumes that government responses to trade openness are, on net, revenue-neutral. In contrast, the tax capacity hypothesis is motivated

[41]Cross-border mobility is likely to be stronger for high-income labor earners than for middle and low income labor earners (Egger, Nigai, and Strecker (2019)). A meaningful analysis of trade's impacts on statutory labor tax systems would thus require the measure of incentives by income percentiles. Such a database does not currently exist, to the best of our knowledge, but is an exciting area for future research.

by the idea that developing countries are revenue constrained and take advantage of opportunities to widen their tax base. In Table A13, we study the heterogeneous impacts of trade on *overall* taxation by development level. We relate all taxes collected by the government (on labor, capital and consumption) to the country's GDP. We again find stark differences by development levels: in developing countries, trade openness leads to a significant increase in overall tax to GDP, while there is a null effect in rich countries.[42] These results are consistent with the interpretation that, in response to increased globalization, governments in developed countries seek to minimize efficiency concerns while maintaining revenue neutrality; developing countries balance efficiency concerns with the need for extra revenue collection.

7.4 Comparison with the Event-Study Estimates

Although the instrumental variable and the event-study estimations (Section 6) rely on entirely different identifying assumptions and methodologies, they yield consistent results. Their results are also comparable in magnitude. Under the assumption that the liberalization events only impacts factor taxation through trade, the event-study results imply a trade impact on ETR_K which equals $\mu^{DiD} = 0.428$ (see Table A3). This estimate is well within the range of IV-based estimates μ^{IV} for capital effective taxation estimated in this section.

These results, combined with the time series patterns (Section 5), show a robust positive impact of trade openness on both the capital share of income and the effective taxation of capital, which particularly affects developing countries.

As a final exercise, we use the IV-results to attempt to quantify trade's role in accounting for the long-run trend of capital taxation observed in developing countries. We combine the IV-coefficient for ETR_K in developing countries with the observed change in trade openness between 1965 and 2018 in those countries. This exercise reveals that the long-run

[42]Studies that investigate whether trade-induced revenue losses at the border are recovered on non-border bases produce mixed results: while Baunsgaard and Keen (2009) and Cage and Gadenne (2018) find incomplete recovery, implying a net revenue loss from trade liberalization, Buettner and Madzharova (2018) find more than complete recovery. Our database does not separate between border and non-border sources of indirect taxation.

increase in cross-border trade leads to a 3.1 percentage point increase in ETR_K, which accounts for 30.5 percent of the long-run rise of effective capital taxation in developing countries.[43] This number should be taken with caution, but suggests that about a third of the rise in effective capital taxation can be accounted for by trade globalization.

8 Conclusion

In this paper, we combine a new global database with several identification strategies to provide novel evidence on trends and causal effects of globalization on tax structures worldwide. Our starting point is the systematic harmonization of novel historical data on national accounts components and disaggregated government revenues. This data collection permitted the construction of a new measures of factor shares and of effective tax rates on capital and labor in 156 countries between 1965 and 2018.

Using this database, we make two contributions. First, we establish new descriptive facts. Taking a global perspective, the average effective tax rates on labor and capital converged, due to a large increase in labor taxation and a small reduction in capital taxation. We find fundamental long-run differences between developed and developing countries: while the effective tax rate on capital mildly declined in OECD countries, it strongly increased in the rest of the world, with a steep acceleration in the post-1995 period of hyper-globalization. Second, using a variety of research designs, we find that increased integration through cross-border flow of goods and services impacts the effective taxation of both labor and capital. In low and middle-income countries our results reject the conjecture that trade openness leads to an erosion of capital taxation.

We observe heterogeneity in long-run factor taxation across countries, and the presence of several, partly countervailing forces, that determine the impact of trade openness on factor taxation. Thus, our paper emphasizes that globalization has nuanced effects on the

[43] Concretely, the long-run increase in trade openness is 7.01 percentage points (Figure A11) and the trade-coefficient for ETR_K is 0.44 (Table 5), hence $7.01 * 0.44 = 3.08ppt$. The long-run increase in ETR_K is 10.1ppt (Figure 3), thus yielding $3.08/10.1 = 0.305$

relative taxation of labor and capital.[44] These nuances are important when considering tax policy's limitations in a globalized world.

[44]This summary is consistent with the conclusion in Swank and Steinmo (2002): "Globalization has mattered but in much more subtle ways than internationalization theory suggests (...) The new reality is an environment where policymakers confront three related constraints: internationalization, domestic economic stress, and budgetary imperatives. Capital mobility has not led to a race to the bottom or to the evisceration of the revenue-raising capacity of the state."

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program". In: *Journal of the American Statistical Association* 105(490), pp. 493–505.

Abadie, Alberto, Jiaying Gu, and Shu Shen (2019). *Instrumental Variable Estimation with First-Stage Heterogeneity*. Tech. rep.

Adam, Antonis et al. (2013). "The effect of globalization on capital taxation: What have we learned after 20 years of empirical studies?" In: *Journal of Macroeconomics* 35, pp. 199–209.

Akcigit, Ufuk et al. (2021). "Taxation and Innovation in the 20th Century." In: *NBER Working Paper Series*(24982).

Alvaredo, Facundo et al. (2020). *Distributional National Accounts Guidelines: Methods and Concepts Used in the World Inequality Database*. URL: https://wid.world/document/distributional-national-accounts-guidelines-2020-concepts-and-methods-used-in-the-world-inequality-database/.

Anderson, James and Eric van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle". In: *American Economic Review* 93(1), pp. 170–92.

Angrist, Joshua and Guido Imbens (1995). "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity". In: *Journal of the American Statistical Association* 90(430), pp. 431–45.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare (2012). "New Trade Models, Same Old Gains?" In: *American Economic Review* 102(1), pp. 94–130.

Basri, Chatib et al. (2019). "Tax Administration vs. Tax Rates: Evidence from Corporate Taxation in Indonesia". In: *NBER Working Paper*(26150).

Bates, Robert, Da-Hsiang, and Donald Lien (1985). "A Note on Taxation, Development and Representative Government". In: *Politics and Society* 14(1), pp. 53–70.

Baunsgaard, Thomas and Mick Keen (2009). "Tax revenue and (or?) trade liberalization". In: *Journal of Public Economics* 94(9), pp. 563–577.

Besley, Timothy and Torsten Persson (2014). "Why Do Developing Countries Tax So Little?" In: *Journal of Economic Perspectives* 28(4), pp. 99–120.

Billmeier, Andreas and Tommaso Nannicini (2013). "Assessing economic liberalization episodes: A synthetic control approach". In: *Review of Economics and Statistics* 95(3), pp. 983–1001. DOI: 10.1162/REST_a_00324.

Blanchet, Thomas and Lucas Chancel (2016). "National Accounts Series Methodology." URL: https://wid.world/document/1676/.

Borusyak, K, X Jaravel, and J Spiess (2021). *Revisiting Event Study Designs: Robust and Efficient Estimation*. Tech. rep. Working Paper.

Brandt, Loren et al. (2017). "WTO Accession and Performance of Chinese Manufacturing Firms". In: *American Economic Review* 107(9), pp. 2784–2820.

Browning, Edgar K. (1978). "The Burden of Taxation." In: *Journal of Political Economy* 86(4), pp. 649–671. DOI: 10.1086/260703.

Buettner, Thiess and Boryana Madzharova (2018). "WTO membership and the shift to consumption taxes". In: *World Development* 108(1), pp. 197–218.

Cage, Julia and Lucie Gadenne (2018). "Tax revenues and the fiscal cost of trade liberalization, 1792-2006". In: *Explorations in Economic History* 70, pp. 1–24.

Cameron, Colin, Jonah Gelbach, and Douglas Miller (2008). "Bootstrap-Based Improvements for Inference with Clustered Standard Errors". In: *Review of Economics and Statistics* 90(3), pp. 414–27.

Carey, David and Josette Rabesona (2004). "Tax Ratios on Labor and Capital Income and on Consumption." In: *Measuring the Tax Burden on Capital and Labor.* Ed. by Peter B. Sørensen. MIT Press. DOI: 10.7551/mitpress/4522.001.0001.

China Daily (2018). *Taxation: Historical development*. URL: https://govt.chinadaily.com.cn/s/201809/20/WS5c08d6fe498eefb3fe46e52c/historical-development.html.

Chinn, Menzie and Hiro Ito (2006). "What Matters for Financial Development? Capital Controls, Institutions, and Interactions". In: *Journal of Development Economics* 81(1), pp. 163–192.

CIAT and IDB (2019). *IDB-CIAT Revenue Collection Database*. URL: https://www.ciat.org/idb-ciat-revenue-collection-database/?lang=en.

Cogneau, Denis, Yannick Dupraz, Justine Knebelmann, et al. (2021). "Taxation in Africa from Colonial Times to Present Evidence from former French colonies 1900-2018". In: *Mimeo*.

Cogneau, Denis, Yannick Dupraz, and Sandrine Mesple-Somps (2021). "Fiscal Capacity and Dualism in Colonial States: The French Empire 1830–1962". In: *The Journal of Economic History* 81(2), pp. 441–480. DOI: 10.1017/S0022050721000140.

Dao, Mai Chi, Zsoka Koczan, and Weicheng Lian (2017). "Why Is Labor Receiving a Smaller Share of Global Income? Theory and Empirical Evidence". In: *IMF Working Papers* 17(169).

Devereux, Michael P. (2004). "Measuring Taxes on Income from Capital." In: *Measuring the Tax Burden on Capital and Labor.* Ed. by Peter B. Sørensen. MIT Press. DOI: 10.7551/mitpress/4522.001.0001.

Dix-Carneiro, Rafael and Brian Kovak (2017). "Trade Liberalization and Regional Dynamics". In: *American Economic Review* 107(10), pp. 2908–46.

Dix-Carneiro, Rafal, Pinelopi Goldberg, et al. (2021). "Trade and Informality in the Presence of Labor Market Frictions and Regulations". In: *NBER Working Paper*(28391).

Eaton, Jonathan and Samuel Kortum (2002). "Technology, Geography, and Trade." In: *Econometrica* 70(5), pp. 1741–79.

Egger, Peter H., Sergey Nigai, and Nora M. Strecker (2019). "The Taxing Deed of Globalization." In: *American Economic Review* 109(2), pp. 353–90. DOI: 10.1257/aer.20160600.

Esberg, Jane and Rebbeca Perlman (2020). "Covert Confiscation: How Regimes Differ in Their Strategies of Expropriation." Mimeo.

Estrin, Saul and Adeline Pelletier (2018). "Privatization in Developing Countries: What Are the Lessons of Recent Experience?" In: *World Bank Research Observer* 33(1), pp. 65–102. DOI: 10.1093/wbro/lkx007.

Feenstra, Robert and Gordon Hanson (2001). "Global Production Sharing and Rising Inequality: A Survey of Trade and Wages". In: *NBER Working Paper Series*(8372).

Fisher-Post, Matthew (2020). "Factor Shares in the Long Run." URL: https://wid.world/document/factor-shares-in-the-long-run-wid-world-working-paper-2020-03/.

Fisunoglu, Ali et al. (2011). *Relative Political Capacity Dataset*. Version 2.4 (August 2020). DOI: 10.7910/DVN/NRR7MB.

Frankel, Jeffrey and David Romer (1999). "Does Trade Cause Growth?" In: *American Economic Review* 89(13), pp. 379–399.

Goldberg, Pinelopi and Nina Pavcnik (2007). "Distributional Effects of Globalization in Developing Countries." In: *Journal of Economic Literature* 45(1), pp. 39–82.

Goldberg, Pinelopi and Nina Pavcnik (2016). "The Effects of Trade Policy". In: *NBER Working Paper Series*(21957).

Gollin, Douglas (2002). "Getting Income Shares Right". In: *Journal of Political Economy* 110(2). DOI: 10.1086/338747.

Gross, Wendelin, Cristina Hayden, and Christian Butz (2012). "About the impact of rising oil price on logistics networks and transportation greenhouse gas emission". In: *Logistics Research* 4, pp. 147–56.

Gruber, Jonathan, Amalie Jensen, and Henrik Kleven (2021). "Do people respond to the mortgage interest deduction? Quasi-experimental evidence from Denmark". In: *American Economic Journal: Economic Policy* 13(2), pp. 273–303.

Guerriero, Marta (2019). "The Labor Share of Income around the World: Evidence from a Panel Dataset." URL: https://www.adb.org/publications/labor-share-income-around-world-evidence-panel-dataset.

Guriev, Sergei, Anton Kolotilin, and Konstantin Sonin (2011). "Determinants of Nationalization in the Oil Sector: A Theory and Evidence from Panel Data." In: *The Journal of Law, Economics, and Organization* 27(2), pp. 301–323. DOI: 10.1093/jleo/ewp011.

Haggarty, Luke and Mary Shirley (1997). "A New Database on State-Owned Enterprises." In: *The World Bank Economic Review* 11(3), pp. 491–513. DOI: 10.1093/wber/11.3.491.

Harrison, Ann (2005). *Has Globalization Eroded Labor's Share? Some Cross-Country Evidence*. Working Paper.

Havnes, Tarjei and Magne Mogstad (2015). "Is universal child care leveling the playing field?" In: *Journal of Public Economics* 127, pp. 100–114. DOI: 10.1016/j.jpubeco.2014.04.007.

ICTD/UNU-WIDER (2020). *Government Revenue Dataset*. URL: https://www.wider.unu.edu/project/government-revenue-dataset.

ILO (2019). "The Global Labour Income Share and Distribution." URL: https://www.ilo.org/ilostat-files/Documents/Labour%5C%20income%5C%20share%5C%20and%5C%20distribution.pdf.

ILOSTAT (2021). *Statistics on labour income and inequality.* URL: https://ilostat.ilo.org/topics/labour-income/.

IMF GFS (2005). *Historical Government Finance Statistics (1972-1989 in GFSM 1986 format)*. CD-ROM. URL: https://www.imf.org/external/pubs/cat/longres.aspx?sk=18674.0.

Jaeger, Simon, Shakked Noy, and Benjamin Schoefer (2021). "What Does Co-Determination Do?" In: *NBER Working Paper Series*(28291).

Jefferson, Gary (2016). "State-Owned Enterprise in China: Reform, Performance, and Prospects". In: *Working Paper*.

Jensen, Anders (2021). "Employment Structure and the Rise of the Modern Tax System". In: *American Economic Review* Forthcoming.

Kanbur, Ravi and Michael Keen (1993). "Jeux Sans Frontier: Tax Competition and Tax Coordination when Countries Differ in Size". In: *American Economic Review* 83(1), pp. 877–92.

Karabarbounis, Loukas and Brent Neiman (2014). "The Global Decline of the Labor Share." In: *Quarterly Journal of Economics* 129(1), pp. 61–103. DOI: 10.1093/qje/qjt032.

Kawano, Laura and Joel Slemrod (2016). "How do corporate tax bases change when corporate tax rates change? With implications for the tax rate elasticity of corporate tax revenues". In: *International Tax and Public Finance* 23(1), pp. 401–433.

King, Mervyn A. and Don Fullerton (1984). "The Theoretical Framework." In: *The Taxation of Income from Capital: A Comparative Study of the United States, the United Kingdom, Sweden, and Germany*. National Bureau of Economic Research, pp. 7–30. ISBN: 9780226436319. DOI: 10.7208/chicago/9780226436319.001.0001.

Kleven, Henrik, Martin Knudsen, et al. (2011). "Unable or Unwilling to Cheat? Evidence from a Tax Audit Experiment in Denmark". In: *Econometrica* 79(3), pp. 651–692.

Kleven, Henrik, Claus Kreiner, and Emmanuel Saez (2016). "Why Can Modern Governments Tax so Much? An Agency Model of Firms as Fiscal Intermediaries". In: *Economica* 83(330), pp. 219–246.

Lequiller, François and Derek Blades (2014). *Understanding National Accounts*. 2nd ed. OECD. DOI: 10.1787/9789264214637-en.

Lotz, Joergen R and Elliott R Morss (1970). "A theory of tax level determinants for developing countries". In: *Economic Development and cultural change* 18(3), pp. 328–341.

Mahdavi, Paasha (2020). *Power Grab: Political Survival through Extractive Resource Nationalization*. Cambridge University Press. DOI: 10.1017/9781108781350.

McCaig, Brian and Nina Pavcnik (2018). "Export Markets and Labor Allocation in a Low-Income Country". In: *American Economic Review* 108(7), pp. 1899–1941.

McDaniel, Cara (2007). "Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003". In:

Mendoza, Enrique G., Assaf Razin, and Linda L. Tesar (1994). "Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption." In: *Journal of Monetary Economics* 34(3), pp. 297–323. DOI: 10.1016/0304-3932(94)90021-3.

Mogstad, Magne, Alexander Torgovitsky, and Christopher Walter (2020). "The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables". In: *NBER Working Papers*(25691).

Naritomi, Joana (2019). "Consumers as Tax Auditors". In: *American Economic Review* 109(9), pp. 3031–72.

OECD (2020). *Revenue Statistics*. URL: https://www.oecd.org/tax/revenue-statistics-2522770x.htm.

OECD-UCLG (2019). *World Observatory on Subnational Government Finance and Investment*. URL: http://www.sng-wofi.org/data/.

OPEC (2021). *Average annual OPEC crude oil price from 1960 to 2021*. URL: https://www.statista.com/topics/1830/opec/.

Pomeranz, Dina (2015). "No Taxation without Information: Deterrence and Self-Enforcement in the Value-Added Tax". In: *American Economic Review* 105(8), pp. 2539–69.

Rodriguez, Francisco and Daniel Ortega (2006). "Are Capital Shares Higher in Poor Countries? Evidence from Industrial Surveys". In: *Working Paper*.

Rodriguez, Francisco and Dani Rodrik (2001). "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence". In: *NBER Macroeconomic Annual* 15, pp. 261–338.

Rodrik, Dani (1997). "Trade, Social Insurance, and the Limits to Globalization". In: *NBER Working Paper Series*(5905).

Rodrik, Dani (1998). "Why do More Open Economies Have Bigger Governments?" In: *Journal of Political Economy* 106(5), pp. 997–1032. doi: 10.1086/250038.

Romer, Christina and David Romer (2010). "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks". In: *American Economic Review* 100(1), pp. 763–801.

Ross, Michael and Paasha Mahdavi (2015). "Oil and gas data, 1932–2014". In: *Harvard Dataverse* 2.

Saez, Emmanuel and Gabriel Zucman (2019). *The Triumph of Injustice: How the Rich Dodge Taxes and How to Make Them Pay*. W.W.Norton. isbn: 9780393531732. url: https://taxjusticenow.org/#/.

SeaRates (2021). *Online Freight Shipping and Transit Time Calculator.* url: https://www.searates.com/services/distances-time/.

Slemrod, Joel (2004). "Are Corporate Tax Rates, Or Countries, Converging?" In: *Journal of Public Economics* 88(6), pp. 1169–1186.

Smith, Brock (2015). "The resource curse exorcised: Evidence from a panel of countries". In: *Journal of Development Economics* 116, pp. 57–73. doi: 10.1016/j.jdeveco.2015.04.001.

Storeygard, Adam (2016). "Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa". In: *Review of Economic Studies* 83(1), pp. 1263–95.

Swank, Duane and Sven Steinmo (2002). "The New Political Economy of Taxation in Advanced Capitalist Democracries". In: *American Journal of Political Science* 46(3), pp. 642–655.

Topalova, Petia and Amit Khandelwal (2011). "Trade Liberalization and Firm Productivity: The Case of India". In: *Review of Economics and Statistics* 93(3), pp. 995–1009.

UN (2018). *Handbook on Backcasting*. url: https://unstats.un.org/unsd/nationalaccount/aeg/2018/M12_8iiib_Backcasting.pdf.

UN COMTRADE (2020). *UN Comtrade Database*. URL: https://comtrade.un.org/data/.

UN DESA (2019). *World Urbanization Prospects: The 2018 Revision.* United Nations, Department of Economic and Social Affairs, Population Division.

UN SNA (1968). *A System of National Accounts*. URL: https://unstats.un.org/unsd/nationalaccount/docs/1968SNA.pdf.

UN SNA (2008). *A System of National Accounts*. URL: https://unstats.un.org/unsd/nationalaccount/docs/sna2008.pdf.

UNSD (2018). *Handbook on Backcasting*. URL: https://unstats.un.org/unsd/nationalaccount/aeg/2018/M12_8iiib_Backcasting.pdf.

Wacziarg, Romain (2001). "Measuring the Dynamic Gains from Trade". In: *World Bank Economic Review* 15(3), pp. 393–429.

Wacziarg, Romain and Jessica Wallack (2004). "Trade Liberalization and Intersectoral Labor Movements". In: *Journal of International Economics* 64, pp. 411–439.

Wacziarg, Romain and Karen Horn Welch (2008). "Trade Liberalization and Growth: New Evidence". In: *World Bank Economic Review* 22(2), pp. 187–231.

WID (2020). *World Inequality Database*. URL: http://wid.world.

Wilson, John (1999). "Theories of Tax Competition". In: *National Tax Journal* 52(2), pp. 269–304.

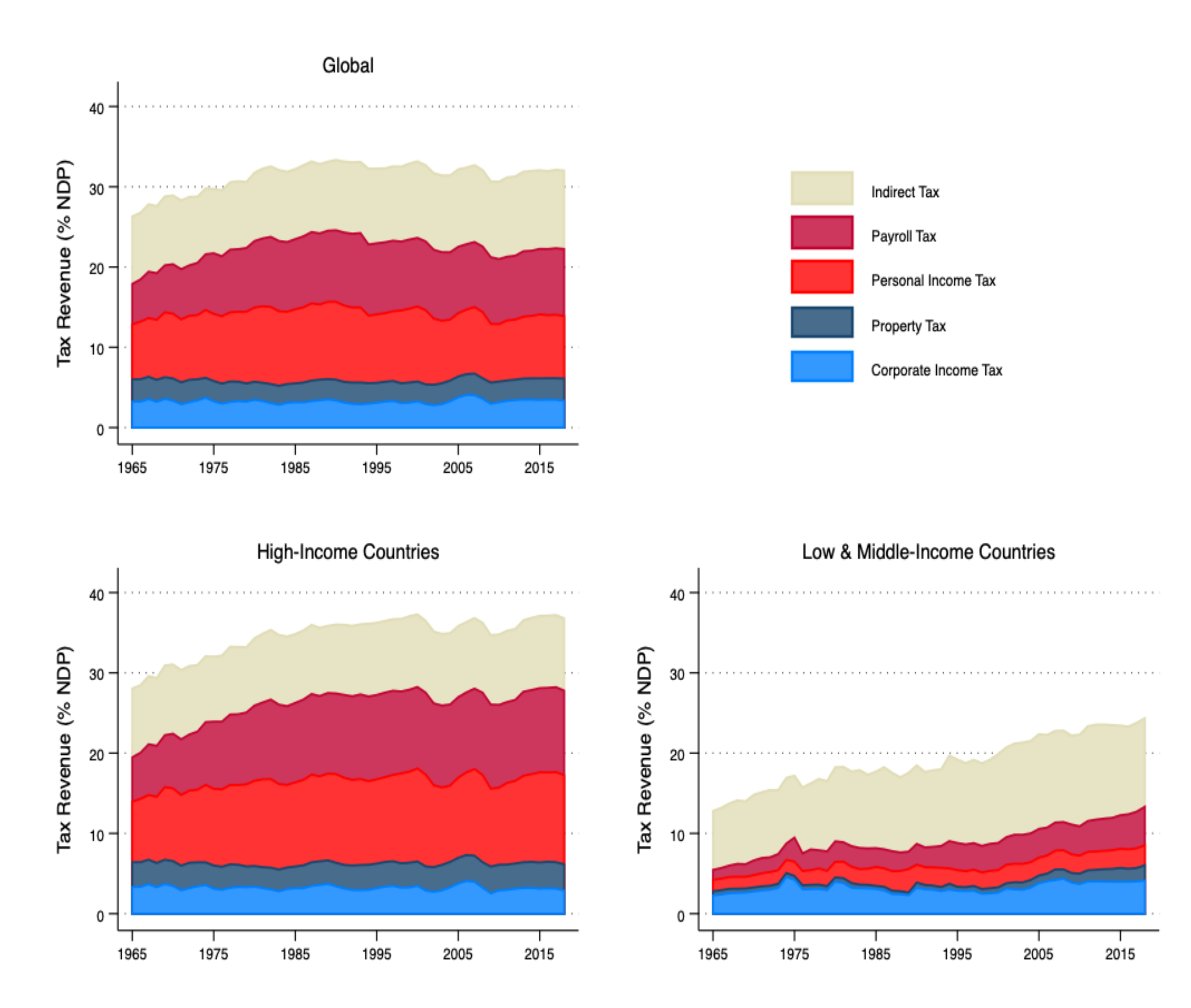
Wong, Christine P. W. and Richard M. Bird (2008). "China's Fiscal System: A Work in Progress". In: *China's Great Economic Transformation*. Ed. by Loren Brandt and Thomas G. Rawski. Cambridge University Press, pp. 429–466. DOI: 10.1017/CBO9780511754234.013.

World Bank (2008). *Public Finance in China: Reform and Growth for a Harmonious Society*. Ed. by Jiwei Lou and Shuilin Wang. The World Bank.

Zidar, Owen (2019). "Tax Cuts for Whom? Heterogeneous Effects of Income Tax Changes on Growth and Employment". In: *Journal of Political Economy* 127(3), pp. 1437–1472.

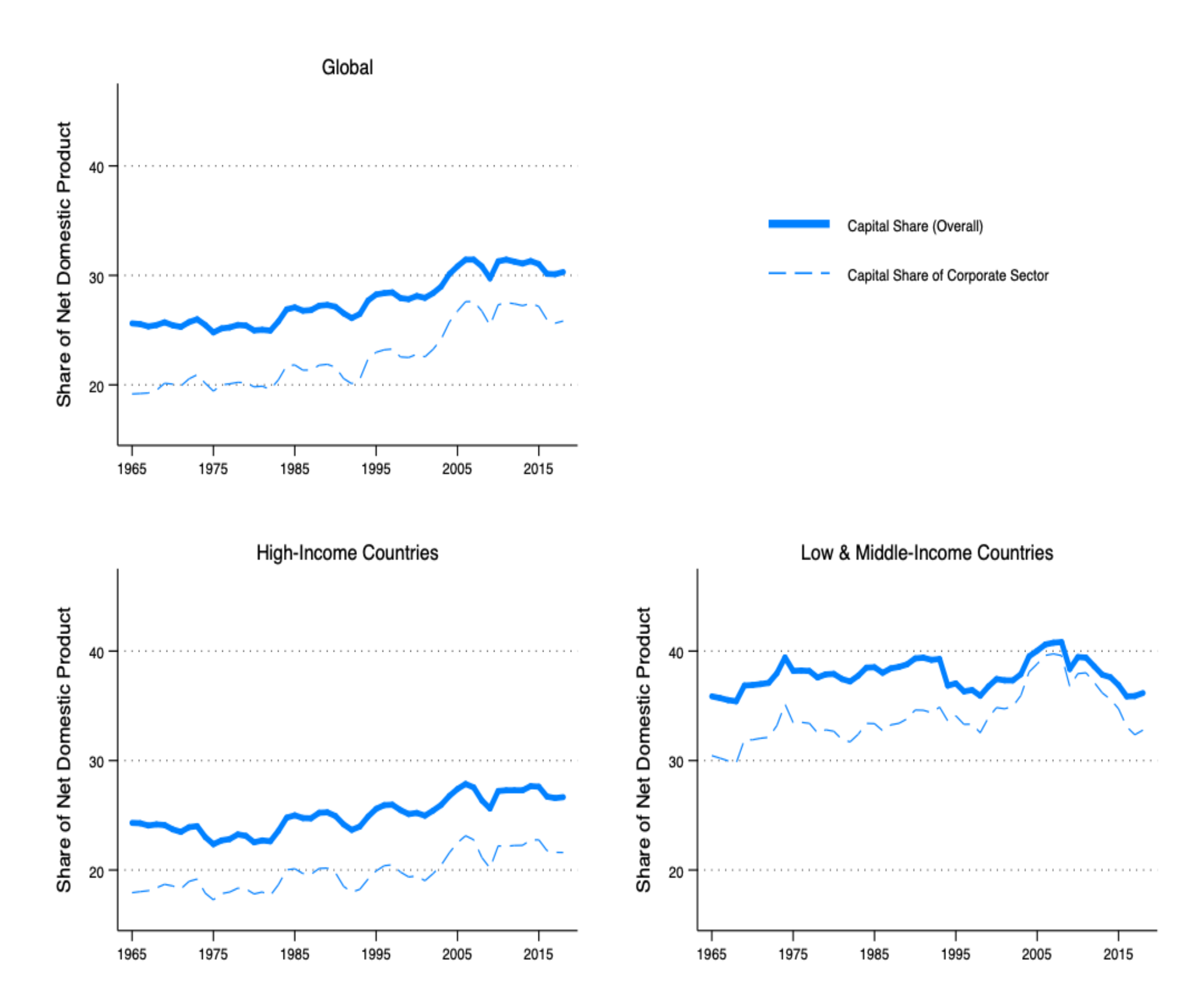
Zucman, Gabriel (2015). *The Hidden Wealth of Nations: The Scourge of Tax Havens*. University of Chicago Press. ISBN: 9780226422640. DOI: 10.7208/chicago/9780226245560.001.0001.

Figure 1: Tax Revenue as a Share of Domestic Product



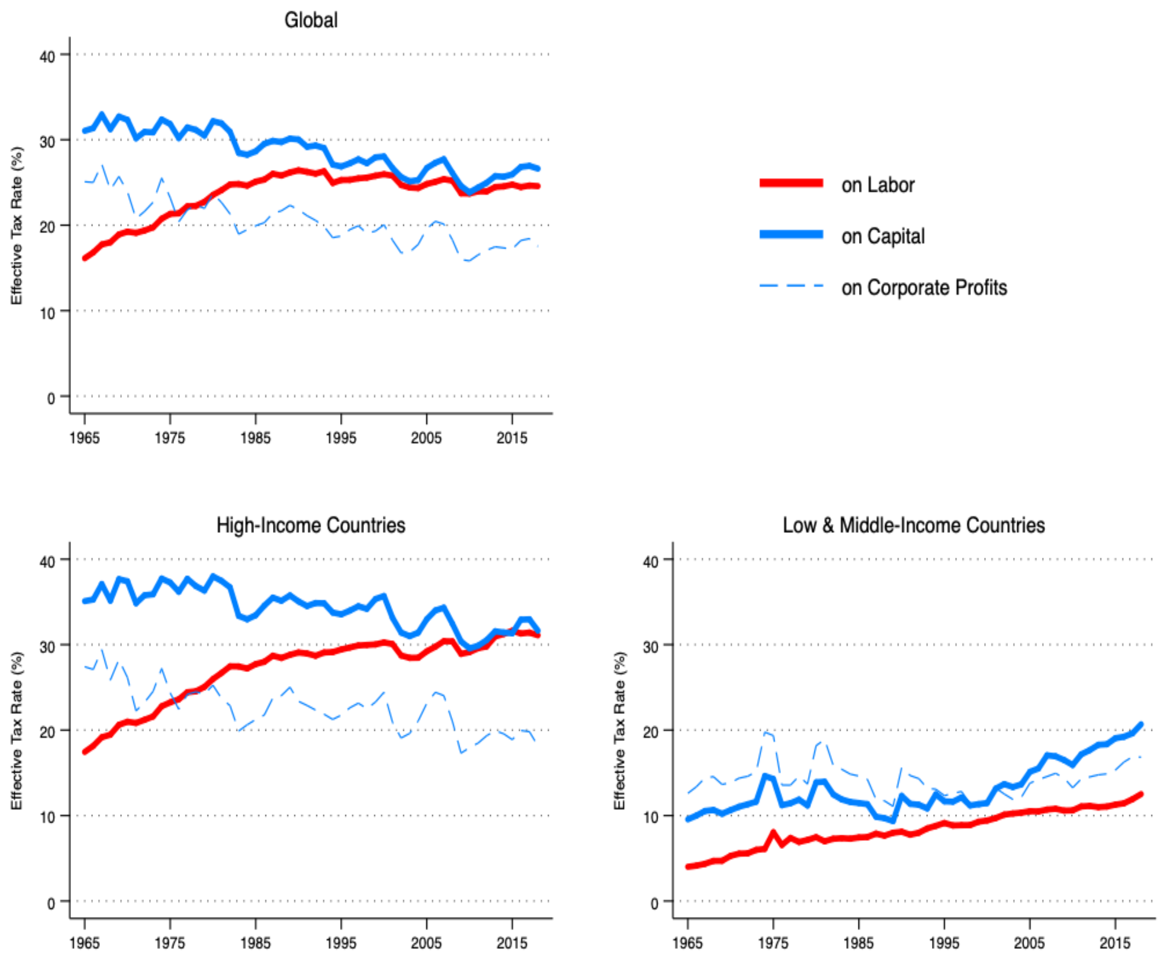
Notes: This figure plots the time series of tax revenue as a share of net domestic product (NDP), separated into five revenue sources. The top left panel corresponds to the global average, weighting country-year observations by their share in that year's total NDP, in constant 2019 USD (N=156). The bottom-left panel shows the results for high-income OECD countries (N=37), and the bottom right for low- and middle-income countries (N=119). We consider as high-income, all OECD countries that meet the World Bank's classification of high-income. Tax revenues are separated into five main categories: indirect taxes (including domestic consumption taxes, excises, and tariffs), payroll 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. It accounts for 85-90% of World GDP during those years. The second, covers 1994-2018 and integrates former communist countries, and in particular China and Russia, and accounts for 98% of World GDP.

Figure 2: Capital Share of Domestic Product



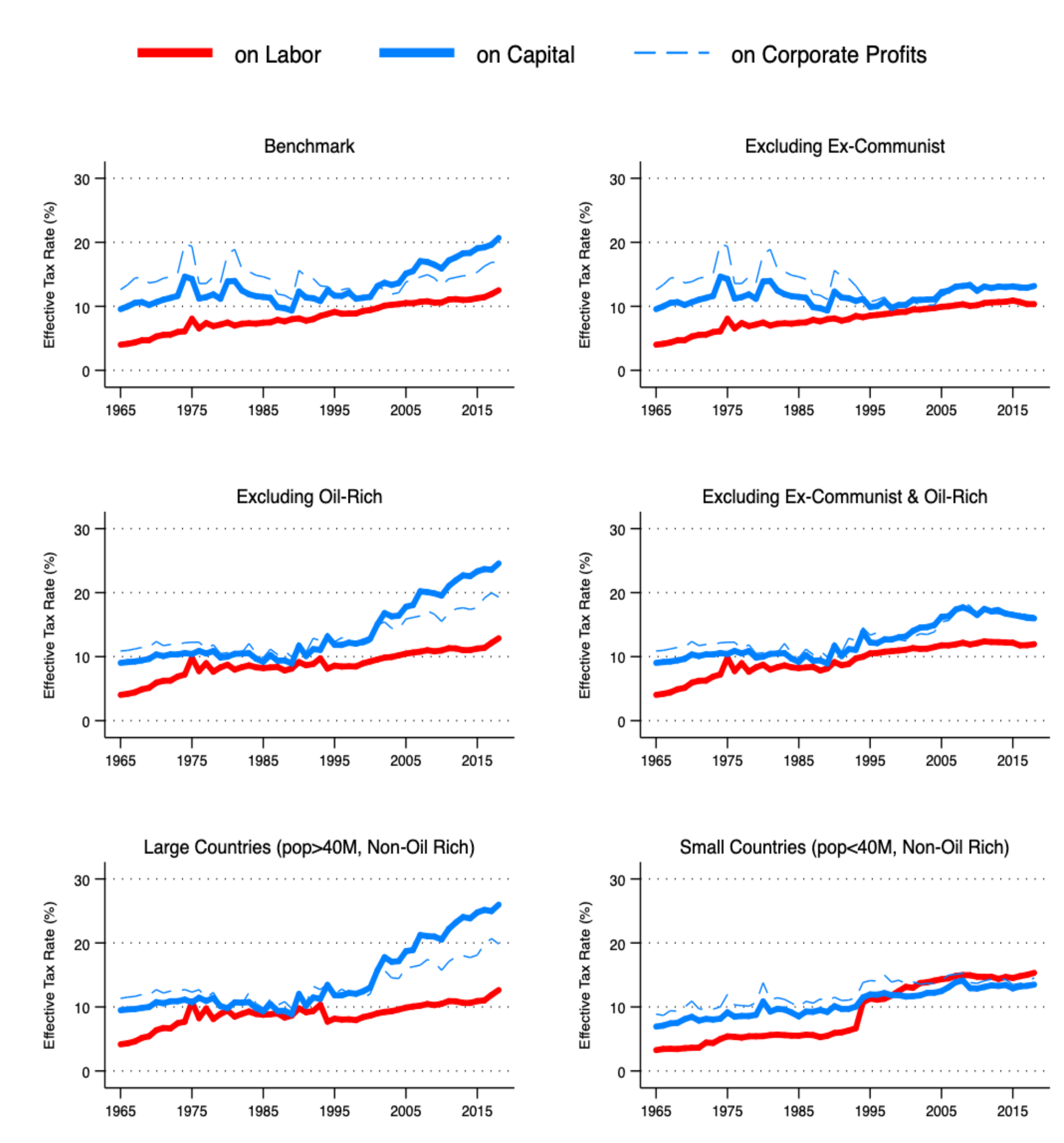
Notes: This figure plots the time series of the capital share as a percentage of net domestic product (NDP). The solid line corresponds to the overall capital share, and the dotted line to the capital share within the corporate sector. The top left panel corresponds to the global average, weighting country-year observations by their share in that year's total NDP, in constant 2019 USD (N=156). The bottom-left panel shows the results for high-income OECD countries (N=37), and the bottom right for low- and middle-income countries (N=119). We consider as high-income, all OECD countries that meet the World Bank's classification of high-income. 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, and in particular China and Russia, and accounts for 98% of World GDP.

Figure 3: Effective Taxation of Capital and Labor



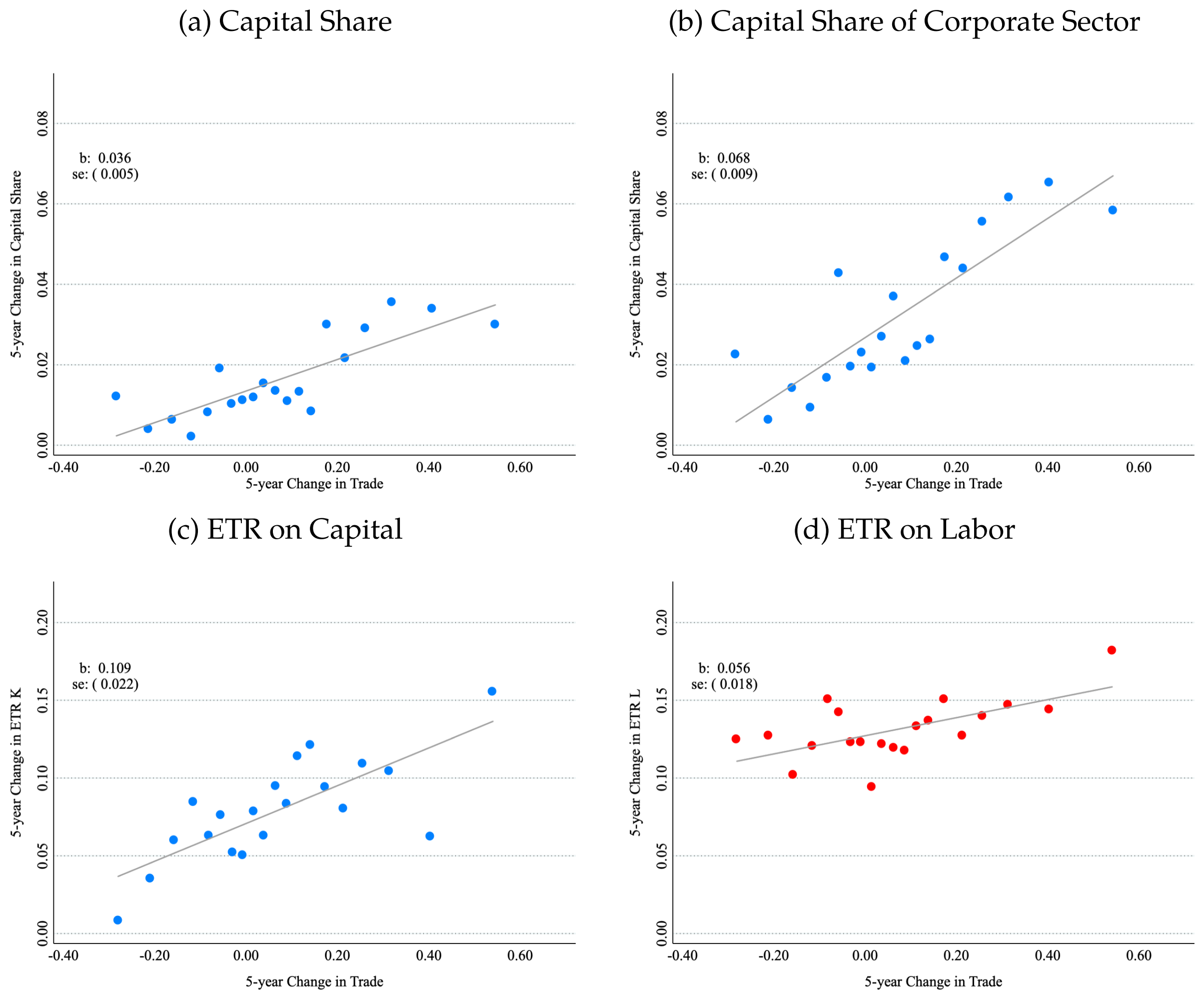
Notes: This figure plots the time series of average effective tax rates on labor (blue) and capital (red), as well as the effective tax rate on corporate profits (red dashed line). The top-left panel corresponds to the global average, weighting country-year observations by their share in that year's total NDP, in constant 2019 USD (N=156). The bottom-left panel shows the results for high-income OECD countries (N=37), and the bottom-right panel for low- and middle-income countries (N=119). High-income countries are OECD countries that meet the World Bank's income threshold of high-income. 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, and in particular China and Russia, and accounts for 98% of World GDP.

Figure 4: Effective Taxation of Capital and Labor in Low- and Middle-Income Countries



Notes: This figure plots the time series of average effective tax rates on labor and capital, as well as on corporate profits , in the 118 low- and middle-income countries only. It thus excludes OECD countries classified as high-income by the World Bank. The top-left panel is our benchmark result, taken from 3. The top-right panel excludes former communist countries, most notably China and Russia. The mid-left panel excludes oil-rich countries (the 33 countries where average oil production since 1990 has exceeded 6.5% of GDP, per Ross and Mahdavi (2015)). The mid-right panel excludes both ex-communist and oil-rich nations. Finally the bottom panels show the results separately for the 18 largest (non-oil rich) countries to the left, and the 68 small (non-oil rich) countries to the right. Large (small) countries are defined as having a population above (below) 40 Million in 2018.

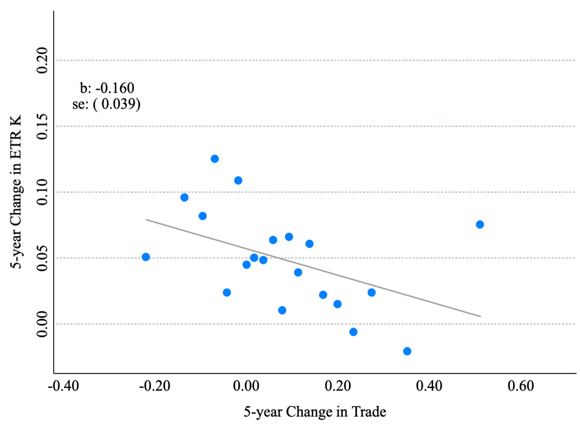
Figure 5: Change in Capital Shares and Factor Taxation vs. Change in Trade



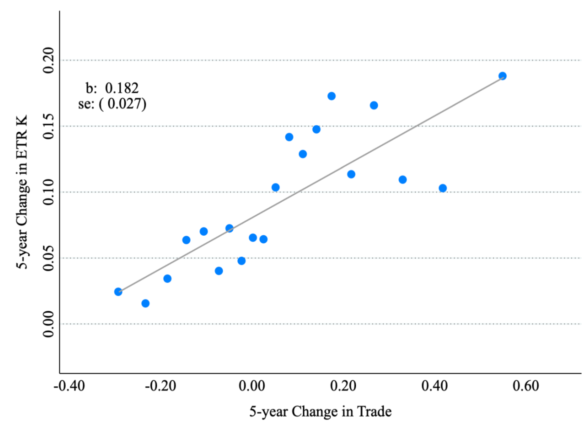
Notes: These panels show the relationship between trade and the capital share of domestic product (a); the capital share of the corporate sector (b); the effective tax rate on capital income (c); and the effective tax rate on labor income (d). Trade is measured as a percentage of net domestic product. 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. For ease of interpretation, we add back the (non-residualized) mean of the given variable. Linear trend lines are unweighted, with year fixed effects, and are estimated based on the underlying country-year panel data. The corresponding slope and standard error are shown top-left in each panel.

Figure 6: Change in Factor Taxation vs. Change in Trade, by Income Level

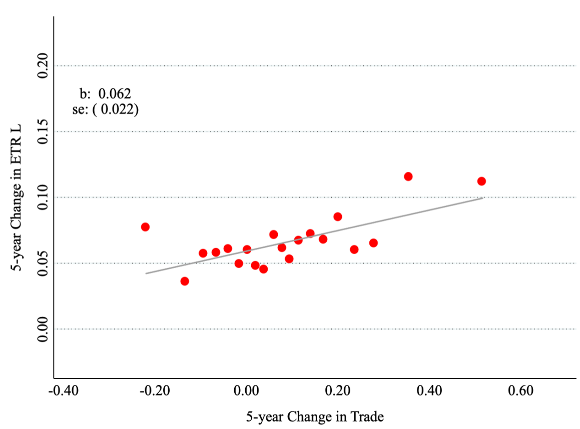
(a) ETR K: High-Income



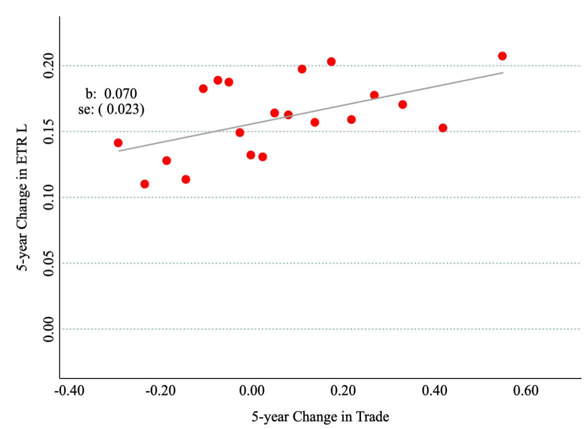
(b) ETR K: Low & Middle-Income



(c) ETR L: High-Income

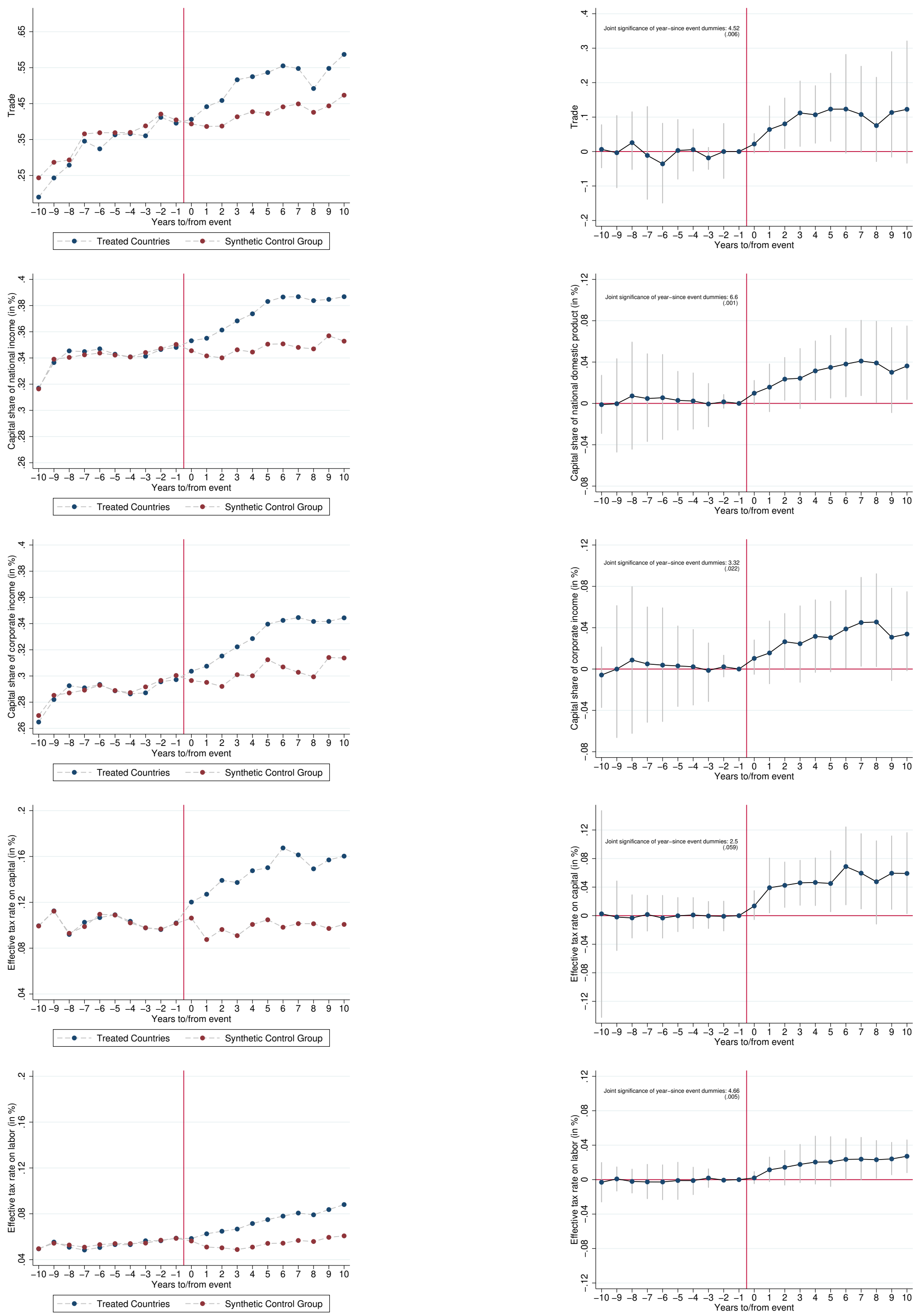


(d) ETR L: Low & Middle-Income



Notes: These panels show the association between changes in trade and changes and changes in effective tax rates of capital (panels a and b) and labor (panels c and d), respectively for high income OECD countries and for low and middle income countries. Trade is measured as a percentage of net domestic product. Both the x-axis and y-axis are measured as within-country percent changes over 5 years. Each graph plots binned scatter plots of the 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. For ease of interpretation, we add back the (non-residualized) mean of the given variable. Linear trend lines are unweighted, with year fixed effects, and are estimated based on the underlying country-year panel data. The corresponding slope and standard error are shown top-left in each panel.

Figure 7: Event Study on Tariff-Based Trade Liberalization Reforms



Notes: These figures show event-studies for trade liberalization reforms in seven countries, over five outcomes: trade; capital shares (overall and within the corporate sector); and effective tax rates (on capital and labor). The left-hand 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 controls. The right-hand graphs show the coefficients on the 'to' ('since') dummies, in a regression model with country fixed effects, year 'to' ('since') fixed effects, and calendar year fixed effects. The bars represent the 95% confidence intervals for 'to' ('since') reform coefficients, while standard errors are clustered at the country-reform level and estimated using the wild bootstrap method. In the top-left corner, we report the F-statistic on joint significance of the post0reform dummies, with the p-value in parentheses below. Details on methodology in Section 6.1 and Appendix C.2.

Table 1: Allocation of Taxes to Factor Incomes, by Type of Tax (Benchmark Estimates)

Type of tax, τ	Series	Allocation, λ_τ	Notes
Panel A: Direct Taxes			
Personal income tax (PIT)	1100	see 3.1.2	most PIT attributed to L
Corporate income tax (CIT)	1200	0%	all CIT attributed to K
Other (unallocable) income tax	1300	50% PIT	rare; small magnitude
Social security & payroll taxes	2000	100%	all payroll taxes attributed to L
Property & wealth taxes	4000	0%	all asset taxes attributed to K
Panel B: Indirect Taxes & Other Revenues			
Indirect taxes	5000	—	assumed proportional to θ_L
Other taxes	6000	—	minor; assumed proportional
Non-tax revenues	7000	—	excluded

Notes: This table shows our benchmark assignment of statutory tax incidence λ_τ on labor (where the incidence on capital is $1 - \lambda_\tau$), for each of the types of taxes in our modified OECD (2020) classification. For the purposes of assigning tax incidence onto factor incomes (see 3.1), we consider here only direct taxes, and implicitly assume that indirect taxation falls proportionally onto labor and capital factor incomes (cf. Browning, 1978; Saez and Zucman, 2019). We treat 'other taxes' similarly (these are rare and insignificant), and ignore non-tax revenue. For income tax revenues whose provenance cannot be understood as either personal income tax (PIT) or corporate income tax (CIT), we assign them as a 50-50% split between the two; these 'unallocable' income tax revenues are rare in occurrence and small in magnitude. Taxes in the 4000 series (largely property taxes) also include wealth and financial transaction taxes.

Table 2: Principal Data Sources

	country-year obs.	%
Panel A: Factor Share Data		
SNA2008	2403	34.8%
SNA1968	1484	21.5%
composite/imputed	3016	43.7%
N	*6903*	*100%*
Panel B: Tax Revenue Data		
OECD	2881	41.7%
Harvard/archives	2092	30.3%
ICTD	1276	18.5%
IMF historical	654	9.5%
N	*6903*	*100%*

Notes: For the N=6903 country-year observations in which we estimate effective tax rates on capital and labor income (over 156 countries since 1965), Panel A presents the sources of our factor share data (on national income components), while Panel B presents the sources our tax revenue data (on total revenues disaggregated by type of tax). In the former, we use online data from UN SNA (2008) and archival data from UN SNA (1968). In the latter, we draw tax revenue data from sources including OECD (2020), ICTD/UNU-WIDER (2020), and IMF GFS (2005), as well as extensive archival research in the Harvard University Library and online. (See Appendix A.)

Table 3: Trade Impacts on Factor Shares and Factor Taxation

	Capital Share		Effective Tax Rate	
	overall	corp. sector	on capital	on labor
Panel A: OLS				
Trade	0.0195*	0.0217	0.0168	0.0246**
	(0.0109)	(0.0148)	(0.0302)	(0.0101)
Panel B: IV estimate (NDP-weighted)				
Trade	0.151**	0.184**	0.375*	0.163***
	(0.0698)	(0.0800)	(0.213)	(0.0538)
First-stage F-statistic	26.07	26.07	26.07	26.07
Panel C: IV estimate (unweighted)				
Trade	0.118*	0.122	0.250**	0.133**
	(0.0681)	(0.0826)	(0.105)	(0.0526)
First-stage F-statistic	8.415	8.415	8.415	8.415
Panel D: IV estimate (NDP-weighted, with controls)				
Trade	0.115**	0.142**	0.400***	0.226***
	(0.0475)	(0.0546)	(0.112)	(0.0551)
First-stage F-statistic	19.02	19.02	19.02	19.02
N	*4518*	*4518*	*4518*	*4518*

Notes: This table presents results from estimating the effect of trade on factor shares and factor taxation. In Panel A, we present results from estimating equation (6) using OLS, while Panels B and C and D present IV estimates—weighted by NDP; unweighted; and weighted with controls, respectively. Across columns, the outcome is the capital share of national domestic product and within the corporate sector, and the effective tax rate on capital and on labor. Trade is measured as the sum of export and imports divided by net domestic product. In Panels B through D, we instrument for trade using the oil-price and the gravity-instruments from Egger, Nigai, and Strecker (2019). All estimates include country and year fixed effects and observations are weighted by net domestic product in constant 2019 USD at PPP (except in Panel C). We use dummy variable controls for significantly interpolated revenue data (rare) or imputed factor share data (frequent). For more details, see Section 7. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table 4: Mechanisms

	OLS	IV		
		Benchmark	Unweighted	Controls
	(1)	(2)	(3)	(4)
Panel A: Corporate income tax rate				
CIT rate	-0.002	-0.064***	-0.051*	-0.061***
	(0.003)	(0.017)	(0.028)	(0.017)
Panel B: Self-Employed as a Share of the Workforce				
Self-employment	-0.0117	-0.220*	-0.185***	-0.174***
	(0.0145)	(0.126)	(0.0460)	(0.0560)
Panel C: National Income Components				
Corporate profits	0.0339***	0.175**	0.124***	0.206***
	(0.0128)	(0.0767)	(0.0321)	(0.0726)
Employee compensation	0.00848	-0.0749	-0.0964	0.0485
	(0.0175)	(0.0904)	(0.0669)	(0.0785)
Mixed income	-0.0231	-0.0685	-0.0391	-0.202**
	(0.0182)	(0.105)	(0.0301)	(0.0816)
Panel D: Effective Tax Rate on Corporate Profits				
ETR_{cit}	0.0120	0.342***	0.359***	0.205
	(0.0220)	(0.121)	(0.0870)	(0.129)

Notes: This table reports estimates the impacts of trade on additional outcomes. Each cell corresponds to a coefficient on trade from a regression model which varies in the outcome (across rows) and estimation model (across columns). Across columns, the coefficients are based on estimating equation 6 using, respectively: OLS; IV; IV with controls; and, IV without weights. The controls included in column (3) are: USD exchange rate; gross fixed capital formation (as a percentage of NDP); (log) population; (log) gross domestic product per capita; and *de jure* capital accounts mobility. In addition, all regressions are weighted by annual net domestic product in constant 2019 USD at PPP, except in column (4) where the weights are removed. Across panels, the outcome is: the statutory corporate income tax rate (Panel A); the self-employed share of the active workforce (Panel B); the share in national income of corporate profits, employee compensation, and mixed income (Panel C); and, the effective tax rate on corporate profits (Panel D). Trade is measured as the sum of export and imports divided by net domestic product. We instrument for trade using the oil-price and the gravity-instruments from Egger, Nigai, and Strecker (2019). All estimates include country and year fixed effects, as well as dummy variable controls for significantly interpolated revenue data (rare) or imputed factor share data (frequent). For more details, see Section 7. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table 5: Heterogeneous Impacts of Trade by Development Level

	ETR_K	ETR_L	CIT-rate	K-share	SE-share	Corp-Share
	(1)	(2)	(3)	(4)	(5)	(6)
Trade	0.444**	0.145	-0.043*	0.182**	-0.252**	0.219***
	(0.181)	(0.093)	(0.024)	(0.077)	(0.107)	(0.063)
Trade$*\mathbb{1}$(High-income)	-0.441	0.120	-0.032	-0.219	0.232	-0.298
	(0.347)	(0.194)	(0.047)	(0.137)	(0.209)	(0.143)
Trade in High-income	0.003	0.265**	-0.075***	-0.036	-0.021	-0.079
	(0.231)	(0.122)	(0.457)	(0.083)	(0.151)	(0.102)
N	*4518*	*4518*	*3810*	*4518*	*4518*	*4518*

Notes: This table presents results from the heterogeneous IV analysis based on estimating equation (7). The top row denotes the outcome variable: effective tax rate on capital; effective tax rate on labor; statutory corporate income tax rate; capital share of domestic product; self-employed share of workforce; corporate share of domestic product. The regression coefficients for $Trade$ as well as the interaction with a dummy for high-income countries, $Trade * \mathbb{1}(High - income)$ are presented. The bottom row reports the coefficient for the linear combination of $Trade$ and the interaction term. Trade is measured as the sum of export and imports divided by net domestic product. We instrument for trade using the oil-price and the gravity-instruments from Egger, Nigai, and Strecker (2019). All estimates include country and year fixed effects and observations are weighted by net domestic product in constant 2019 USD at PPP. We use dummy variable controls for significantly interpolated revenue data (rare) or imputed factor share data (frequent) in all columns, except in column (3) where we use dummies for the data-source of the tax rate. For more details, see Section 7. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table 6: Additional Heterogeneous Impacts of Trade

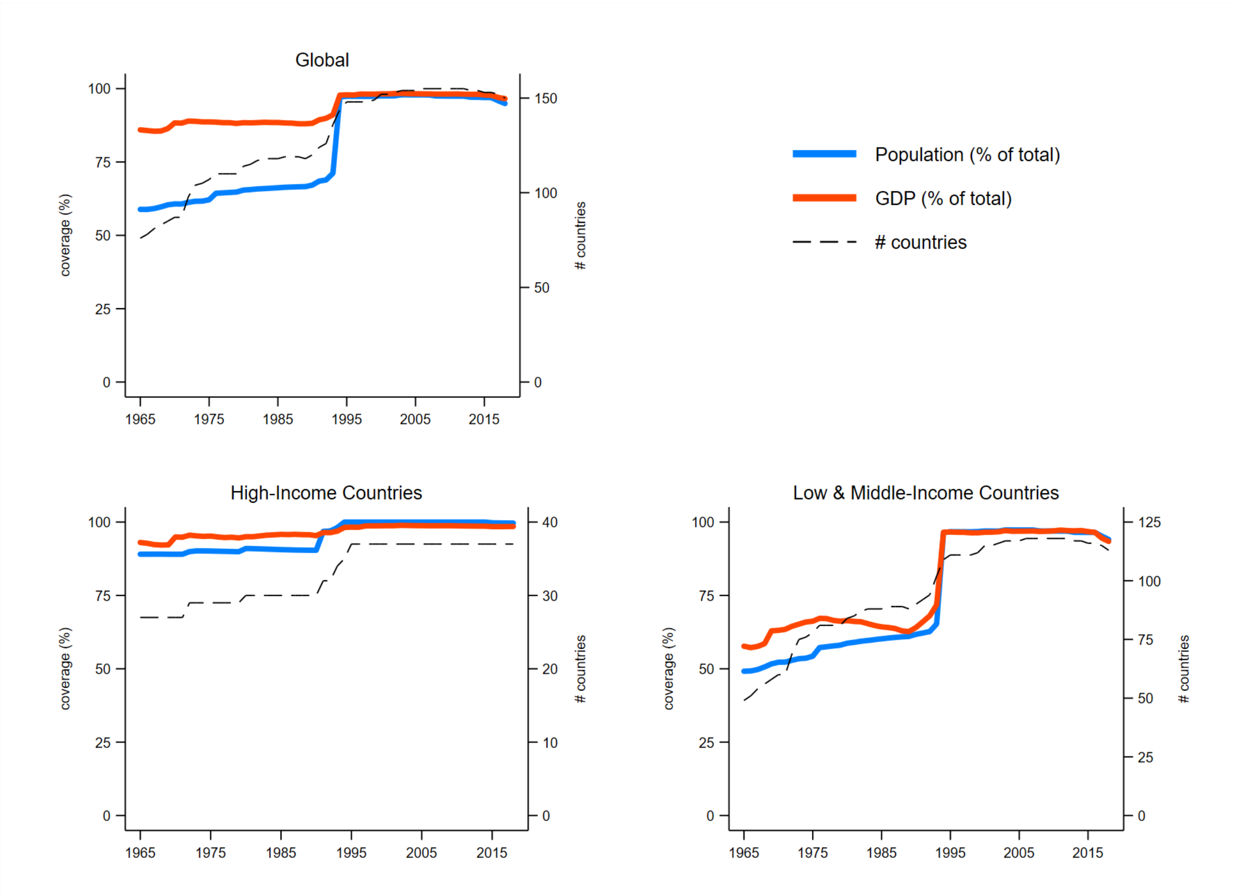
Heterogeneity H_c :	$\mathbb{1}$(Small population)	Capital openness	$\mathbb{1}$(Post-1995)
Panel A: CIT rate			
Trade	-0.053***	-0.063***	-0.044*
	(0.014)	(0.018)	(0.024)
Trade$*H_c$	-0.034	-0.034	-0.014
	(0.054)	(0.079)	(0.022)
Coefficient on Trade in H_c	-0.088*	-0.094	-0.058***
	(0.049)	(0.072)	(0.012)
Panel B: ETR_K			
Trade	0.357**	0.617**	0.504*
	(0.177)	(0.274)	(0.279)
Trade$*H_c$	-0.491	-0.483	-0.218
	(0.544)	(0.456)	(0.207)
Coefficient on Trade in H_c	-0.134	0.133	0.285**
	(0.456)	(0.224)	(0.131)
Panel C: ETR_L			
Trade	0.169***	0.144	0.124
	(0.061)	(0.158)	(0.115)
Trade$*H_c$	0.145	0.159	0.059
	(0.282)	(0.275)	(0.098)
Coefficient on Trade in H_c	0.314	0.304**	0.183***
	(0.242)	(0.139)	(0.044)

Notes: This table presents results from the heterogeneous IV analysis based on estimating equation (7). The top row denotes the source of heterogeneity H_c, respectively across columns: a dummy for small population size (below 40 million); the Chinn-Ito index of capital account openness (Chinn and Ito, 2006), which is a continuous variable between 0 and 1; and, a dummy indicator for the post-1995 period. Across Panels, we estimate the effects of trade on the statutory corporate income tax rate (Panel A), the effective tax rate on capital (Panel B), and the effective tax rate of labor (Panel C). At the bottom of each panel, we report the coefficient on trade (and standard error) in the heterogeneity sub-sample as the linear combination of the coefficients on $Trade$ and $Trade * H_c$. Trade is measured as the sum of export and imports divided by net domestic product. We instrument for trade using the oil-price and the gravity-instruments from Egger, Nigai, and Strecker (2019). All estimates include country and year fixed effects and observations are weighted by net domestic product in constant 2019 USD at PPP. We use dummy variable controls for significantly interpolated revenue data (rare) or imputed factor share data (frequent). For more details, see Section 7. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Appendices

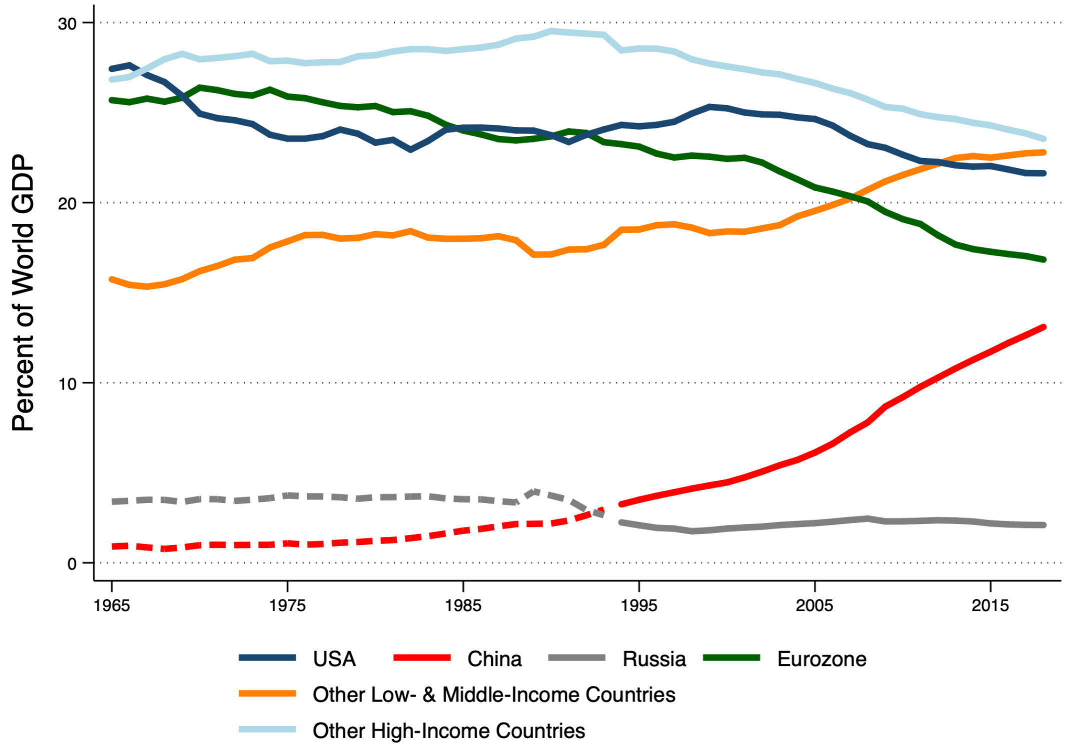
Appendix Figures and Tables

Figure A1: Data Coverage for Effective Tax Rates



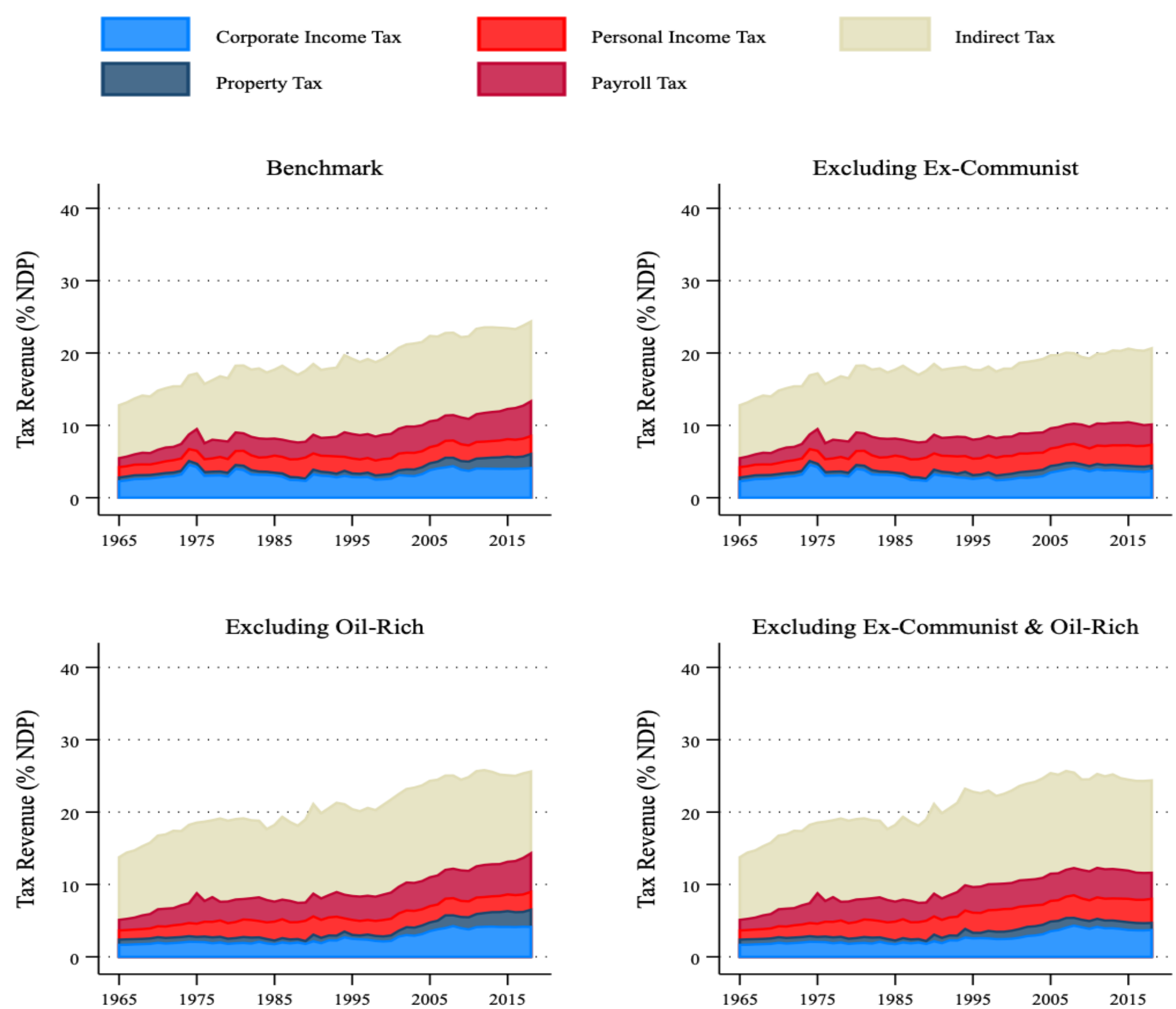
Notes: This figure shows the coverage of our effective tax rate data between 1965 and 2018, globally and for high vs. low- and middle-income countries. The solid lines plot the percentage of total population and GDP that is covered in our data (left axis). The dashed lines show the number of countries in the data (right axis). The missing 'missing' income (and population) prior to the 1990s corresponds to communist countries, particularly China, Russia and the ex-Soviet republics, and Vietnam. In addition to limited data on public revenue, communist country present a conceptual mismatch with our framework for factor income taxation (see Appendix A). Other missing country-years correspond to conflicts, and in a few cases to missing data.

Figure A2: Share of Worldwide GDP by Country



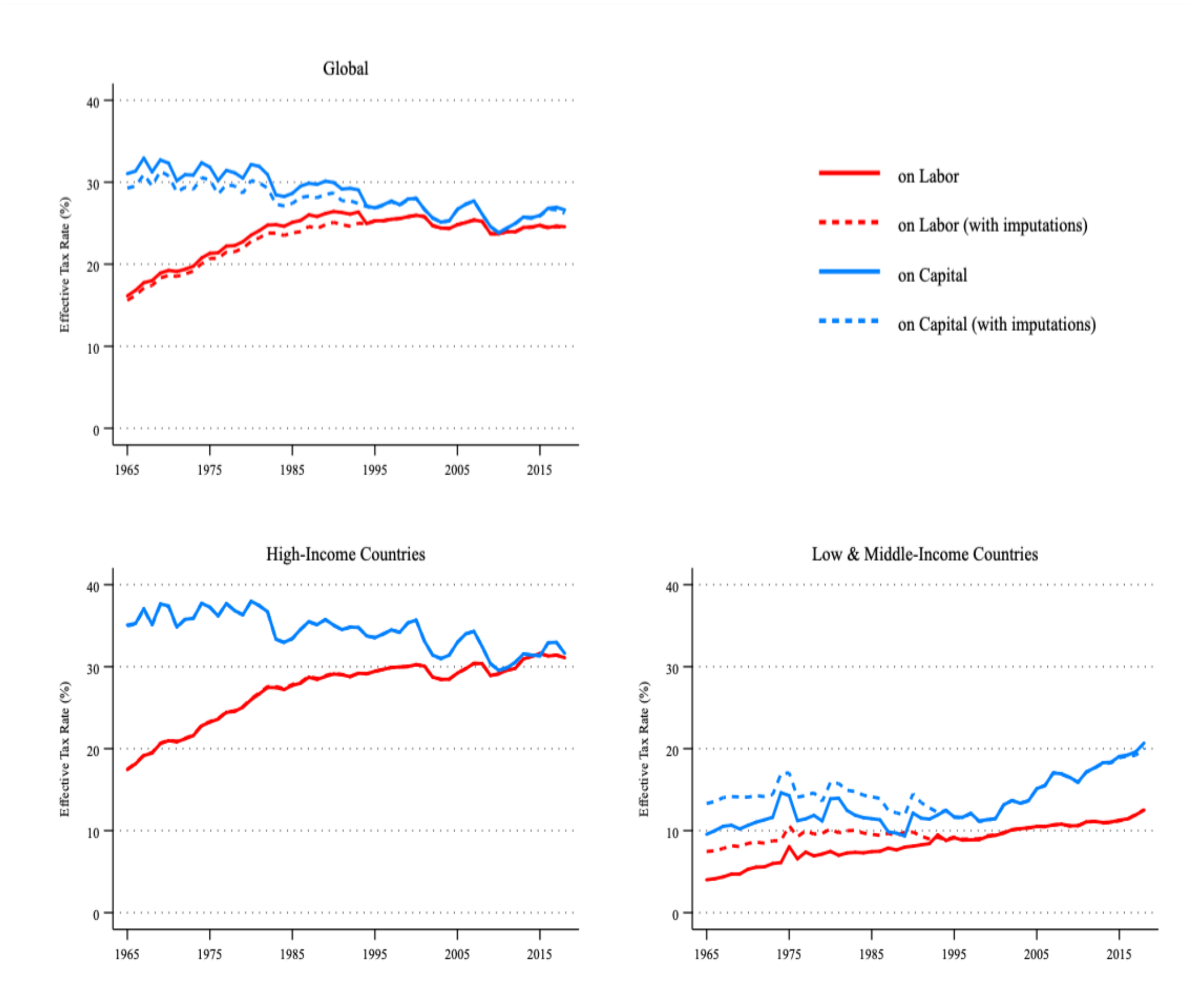
Notes: This figure plots the relative weight in world GDP of important countries. The countries are the USA, Eurozone 19, China, Russia, the remaining high-income countries and the remaining low and middle-income countries (less China and Russia). By construction these lines add to 100% of world GDP in each year. The dotted lines for China and Russia are to show that do not have data on effective tax rates for these countries pre-1994, as discussed in Appendix A and Figure A1, but we know for every year their contribution to world GDP.

Figure A3: Tax Revenue in Low- & Middle-Income Countries



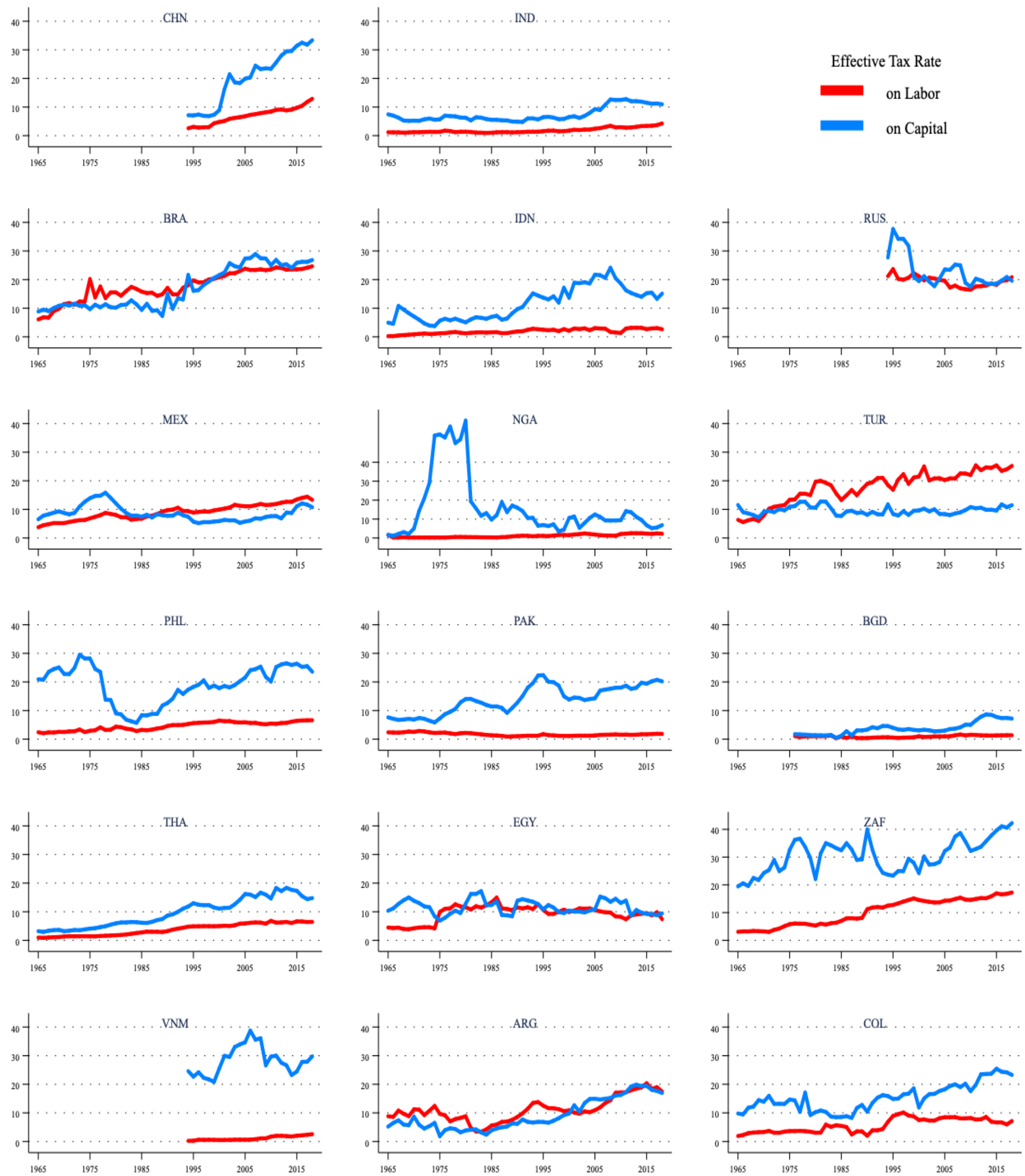
Notes: This figure plots the time series of tax revenue as a share of net domestic product (NDP), separated by revenue source, for the 119 low- and middle-income countries (sample countries that are not high-income OECD countries), weighting country-year observations by their share in that year's total NDP, in constant 2019 USD. The top left-panel repeats the benchmark figure for low and middle income countries. The top-right panel excludes former communist countries, most notably China and Russia. The mid-left panel excludes oil-rich countries (the 33 countries where average oil production since 1990 has exceeded 6.5% of GDP, per Ross and Mahdavi (2015)). The mid-right panel excludes both ex-communist and oil-rich nations.

Figure A4: Impact of Imputations on Effective Tax Rates



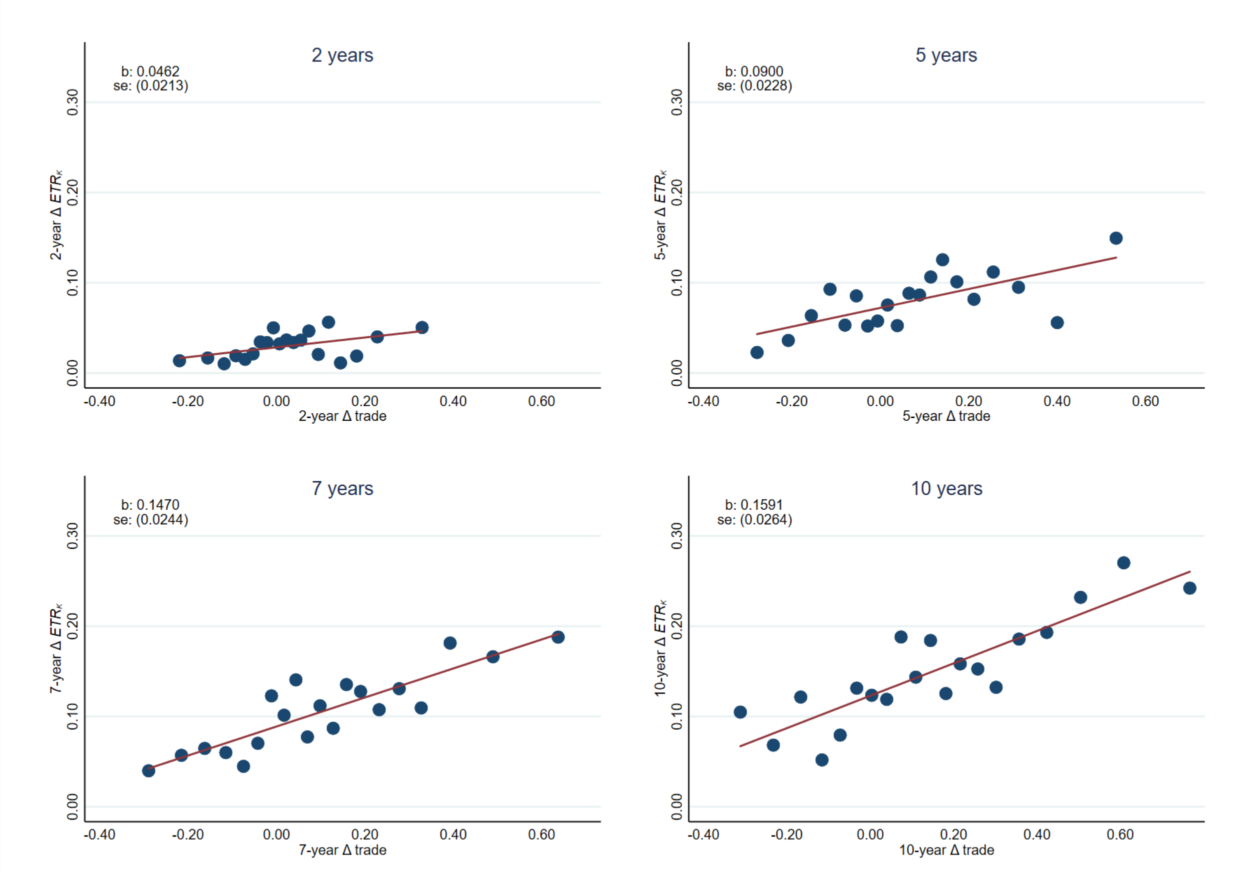
Notes: This figure plots the time series of effective tax rates for labor and capital series with and without imputations for missing country-years. This allows us to control for the sample composition changing over time and shows the plausible impact that having a balanced panel might have had on the time-series of ETRs. The solid lines correspond to the ETR series as shown in the paper and are derived from the unbalanced panel of countries, where ex-communist countries enter in 1994, including China and Russia. The dotted lines show the results of imputing values for the missing country-years in the pre-1994 data, as to have a fully balanced panel. We impute ETR values of the missing years based on a simple prediction model, which extends the trends observed in the data to the missing observation. Concretely, we decompose the ETR on capital and on labor into year and country fixed effects, separately for high versus low and middle-income countries: $ETR_{i,t} = \beta_t + \beta_i + \epsilon_{i,t}$, where β_t and β_i are year and country fixed effects. Each country-year is weighted by its share of worldwide income (in constant 2019 USD). We impute missing data by adding, to the first year of available data for a given country, the difference between the year fixed-effect of a missing year and that of the first available year in the country's time series. For example, the imputed 1993 value of ETR_K for China (whose series begins in 1994) is constructed as $ETR^K_{CHN,1993} = ETR^K_{CHN,1994} + \hat{\beta}_{1993} - \hat{\beta}_{1994}$.

Figure A5: Effective Tax Rates in Large Developing Countries



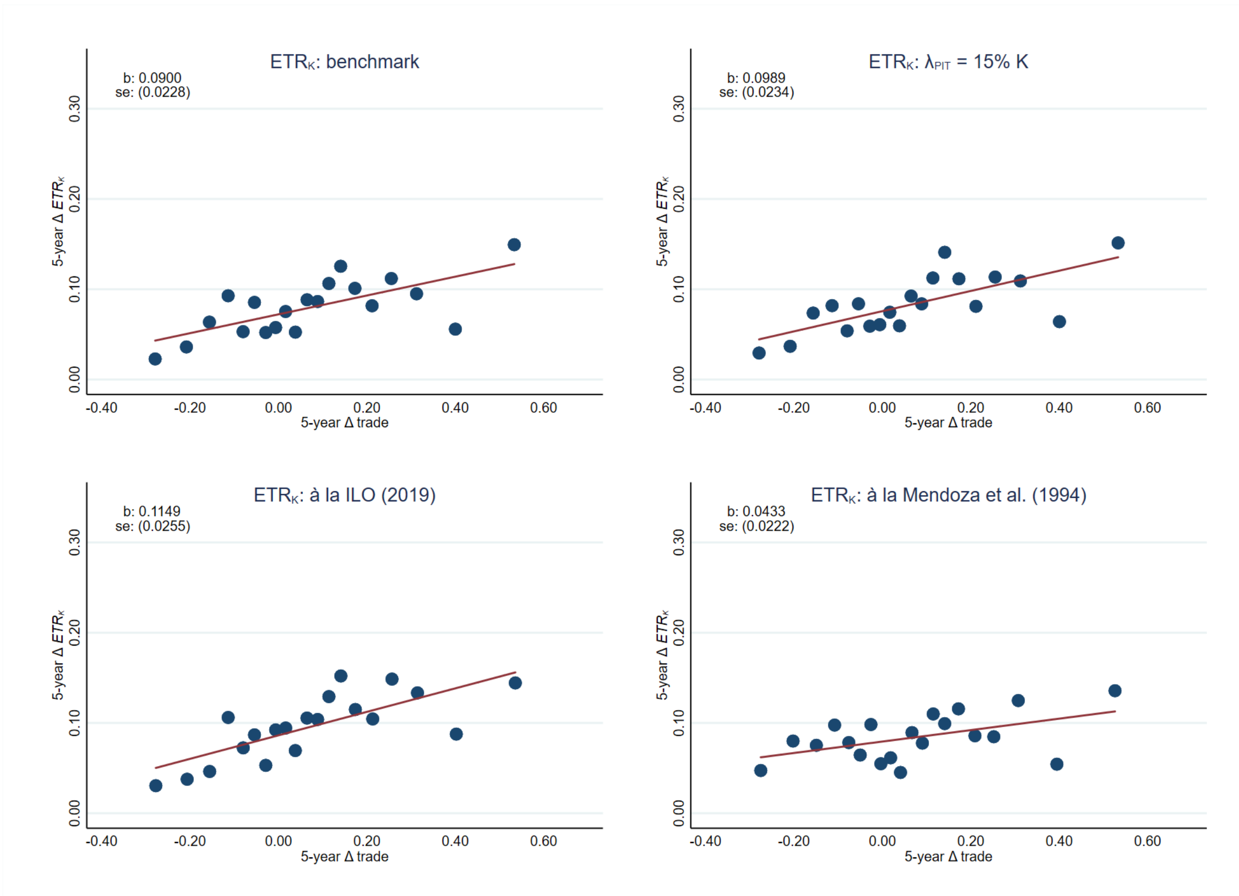
Notes: This figure shows the evolution of effective tax rates on labor and capital for the 17 largest low and middle income countries. Countries are displayed when they rank in the top 20 both in terms of population and GDP, in 2018. Dotted lines correspond to missing data which was imputed based on the fixed effect model presented in section 4.

Figure A6: Change in Capital Taxation vs. Change in Trade - Alternative Time Horizons



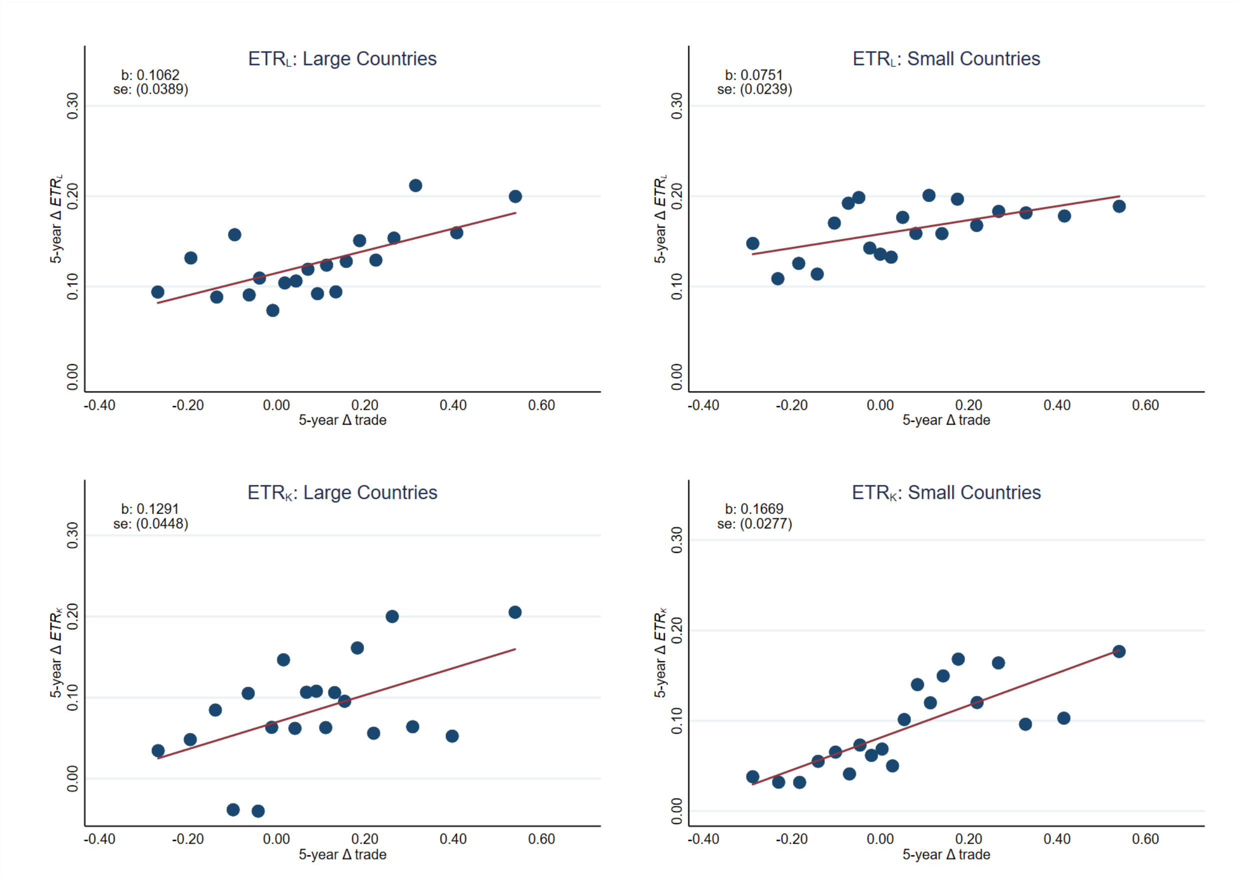
Notes: Clockwise from top left, the panels represent the association between changes in trade and changes in the effective tax rate on capital ETR_K over different time horizons: two years; five years; seven years; and ten years. Trade is measured as sum of imports and exports in goods and services and expressed as a percent of net domestic product. Both the x-axis and y-axis are measured as within-country percent changes over the varying first-difference time-horizons. Each graph plots binned scatters of each outcome against trade, after having residualized all variables against year fixed effects. Each dot corresponds to a ventile (20 equal-sized bins) of the residualized trade variable. For ease of interpretation, we add back the (non-residualized) mean of the given variable. Linear trend lines are unweighted, with year fixed effects, and are estimated based on the underlying country-year panel data. The corresponding slope and standard error are shown top-left in each panel.

Figure A7: Change in Capital Taxation vs. Change in Trade - Alternative Measurements



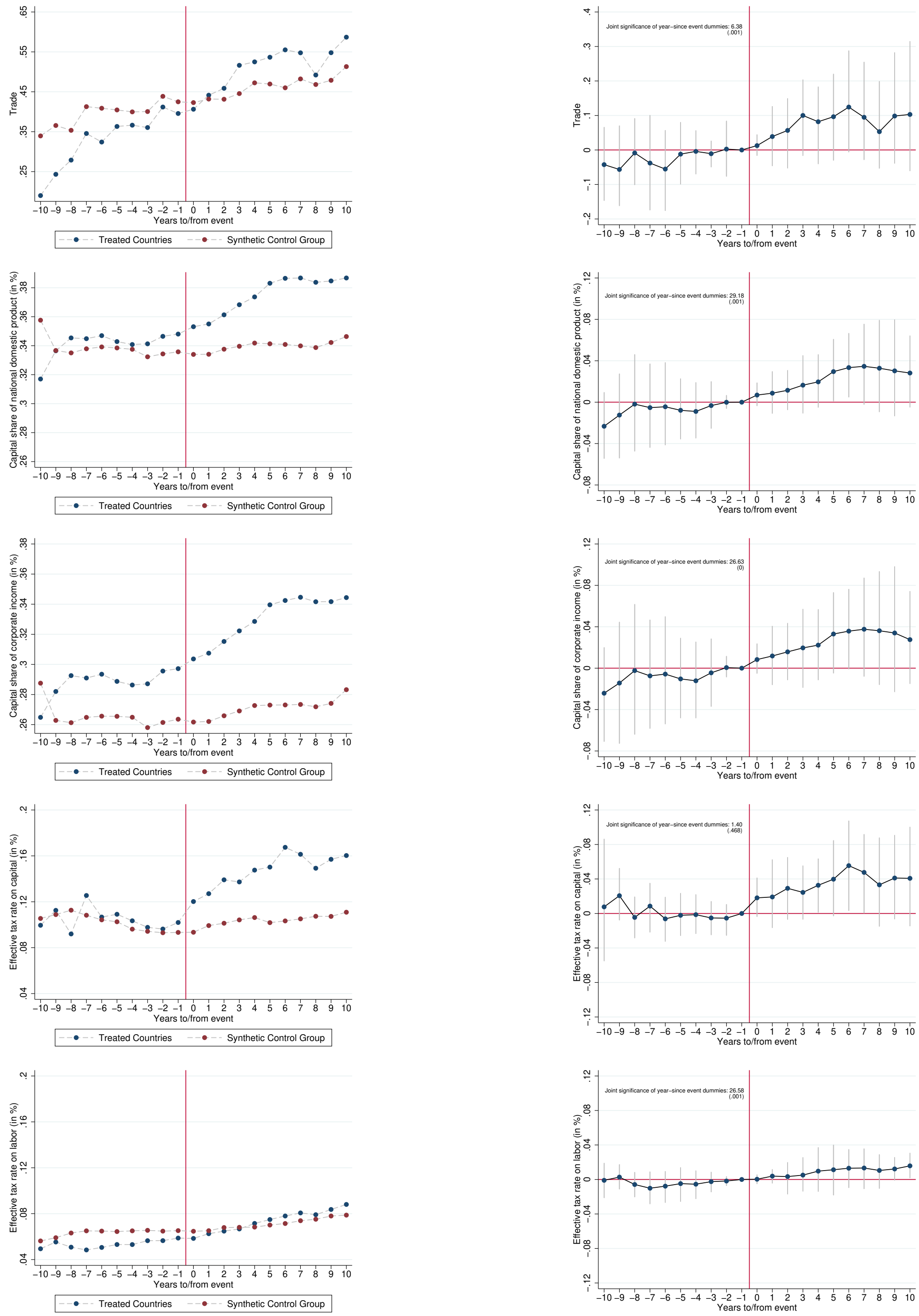
Notes: Clockwise from top left, the panels represent the association between changes in trade and changes in the effective tax rate on capital ETR_K when ETR_K is measured using different approaches: our benchmark measure (discussed in Section 3.1); fixing λ_{PIT} at 15% incidence of PIT on capital income; measuring effective tax rates using the methodology in Mendoza, Razin, and Tesar (1994); and measuring factor shares using the methodology in ILO (2019). Trade is measured as sum of imports and exports in goods and services and expressed as a percent of net domestic product. Both the x-axis and y-axis are measured as within-country percent changes over a 5-year time-horizon. Each graph plots binned scatters of each outcome against trade, after having residualized all variables against year fixed effects. Each dot corresponds to a ventile (20 equal-sized bins) of the residualized trade variable. For ease of interpretation, we add back the (non-residualized) mean of the given variable. Linear trend lines are unweighted, with year fixed effects, and are estimated based on the underlying country-year panel data. The corresponding slope and standard error are shown top-left in each subgraph.

Figure A8: Change in Factor Taxation vs. Change in Trade, by Population



Notes: Clockwise from top left, the panels represent the association between changes in trade and changes in: ETR_L in large, populous countries; ETR_L in small countries; ETR_K in large countries; and ETR_K in small countries. The top panels show changes in the effective taxation of labor income, while the bottom panels show changes in effective capital taxation. A country is defined to be large if the population exceeds 40 million in 2018. Trade is measured as a percentage of net domestic product. Both the x-axis and y-axis are measured as within-country percent changes over 5 years. Each graph plots binned scatters of each outcome against trade, after having residualized all variables against year fixed effects. Each dot corresponds to a ventile (20 equal-sized bins) of the residualized trade variable. For ease of interpretation, we add back the (non-residualized) mean of the given variable. Linear trend lines are unweighted, with year fixed effects, and are estimated based on the underlying country-year panel data. The corresponding slope and standard error are shown top-left in each subgraph.

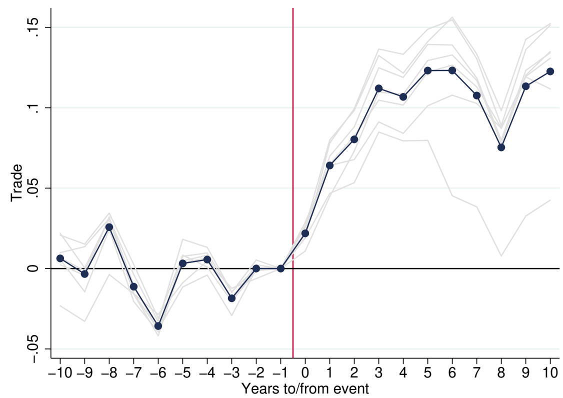
Figure A9: Event Studies Around Policy-Induced Trade Liberalization Events With Simultaneous Matching On Main Outcomes



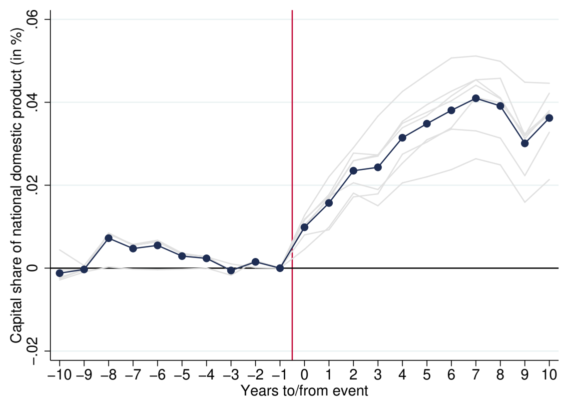
Notes: These figures show event-studies for trade liberalization reforms in seven countries, over five outcomes: trade (as a percentage of domestic product); capital shares (overall and within the corporate sector); and effective tax rates (on capital and labor). The left-hand graphs show the average level of the outcome in every year relative to the event, for the treated group and for the group of synthetic controls. The right-hand graphs show the coefficients on the 'to' ('since') dummies, in a regression model with country fixed effects; year 'to' ('since') fixed effects; and calendar-year fixed effects. The bars represent the 95% confidence intervals for 'to' ('since') reform coefficients, while standard errors are clustered at the country-reform level and estimated using the wild bootstrap method. In the top-left corner, we report the F-statistic on joint significance of the post-reform dummies, with the p-value in parentheses below. These graphs are constructed similar to Figure 7, with the exception that the synthetic control for each event-country is based on matching simultaneously on all outcomes.

Figure A10: Robustness of Event-Study to Changing Set of Treatment Countries

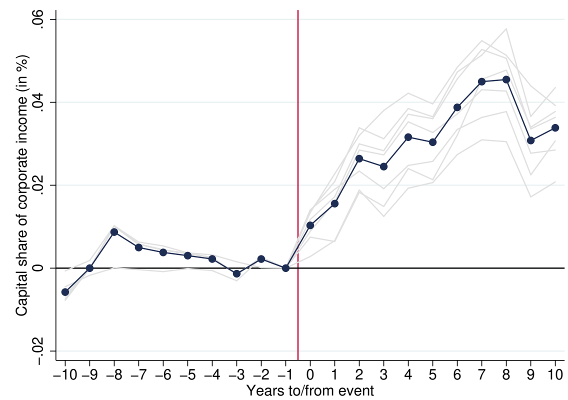
(a) Trade as % of NDP, Event Study



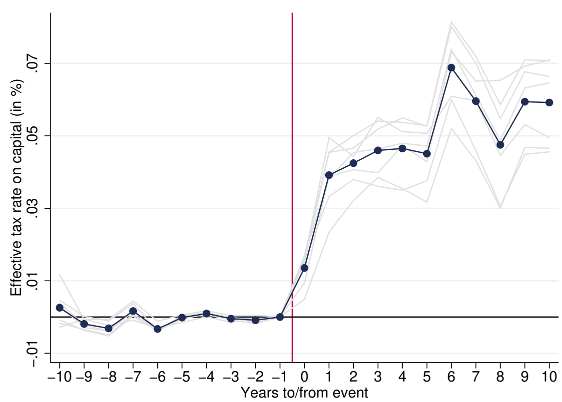
(b) Capital Share of NDP, Event Study



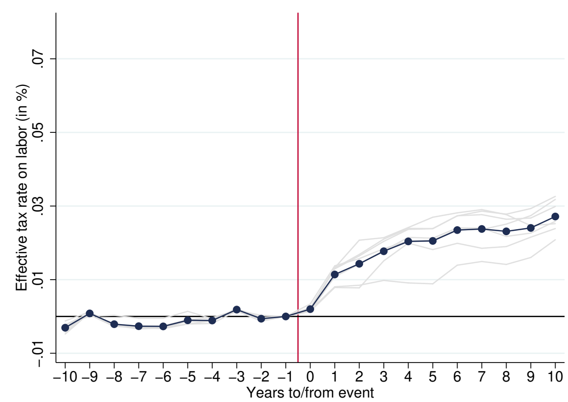
(c) Capital Share of Corp. Income, Event Study



(d) Effective Tax Rate on Capital, Event Study

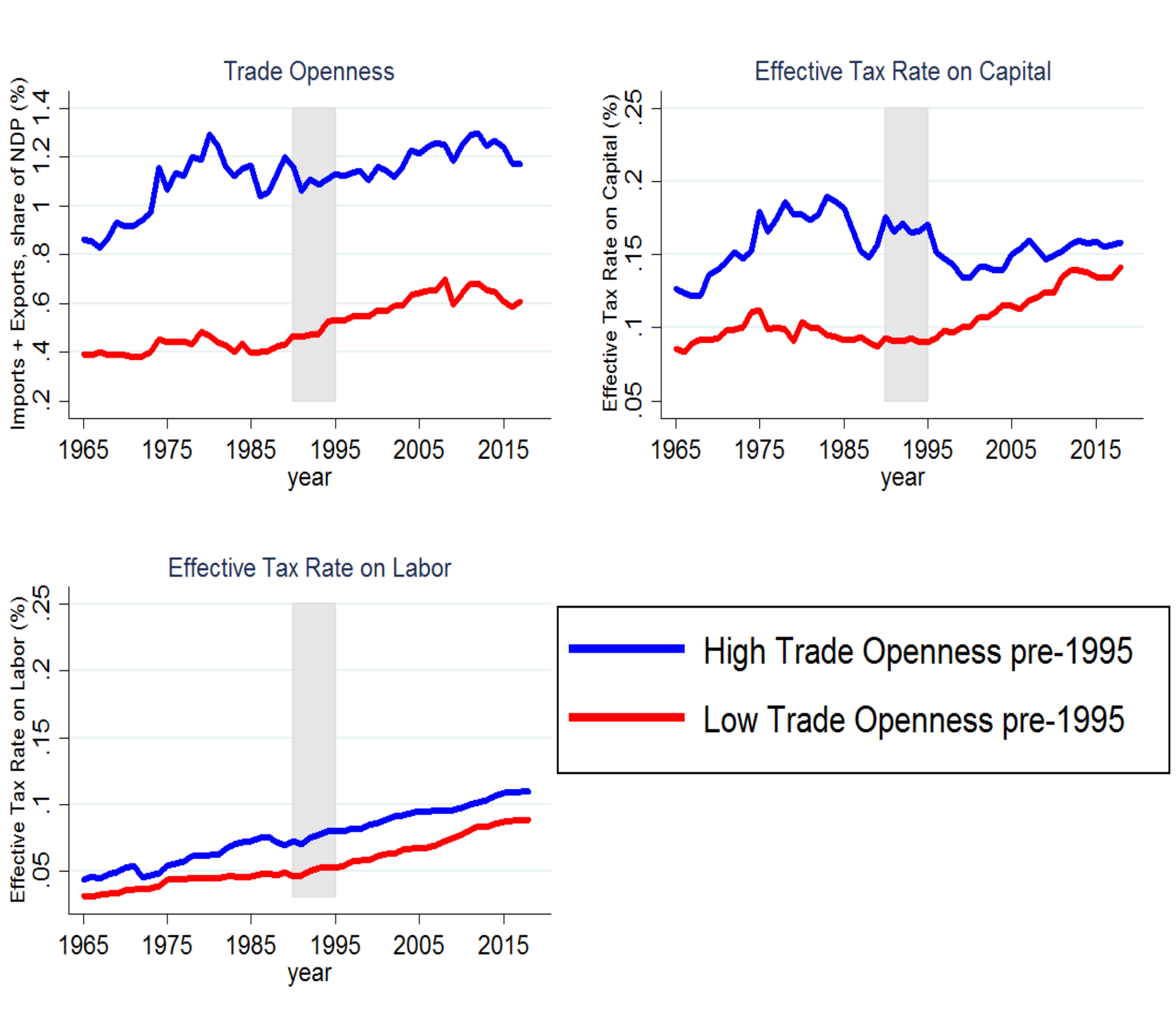


(e) Effective Tax Rate on Labor, Event Study



Notes: These figures show event studies for trade liberalization reforms in seven countries, over five outcomes: trade (as a percentage of domestic product); capital shares (overall and within the corporate sector); and effective tax rates (on capital and labor). In each figure, the solid blue line displays the estimated coefficients for the interaction between a treatment dummy and a year 'to' ('since') dummy [note the omitted period is $t-1$], corresponding to the graphs displayed in the right column of Figure 7. Each lightly-shaded gray line repeats the estimation procedure based on a sample that removes one of the seven treated countries, one at a time. All the gray lines thus represent the dynamic treatment effects but for different subsets of the treated countries. More details can be found in Appendix C.2.

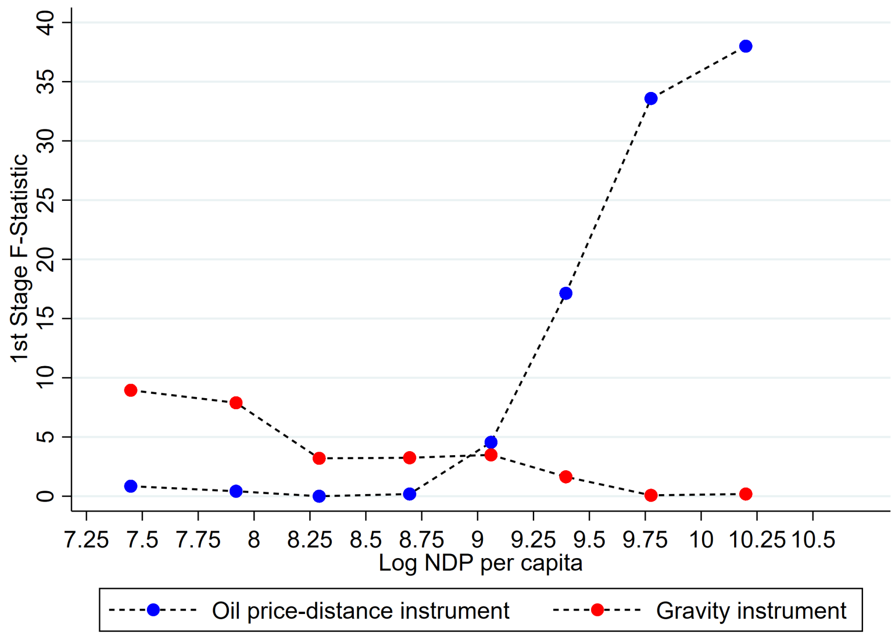
Figure A11: Trends in Factor Taxation by Initial Trade Openness, in Developing Countries



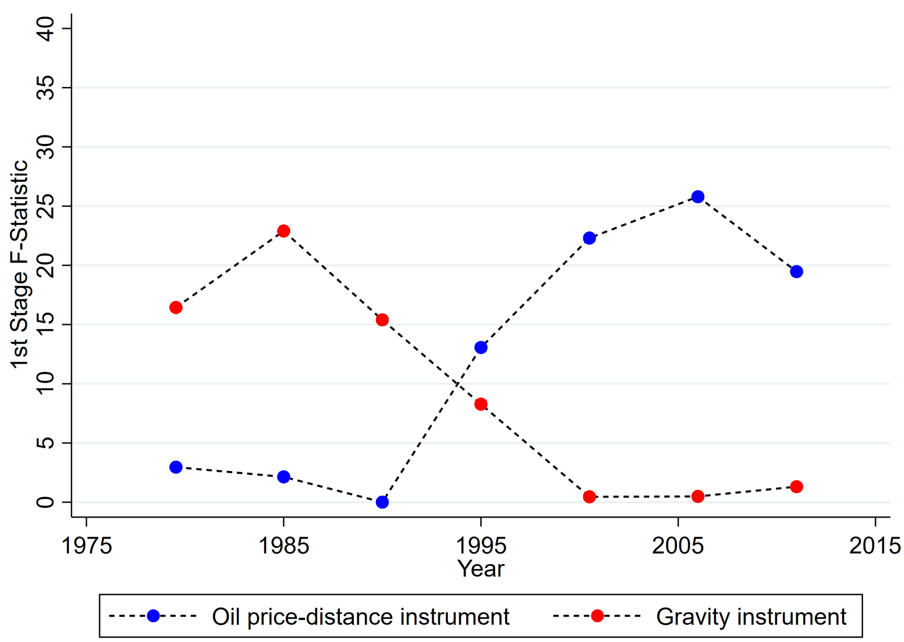
Notes: This figure plot the time series of trade openness (top-left panel), average effective tax rates on capital (top-right panel) and labor (bottom-left panel). The sample is limited to low- and middle-income countries. Within each panel, the blue line (red line) traces the evolution of the group of developing countries which had relatively high (low) trade openness prior to 1995. Specifically, high (low) trade openness is defined as having average yearly trade openness which lies above (below) the global yearly 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 unweighted OLS regression, in the relevant sub-sample of the outcome, on country and year fixed effects. The inclusion of country fixed effects eliminates 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, featuring a proliferation of bilateral and multilateral trade agreements (Egger, Nigai, and Strecker, 2019).

Figure A12: Strength of Individual Instruments Across Subsamples

(a) Sub-samples of NDP per capita

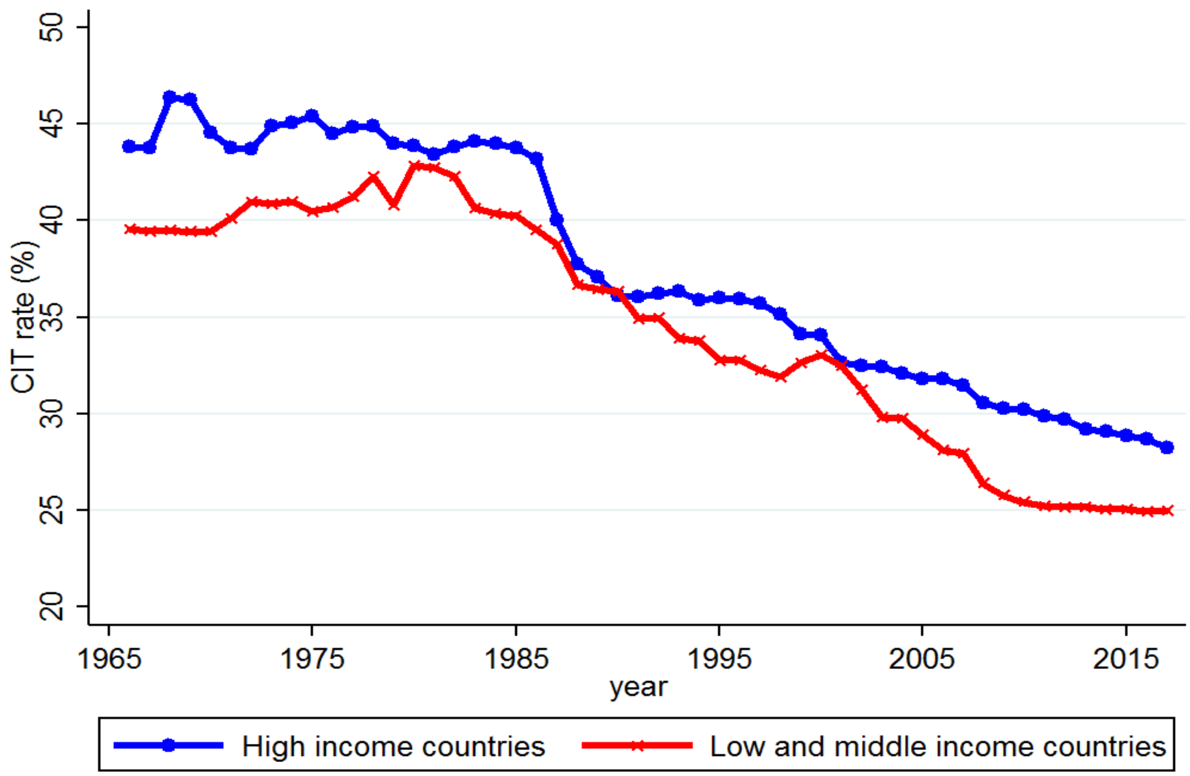


(b) Sub-samples of time-periods



Notes: These figures show the individual statistical strength of the two instruments, denoted 'oil-distance' and 'gravity'. The outcome is the first-stage F-statistic from a regression of trade openness on the individual instruments (see Section 7). The outcome is shown across subsamples of log GDP per capita (Panel A) and years (Panel B). To construct each figure, the x-axis is first partitioned into ten deciles (ten bins of equal size). The first-stage F-statistic is then separately estimated in samples centered on each decile. The estimation is done in increments of one decile, and the bandwidth uses one decile of data on either side of the decile-center. To maintain an equal size in all estimation samples, estimation centered on the first and the tenth decile are therefore dropped. The value on the x-axis is the average value of the partitioning variable in each estimation sample.

Figure A13: Trends in Corporate Income Tax Rates



Notes: This figure plot the time series of the statutory corporate income tax rate, separately for high income countries (blue dots, $N = 35$) and for middle and low-income countries (red crosses, $N = 118$). Each line plots the year fixed effects from an OLS regression in the relevant sub-sample of the CIT rate on country and year fixed effects. The inclusion of country fixed effects overcomes 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 observations are weighted with their share in the year's total NDP in constant 2019 USD.

Table A1: Weights in Synthetic Control, Matched on All Outcomes Simultaneously

Country	Matched on trade, capital share, effective labor and capital tax rates	Weight
Argentina	Iran	66.8 %
	Bangladesh	24.6 %
	Ghana	8.6 %
	.	.
Brazil	Greece	28.0 %
	United States	27.2 %
	Bangladesh	21.3 %
	...	...
China	Bangladesh	50.5 %
	Nepal	22.5 %
	Iran	10.2 %
	...	...
Colombia	Bangladesh	50.7 %
	Iran	22.6 %
	Guatemala	12.5 %
	...	...
India	Sudan	61.9 %
	Bangladesh	23.6 %
	Chad	6.4 %
	...	...
Mexico	Bangladesh	47.0 %
	United States	28.8 %
	Uruguay	18.5 %
	...	...
Vietnam	Thailand	42.4 %
	Ghana	22.6 %
	Venezuela	21.7 %
	...	...

Notes: The seven treated countries and the corresponding weighted control groups form the panel in time relative to the event. For this robustness check, where we predict simultaneously the synthetic control using the main four outcomes of interest (i.e., trade volume, capital share of domestic product, and effective tax rates on labor and capital), we obtain the same synthetic control across all outcomes for each event, by construction. (For example, here Argentina has the same synthetic control group across all outcomes.) Since we require all predictors to be strictly balanced throughout the donor pool, we use the sample from the most restrictive outcome, trade, which yields a donor pool of 90 countries. One additional restriction applies with respect to this sample. For the outcomes on trade and ETR_K, the extrapolation of Vietnam to the years 1991-93 lead to outlier values in those years, so we do not use these imputed values. This results in the panel for these outcomes to be slightly unbalanced in the years $t-10$ to $t-8$.

Table A2: Weights in Synthetic Control Matched on Each Outcome Separately

Country	Factor share of national domestic product	Weight	Factor share in corporate sector	Weight	Effective tax rate on capital	Weight	Effective tax rate on labor	Weight	Trade	Weight	Event Year	Literature Reference
	Uruguay	28.7 %	Uruguay	27.8 %	Bangladesh	41.6 %	Chile	35.9 %	Bangladesh	97.3 %		
Argentina	Oman	18.6 %	Oman	16.5 %	Haiti	14.1 %	Togo	31.6 %	United States	2.7 %	1989	Goldberg and Pavcnik (2006)
	Bolivia	16.2%	Bolivia	13.9 %	Bolivia	13.4 %	Jordan	16.8 %	.	.		
	...	...	...	...	...	...	...	...	.	.		
	Bolivia	9.1 %	Mauritius	12.2 %	Jordan	35.7 %	Panama	25.7 %	Bangladesh	59.8 %		
Brazil	Eswatini	7.6 %	Nepal	10.3 %	Sudan	21.2 %	Guyana	21.7 %	United States	32.2 %	1988	Goldberg and Pavcnik (2006), Dix-Carneiro and Kovak (2017)
	Mauritius	7.0 %	Fiji	8.9 %	Zimbabwe	12.7 %	Chile	14.5 %	Japan	6.1 %		
	...	...	...	...	...	...	...	...	...	...		
	Sweden	37.0 %	Sweden	66.1 %	Congo	41.8 %	Kuwait	31.1 %	United States	36.2 %		
China	Switzerland	24.3 %	Lesotho	27.9 %	Nicaragua	26.3 %	Pakistan	22.9 %	Bangladesh	36.0 %	2001	Brandt et al. (2006)
	Sierra Leone	21.1 %	Senegal	3.3 %	Gabon	14.2 %	Uganda	20.2 %	Dominican Republic	12.2 %		
	...	...	...	...	...	...	...	...	...	...		
	Lesotho	17.0 %	Lesotho	17.6 %	Kuwait	67.9 %	Paraguay	45.5 %	Bangladesh	50.7 %		
Colombia	Nicaragua	9.5 %	Bahrain	16.2	Gabon	14.6 %	Sudan	15.0 %	Iran	22.6 %	1985	Goldberg and Pavcnik (2006; 2016)
	Bahrain	8.1 %	France	12.6 %	Sierra Leone	12.6 %	Cameroon	11.5 %	Guatemala	12.5 %		
	...	...	...	...	...	...	...	...	...	...		
	Nepal	8.0 %	Senegal	12.8 %	Uganda	41.4 %	Lebanon	37.9 %	United States	76.4 %		
India	Nicaragua	6.9 %	Iceland	5.7 %	Bolivia	14.0 %	Oman	17.6 %	Bangladesh	23.6 %	1991	Goldberg and Pavcnik (2006, 2016); Topalova et al. (2009)
	Iceland	5.8 %	Iran	5.4 %	Haiti	4.6 %	Jordan	16.2 %	.	.		
	...	...	...	...	...	...	...	...	.	.		
	Paraguay	36.3 %	Paraguay	26.8 %	Sierra Leone	33.2 %	Tunisia	31.1	Bangladesh	72.0 %		
Mexico	Botswana	18.7 %	New Zealand	25.0 %	Bahrain	23.6 %	Zimbabwe	25.8 %	Uruguay	9.6 %	1985	Feenstra and Hanson (1997); Goldberg and Pavcnik (2006, 2016)
	Philippines	14.3 %	Botswana	18.7 %	Bolivia	14.7 %	Uruguay	15.9 %	Spain	8.0 %		
	...	...	...	...	...	...	...	...	...	...		
	Nepal	27.5 %	Ethiopia	28.7 %	Korea	45.8 %	Bangladesh	72.8 %	Thailand	42.4 %		
Vietnam	United States	10.1 %	Myanmar	10.4 %	Luxembourg	19.2 %	Myanmar	22.6 %	Ghana	22.6 %	2001	Goldberg and Pavcnik (2016), McCaig and Pavcnik (2018)
	Switzerland	4.7 %	Nepal	8.8 %	Trinidad and Tobago	17.3 %	Haiti	4.6 %	Venezuela	21.7 %		
	...	...	...	...	...	...	.	.	...	...		

Notes: The seven treated countries and the corresponding weighted control groups form the panel in time relative to the event. For each outcome, the pool of possible donor countries consists of 103 countries, with the exception of the trade variable, where we have only 90 countries with a balanced panel over the period considered. Note that the synthetic control method requires the panel of possible donor countries to be strictly balanced in all 'pre' periods that are used in the matching procedure. One additional restriction applies with respect to this sample. For the outcomes on trade and ETR_K, the extrapolation of Vietnam to the years 1991-93 lead to outlier values in those years, so we do not use these imputed values. This results in the panel for these outcomes to be slightly unbalanced in the years $t-10$ to $t-8$.

Table A3: Event Study Difference-in-Difference, Matched on Each Outcome Separately

	Trade	Capital Share of Income	Capital Share Corp. Income	ETR Capital	ETR Labor
Pre 10	0.00632	-0.00121	-0.00577	0.00257	-0.00306
	(0.0395)	(0.0176)	(0.0177)	(0.0642)	(0.0130)
Pre 9	-0.00338	-0.000302	-0.00000649	-0.00190	0.000832
	(0.0618)	(0.0276)	(0.0391)	(0.0286)	(0.00853)
Pre 8	0.0257	0.00723	0.00871	-0.00313	-0.00210
	(0.0503)	(0.0309)	(0.0424)	(0.0179)	(0.00891)
Pre 7	-0.0113	0.00474	0.00498	0.00166	-0.00264
	(0.0890)	(0.0259)	(0.0346)	(0.0158)	(0.0125)
Pre 6	-0.0357	0.00549	0.00379	-0.00329	-0.00270
	(0.0783)	(0.0252)	(0.0334)	(0.0186)	(0.0120)
Pre 5	0.00318	0.00291	0.00302	-0.000146	-0.00103
	(0.0536)	(0.0179)	(0.0243)	(0.0142)	(0.0128)
Pre 4	0.00560	0.00236	0.00221	0.000962	-0.00110
	(0.0390)	(0.0170)	(0.0231)	(0.0126)	(0.00954)
Pre 3	-0.0185	-0.000561	-0.00134	-0.000453	0.00187
	(0.0214)	(0.0131)	(0.0178)	(0.0115)	(0.00652)
Pre 2	0.00000838	0.00152	0.00220	-0.000867	-0.000604
	(0.0508)	(0.00415)	(0.00648)	(0.0133)	(0.00183)
Zero	0.0219	0.00986	0.0103	0.0136	0.00200
	(0.0174)	(0.00696)	(0.00999)	(0.0127)	(0.00443)
Post 1	0.0641*	0.0157	0.0156	0.0392*	0.0114
	(0.0386)	(0.0123)	(0.0168)	(0.0230)	(0.00871)
Post 2	0.0803**	0.0235**	0.0264*	0.0425**	0.0143
	(0.0423)	(0.0123)	(0.0156)	(0.0181)	(0.0124)
Post 3	0.112**	0.0243	0.0245	0.0460***	0.0177
	(0.0547)	(0.0176)	(0.0220)	(0.0188)	(0.0132)
Post 4	0.107**	0.0315**	0.0316*	0.0465**	0.0204
	(0.0485)	(0.0173)	(0.0209)	(0.0204)	(0.0160)
Post 5	0.123**	0.0348**	0.0304*	0.0451**	0.0205
	(0.0598)	(0.0186)	(0.0199)	(0.0249)	(0.0175)
Post 6	0.123	0.0381**	0.0388**	0.0688**	0.0235*
	(0.0867)	(0.0203)	(0.0219)	(0.0331)	(0.0145)
Post 7	0.108	0.0410**	0.0450**	0.0596**	0.0238*
	(0.0772)	(0.0215)	(0.0249)	(0.0312)	(0.0152)
Post 8	0.0753	0.0391*	0.0455*	0.0475	0.0231*
	(0.0747)	(0.0231)	(0.0261)	(0.0336)	(0.0138)
Post 9	0.113	0.0301	0.0308	0.0594**	0.0241**
	(0.0900)	(0.0241)	(0.0258)	(0.0298)	(0.0117)
Post 10	0.123	0.0362**	0.0338*	0.0592**	0.0272**
	(0.101)	(0.0206)	(0.0223)	(0.0321)	(0.0114
F-Test Statistic on joint insignificance of the coefficients *Zero* to *Post 10*	4.520	6.060	3.320	2.500	4.660
P-values of F-Statistic	0.00610	0.00158	0.0217	0.0593	0.00532
Post*Treat	0.0992**	0.0271*	0.0283	0.0485**	0.0199*
	(0.0530)	(0.0165)	(0.0202)	(0.0230)	(0.0111)
Imputed treatment effect	0.107***	0.0269**	0.0284**	0.0490***	0.0198***
	(0.0291)	(0.0115)	(0.0142)	(0.0146)	(0.00512)
N	284	290	292	282	290

Notes: * p<0.10 ** p<0.05 *** p<0.01. The table presents the event study estimates displayed in 7. The Post*Treat coefficient is the corresponding difference-in-difference estimate, representing the average treatment effect from period 0 through 10 post the trade liberalization event. Finally, as a further robustness check, we present an additional difference-on-difference estimate proposed by Borusyak, Jaravel, and Spiess (2021). This estimate is imputed by first estimating country and time fixed effects, using non-treated countries as well as treated countries before their respective event. Those unit and year specific estimates are then used to impute the treatment effect for every treated country, and the imputed coefficient is then the average of the individual treatment effects. Due to the small sample size, we present wild bootstrap standard errors in parentheses (Cameron, Gelbach, and Miller, 2008), except for the imputed treatment effect according to Borusyak, Jaravel, and Spiess (2021), where we report the default standard errors produced by the Stata command *did_imputation*.

Table A4: Event Study Difference-in-Difference, Matched on All Outcomes Simultaneously

	Trade	Capital Share of Income	Capital Share Corp. Income	ETR Capital	ETR Labor
Pre 10	-0.0425	-0.0233	-0.0242	0.00766	-0.000980
	(0.0601)	(0.0176)	(0.0239)	(0.0341)	(0.0116)
Pre 9	-0.0565	-0.0125	-0.0144	0.0206	0.00276
	(0.0671)	(0.0220)	(0.0316)	(0.0179)	(0.00825)
Pre 8	-0.00872	-0.00190	-0.00234	-0.00455	-0.00585
	(0.0578)	(0.0264)	(0.0357)	(0.0138)	(0.00819)
Pre 7	-0.0380	-0.00527	-0.00745	0.00854	-0.0101
	(0.0897)	(0.0231)	(0.0306)	(0.0159)	(0.0112)
Pre 6	-0.0554	-0.00445	-0.00573	-0.00623	-0.00775
	(0.0774)	(0.0234)	(0.0302)	(0.0152)	(0.0103)
Pre 5	-0.0121	-0.00786	-0.0104	-0.00211	-0.00481
	(0.0545)	(0.0170)	(0.0227)	(0.0144)	(0.0106)
Pre 4	-0.00407	-0.00895	-0.0122	-0.00137	-0.00547
	(0.0381)	(0.0152)	(0.0209)	(0.0134)	(0.00922)
Pre 3	-0.0108	-0.00329	-0.00450	-0.00518	-0.00253
	(0.0230)	(0.0131)	(0.0187)	(0.0114)	(0.00646)
Pre 2	0.00251	-0.000107	0.000555	-0.00542	-0.00179
	(0.0531)	(0.00423	(0.00640	(0.0104	(0.00317
Zero	0.0126	0.00687	0.00832	0.0181	0.000268
	(0.0190	(0.00648	(0.00833	(0.0137	(0.00337
Post 1	0.0389	0.00868	0.0118	0.0192	0.00387
	(0.0496	(0.0117	(0.0165	(0.0234	(0.00501
Post 2	0.0567	0.0115	0.0158	0.0292*	0.00337
	(0.0606	(0.0109	(0.0153	(0.0203	(0.0105
Post 3	0.100*	0.0164	0.0196	0.0244	0.00507
	(0.0682	(0.0155	(0.0211	(0.0184	(0.0111
Post 4	0.0819	0.0196	0.0223	0.0327*	0.00968
	(0.0679	(0.0146	(0.0194	(0.0194	(0.0143
Post 5	0.0964	0.0295*	0.0330*	0.0398*	0.0113
	(0.0746	(0.0170	(0.0222	(0.0264	(0.0164
Post 6	0.124	0.0334**	0.0358*	0.0555**	0.0130
	(0.0888	(0.0182	(0.0230	(0.0308	(0.0125
Post 7	0.0947	0.0346*	0.0376	0.0476**	0.0133
	(0.0842	(0.0220	(0.0279	(0.0267	(0.0124
Post 8	0.0531	0.0328	0.0362	0.0332	0.0105
	(0.0786)	(0.0260)	(0.0316)	(0.0293)	(0.0108)
Post 9	0.0984	0.0302	0.0340	0.0411	0.0122*
	(0.0946)	(0.0266)	(0.0343)	(0.0292)	(0.00772)
Post 10	0.103	0.0282	0.0276	0.0408	0.0158**
	(0.106)	(0.0200)	(0.0251)	(0.0326)	(0.00808)
F-Test Statistic on joint insignificance of the coefficients *Zero* to *Post 10*	6.380	29.18	26.63	1.040	26.58
P-values of F-Statistic	0.00122	0.00001	0.00001	0.468	0.00001
Post*Treat	0.0994*	0.0292***	0.0335***	0.0339*	0.0127
	(0.0561)	(0.00697)	(0.0106)	(0.0201)	(0.00974)
Imputed treatment effect	0.109***	0.0272***	0.0311***	0.0341***	0.0121***
	(0.0292)	(0.00455)	(0.00661)	(0.00647)	(0.00430)
N	284	290	293	287	290

Notes: * p<0.10 ** p<0.05 *** p<0.01. The table presents the event study estimates displayed in A9. The Post*Treat coefficient is the corresponding difference-in-difference estimate, representing the average treatment effect from period 0 through 10 post the trade liberalization event. Finally, as a further robustness check, we present an additional difference-on-difference estimate proposed by Borusyak, Jaravel, and Spiess (2021). This estimate is imputed by first estimating country and time fixed effects, using non-treated countries as well as treated countries before their respective event. Those unit and year specific estimates are then used to impute the treatment effect for every treated country, and the imputed coefficient is then the average of the individual treatment effects. Due to the small sample size, we present wild bootstrap standard errors in parentheses (Cameron, Gelbach, and Miller, 2008), except for the imputed treatment effect according to Borusyak, Jaravel, and Spiess (2021), where we report the default standard errors produced by the Stata command *did_imputation*.

Table A5: First-Stage Regressions

	benchmark (1)	controls (2)	unweighted (3)	pop.-wt. (4)
Outcome: Trade in goods and services (% of NDP)				
Instrument I: $Z^{Oil-Distance}$	-0.0801***	-0.0844***	-0.118***	-0.0813***
	(0.0111)	(0.0143)	(0.0333)	(0.0207)
Instrument II: $Z^{Gravity}$	0.00531	0.00452	0.00769**	0.00884**
	(0.00422)	(0.00329)	(0.00332)	(0.00392)
Outcome: Trade in goods and services (% of NDP), winsorized				
Instrument I: $Z^{Oil-Distance}$	-0.0815***	-0.0851***	-0.119***	-0.0831***
	(0.0107)	(0.0137)	(0.0312)	(0.0205)
Instrument II: $Z^{Gravity}$	0.00468	0.00410	0.00779**	0.00833**
	(0.00400)	(0.00304)	(0.00311)	(0.00379)
Outcome: Trade in goods and services (in log levels)				
Instrument I: $Z^{Oil-Distance}$	-0.0402	-0.0973***	-0.0205	-0.0482
	(0.0582)	(0.0276)	(0.0778)	(0.130)
Instrument II: $Z^{Gravity}$	0.0613***	0.0204***	0.0335***	0.0696***
	(0.0158)	(0.00437)	(0.00659)	(0.0128)
Outcome: Trade in goods (% of NDP)				
Instrument I: $Z^{Oil-Distance}$	-0.0607***	-0.0592***	-0.0797***	-0.0516*
	(0.00958)	(0.0110)	(0.0172)	(0.0294)
Instrument II: $Z^{Gravity}$	0.00656*	0.00640**	0.0176*	0.0107***
	(0.00351)	(0.00285)	(0.00957)	(0.00339)
N	*4518*	*4518*	*4518*	*4518*

Notes: This table shows the first-stage regressions of trade on the two instruments from Egger, Nigai, and Strecker (2019). The two instruments, $Z^{Oil-Distance}$ and $Z^{Gravity}$, are described in detail in Section 7 and Appendix Section B. Columns correspond to different specifications: our benchmark specification (column 1) includes net domestic product weights, without control variables; next (2) we show this benchmark with controls and NDP weights; then (3) without controls and without weights; and finally (4) without controls but with population weights. The controls included in column (2) are: USD exchange rate; gross fixed capital formation (as a percentage of NDP); (log) population; (log) gross domestic product per capita; and *de jure* capital accounts mobility (Chinn and Ito, 2006). Each panel corresponds to a different measure of trade: trade in goods and services, as a share of NDP (1); the same, but winsorized at the 95th percentiles on a yearly basis (2); the log of the value of trade in goods and services, expressed in constant 2019 USD (3); and finally trade in goods only, expressed as a percentage of NDP (4). In each panel and column, we report the coefficients on the two instruments. In this table, for ease of interpretation, the oil price-distance instrument $Z^{Oil-Distance}$ has been multiplied by 10^{10}. All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A6: Robustness of Main Results to Changing the Measurement of Trade

	Capital Share		Effective Tax Rate	
	overall	corp. sector	on capital	on labor
Trade in G&S (%NDP)	0.151**	0.184**	0.375*	0.163***
	(0.0698)	(0.0800)	(0.213)	(0.0538)
First-stage F-statistic	26.07	26.07	26.07	26.07
Trade in G&S (%NDP), winsorized	0.145**	0.178**	0.366*	0.168***
	(0.0666)	(0.0764)	(0.207)	(0.0515)
First-stage F-statistic	29.05	29.05	29.05	29.05
Trade in G&S (log levels)	0.0328***	0.0380**	0.0699*	0.00607
	(0.0124)	(0.0156)	(0.0393)	(0.0166)
First-stage F-statistic	9.025	9.025	9.025	9.025
Trade in Goods Only (%NDP)	0.196**	0.236**	0.469*	0.165**
	(0.0975)	(0.112)	(0.276)	(0.0814)
First-stage F-statistic	21.14	21.14	21.14	21.14
N	*4518*	*4518*	*4518*	*4518*

Notes: This table shows the robustness of the main IV results to changing the measure of trade. Across columns, the outcomes of interest are: capital share of net domestic product (NDP, column 1); capital share within the corporate sector (2); the effective tax rate on capital income (3); and the effective tax rate on labor income (4). Across panels, the measurement of trade varies as follows: trade in goods and services (G&S), expressed as a percentage of NDP; then the same, but winsorized at the 95th percentile; then log of the value of trade in G&S, expressed in constant 2019 USD; and finally trade in goods only, expressed as a percentage of NDP. At the bottom of each panel, we report the first-stage F-statistic. All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A7: Robustness to Changing the Weights

	Capital Share		Effective Tax Rate	
	overall	corp. sector	on capital	on labor
IV: with NDP weights	0.151**	0.184**	0.375*	0.163***
	(0.0698)	(0.0800)	(0.213)	(0.0538)
First-stage F-statistic	26.07	26.07	26.07	26.07
IV: with population weights	0.225*	0.214*	0.553*	0.159**
	(0.119)	(0.129)	(0.296)	(0.0792)
First-stage F-statistic	10.97	10.97	10.97	10.97
IV: unweighted	0.118*	0.122	0.250**	0.133**
	(0.0681)	(0.0826)	(0.105)	(0.0526)
First-stage F-statistic	8.415	8.415	8.415	8.415
N	*4518*	*4518*	*4518*	*4518*

Notes: This table shows the robustness of the main IV results to changing weights. Across columns, the outcomes of interest are: capital share of net domestic product (NDP, column 1); capital share within the corporate sector (2); the effective tax rate on capital income (3); and the effective tax rate on labor income (4). In these IV specifications, we instrument for trade (in goods and services, expressed as a percentage of NDP) with both price-distance and gravity instruments (as discussed in Section 7). Across panels, estimates are weighted here by NDP in constant 2019 USD at PPP; or by national population in each year of the observation; or without weights. All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A8: Robustness to Controls & Oil-Rich*Time Fixed Effects

	Capital Share		Effective Tax Rate	
	overall	corp. sector	on capital	on labor
IV: without controls	0.151**	0.184**	0.375*	0.163***
	(0.0698)	(0.0800)	(0.213)	(0.0538)
First-stage F-statistic	26.07	26.07	26.07	26.07
IV: with controls	0.115**	0.142**	0.400***	0.226***
	(0.0475)	(0.0546)	(0.112)	(0.0551)
First-stage F-statistic	19.02	19.02	19.02	19.02
IV: with controls & oil-rich time FE	0.261**	0.275**	0.573**	0.386**
	(0.121)	(0.131)	(0.232)	(0.151)
First-stage F-statistic	3.896	3.896	3.896	3.896
N	*4518*	*4518*	*4518*	*4518*

Notes: This table shows the robustness of the main IV results to including controls and then additional time fixed effects for countries in which oil is an important part of the domestic product, exceeding 6.5% of GDP (Ross and Mahdavi, 2015; Mahdavi, 2020). Across columns, the outcomes of interest are: capital share of net domestic product (NDP, column 1); capital share within the corporate sector (2); the effective tax rate on capital income (3); and the effective tax rate on labor income (4). Across panels, the first set of results show our benchmark specification from Table 3, then to a vector of control variables, then adding additional time fixed effects for the oil-rich countries. The control variables are as follows: USD exchange rate; gross fixed capital formation (as a percentage of NDP); (log) population; (log) gross domestic product per capita; and *de jure* capital accounts mobility (Chinn and Ito, 2006). In these IV specifications, we instrument for trade (in goods and services, expresses as a percentage of net domestic product) with both price-distance and gravity instruments (as discussed in Section 7). Estimates are weighted by net domestic product in constant 2019 USD at PPP. All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A9: Robustness to Estimating Factor Shares à la ILO

	Capital Share	Effective Tax Rate	
	overall	on capital	on labor
OLS	-0.00602	0.00126	0.0214**
	(0.0207)	(0.0332)	(0.0103)
IV: benchmark specification	0.194	0.231*	0.194***
	(0.135)	(0.137)	(0.0652)
First-stage F-statistic	26.07	26.07	26.07
IV: with controls	0.205*	0.237**	0.264***
	(0.113)	(0.0990)	(0.0622)
First-stage F-statistic	19.02	19.02	19.02
IV: unweighted	-0.106	0.234*	0.0865
	(0.118)	(0.126)	(0.0591)
First-stage F-statistic	8.415	8.415	8.415
N	*4518*	*4518*	*4518*

Notes: This table shows the robustness of the main IV results to measuring factor shares à la ILO (2019). Across columns, the outcomes of interest are: capital share of net domestic product (NDP, column 1); the effective tax rate on capital income (2); and the effective tax rate on labor income (3). (Note that the capital share within the corporate sector does not change in this adjustment, by construction, as the ILO adjustment affects strictly the measurement of self-employment income.) Across panels, we present first the OLS estimate, then our benchmark IV specification, and then to this we add controls or remove the weight (NDP in constant 2019 USD at PPP). The additional control variables include: USD exchange rate; gross fixed capital formation (as a percentage of NDP); net foreign direct investment; (log) population; (log) gross domestic product per capita; and *de jure* capital accounts mobility (Chinn and Ito, 2006). In all IV specifications here, we instrument for trade (in goods and services, expressed as a percentage of NDP) with both price-distance and gravity instruments (as discussed in Section 7). Estimates are weighted by net domestic product in constant 2019 USD at PPP. All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A10: Robustness to Alternative Measurements of Effective Tax Rates

	OLS (1)		IV (2)		IV: with controls (3)		IV: unweighted (4)	
	ETR_K	ETR_L	ETR_K	ETR_L	ETR_K	ETR_L	ETR_K	ETR_L
à la Mendoza	-0.00287	0.0125	0.261	0.138**	0.288***	0.183***	0.155**	0.0741
	(0.0185)	(0.00899)	(0.158)	(0.0527)	(0.0889)	(0.0457)	(0.0676)	(0.0451)
PIT = 0% on capital	0.00830	0.0285***	0.387*	0.158***	0.408***	0.222***	0.272**	0.123**
	(0.0249)	(0.0108)	(0.198)	(0.0559)	(0.107)	(0.0565)	(0.107)	(0.0532)
PIT = 30% on capital	0.00685	0.0245**	0.394*	0.157***	0.430***	0.214***	0.260**	0.126**
	(0.0317)	(0.00979)	(0.231)	(0.0524)	(0.127)	(0.0513)	(0.111)	(0.0501)
PIT = 15% on capital	0.00817	0.0265**	0.390*	0.158***	0.419***	0.218***	0.267**	0.125**
	(0.0282)	(0.0103)	(0.214)	(0.0540)	(0.116)	(0.0538)	(0.108)	(0.0515)
First-stage F-statistic			26.07	26.07	19.02	19.02	8.415	8.415
N	*4518*	*4518*	*4518*	*4518*	*4518*	*4518*	*4518*	*4518*

Notes: This table shows the robustness of the main results alternative measures of effective tax rates—either measured with the Mendoza, Razin, and Tesar (1994) definition, or with the capital share of PIT revenues bounded at levels that are either unrealistically low (0%) or unrealistically high (30%)—or with an intermediate fixed parameter of 15%. Specification (1) presents the OLS estimate. We present the benchmark IV specification in column (2), and then either add controls in (3) or remove NDP weights in (4). Specifications (1)-(3) are weighted by NDP in constant 2019 USD at PPP. The additional control variables in (3) include: USD exchange rate; gross fixed capital formation (as a percentage of NDP); (log) population; (log) gross domestic product per capita; and *de jure* capital accounts mobility (Chinn and Ito, 2006). All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A11: First-stage and Core Results: Individual Instruments vs. Both Instruments

	first-stage			IV	
	(1)	(2)	(3)	ETR_K	ETR_L
Panel A: Both Instruments Together					
$Z^{Gravity}$ and $Z^{Oil-Distance}$				0.250**	0.133**
				(0.105)	(0.0526)
First-stage F-statistic				8.415	8.415
Panel B: Each Instrument Separately					
$Z^{Gravity}$	0.00769**	0.00749**		0.294*	0.134*
	(0.00332)	(0.00332)		(0.163)	(0.0767)
First-stage F-statistic				5.101	5.101
$Z^{Oil-Distance}$	-0.118***		-0.111***	0.151**	0.131***
	(0.0333)		(0.0319)	(0.0645)	(0.0283)
First-stage F-statistic				12.14	12.14
N	*4518*	*4518*	*4518*	*4518*	*4518*

Notes: This table shows the main results for both instruments together (Panel A), and then for each instrument separately (Panel B). In column (1), however, we show OLS estimates of trade (in goods and services, expressed as a percentage of NDP) regressed on both instruments together. Then columns (2) and (3) show OLS estimates of trade on each instrument separately. In the columns to the right, we present IV estimates for the effective tax rate on capital income ETR_K and on labor income ETR_L. See Section 7 and Appendix B for further discussion of these instruments. For ease of interpretation, these estimates are unweighted (therefore, these estimates match Panel B of Table 3 as well as Table A7). The price-distance instrument $Z^{Oil-Distance}$ has been multiplied by 10^{10}, for ease of interpretation. All estimates include country and year fixed effects, with errors clustered at country level. We use dummy variable controls for significantly interpolated revenue data or imputed factor share data. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A12: Association between statutory CIT rate and capital tax collection

	ETR_K			CIT revenue (% GDP)		
	All ctes	High inc	Mid-low inc	All ctes	High-inc	Mid-low inc
CIT rate	0.201***	0.216**	0.188***	0.037**	.036*	.038*
	(0.055)	(0.097)	(0.067)	(0.016)	(0.020)	(.022)
N	*3653*	*1225*	*2428*	*3653*	*1225*	*2428*

Notes: This table presents results from estimating the association between the corporate statutory tax rate and the effective tax rate on capital, ETR_K, as well as the share of corporate income tax revenue in GDP. For both outcomes, the association is estimated for all countries, then separately for high income countries, and for middle and low-income countries. All regressions include country and year fixed effects. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table A13: Heterogeneous impacts on taxes (as % of GDP) by development level

	Corporate income taxes	Capital taxes	Total taxes
	(1)	(2)	(3)
Trade	0.102***	0.168***	0.241**
	(0.028)	(0.061)	(0.150)
Trade$*\mathbb{1}$(High-income)	-0.127**	-0.188	-0.281
	(0.061)	(0.121)	(0.391)
Trade in High-income	-0.025	-0.019	-0.039
	(0.035)	(0.077)	(0.200)
N	*4518*	*4518*	*4518*

Notes: This table presents results for the impact of trade on taxation, expressed as a % of GDP. Across columns, the outcome is: corporate income taxes, capital taxes, total taxes. Total taxes include all taxes collected on capital, labor and consumption. All regressions are estimated using IV. The regression coefficients for $Trade$ as well as the interaction with a dummy for high-income countries, $Trade * \mathbb{1}(High - income)$ are presented. The bottom row reports the coefficient for the linear combination of $Trade$ and the interaction term. Trade is measured as the sum of export and imports divided by GDP. We instrument for trade using the oil-price and the gravity-equation instruments from Egger, Nigai, and Strecker (2019). All estimates include country and year fixed effects and observations are weighted by net domestic product in constant 2019 USD at PPP. We use dummy variable controls for significantly interpolated revenue data (rare) or imputed factor share data (frequent) in all columns. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Appendix A Data-sources and methodology

A.1 Separating national income into capital and labor income

Data We used two datasets for computing factor shares of net domestic product. Both come from the United Nations Statistics Division. The first was the SNA2008 (UN SNA, 2008) online data repository. The second was SNA1968 archival material (UN SNA, 1968). While the SNA2008 data has extensive coverage, we were able to agument this coverage in the historical era using SNA1968 observations. As noted in Section 3.2.1 and Table 2, we nearly double our 'complete' coverage observations by adding the historical data.[45] What we here can describe as 'complete' observations are those that include a split of basic factor income concepts, including the compensation of employees in the household sector and operating surplus in the corporate sector. Let us briefly describe these national income accounting concepts here below, and how we attribute them to labor or capital factors of production, before returning to discuss missing national accounts data and overall coverage.

Income Concepts It may be useful here to enumerate the basic accounting concepts expressed in equation (1) above: 'Compensation of employees' (CE) refers to formal labor income. It includes wages and salaries, employer and employee social contributions, and all payments from employers in a current or deferred income stream to their employees. The 'operating surplus of corporations' (OS_{CORP}) includes all corporate income after paying employees and expenses, and after depreciation, and can also be thought of as formal-sector capital income. The 'operating surplus of private unincorporated enterprises' (OS_{PUE}), or mixed income, includes income from self-employment, unincorporated professionals, household business owners, and the owners of private unincorporated enterprises (often informal enterprise, in developing countries). This income usually represents a return to both labor and capital. 'Household operating surplus' (OS_{HH}) is essentially the imputed rental income accruing to homeowners who live in their own home (so that a country entirely of homeowners would not appear to have a lower national income than a country where landlords earn rental income on all homes).

[45] This of course strengthens the support for later imputations, as well, which we will discuss below.

While these concepts above are sufficient for a simple decomposition of factor income into labor and capital income flows, it is worth also mentioning several concepts that do not figure into our calculations. First, 'net indirect taxes' (indirect taxes net of subsidies, NIT) are excluded from consideration. In our data these frequently comprise on the order of 10% to 15% of national income.[46] These indirect taxes are assumed to be factor-neutral, i.e., levied on the returns to capital and labor (their income) proportional to each factor's share in the affected production process.[47] In considering these basic national accounting concepts in the 'income approach' to domestic product, note also that government surplus is usually zero by construction and so does not have to be distributed to labor or capital.[48] Nonprofit institutions are included in the household sector.[49] Note also that all operating surpluses—of corporations, of unincorporated enterprises (viz. mixed income), and of the household sector (imputed rent)—are estimated net of depreciation (consumption of fixed capital, CFC). We consider that depreciation would not be rightly (or comparably, over time or across countries) regarded as an income accruing to capital as a factor of production. To include depreciation would overestimate the capital share. For this reason we use 'net domestic product' rather than 'gross domestic product' as our total income concept from the national accounts.

Finally, note that we do have available data on net foreign income—the net flow of labor and capital incomes (usually compensation of employees or corporate profits) from abroad—such that we could move from 'net domestic product' to 'net national income' as our total economy income concept. In practice, this seems to be the less relevant tax base for our sample (most countries tax income earned domestically regardless of citizenship, whereas net foreign income is taxed only with difficulty), but the adjustment does not affect our main results.

[46]The interquartile range for net indirect taxes runs from 8.1% to 14.9% of net national income.

[47]If some industries have different indirect tax rates, and different labor vs. capital splits within industry, we do not capture that effect in the aggregate data. It is partly for this reason that one might not consider a cross-country comparison of average effective tax rates as a sufficient statistic for comparing private-sector investment incentives across countries.

[48]Public sector enterprises that operate on the market are included in the corporate sector.

[49]Their employees earn labor income compensation, while the institutions themselves (especially trusts and institutions with significant endowments) do earn imputed rent from land ownership. In national accounts parlance, these non-governmental organizations are referred to as 'non-profit institutions *serving households*' (emphasis added).

Missing Data For values that are still missing after our work with the data sources and accounting identities noted above, we follow the imputation procedure from Blanchet and Chancel (2016). These authors put forward a straightforward, simple and transparent method for imputing consumption of fixed capital (depreciation) in the World Inequality Database (WID) series on national income (albeit without decomposing it into factor shares). In the same way as they observed three stylized facts about CFC to impute its missing values, so we also used a similar set of stylized facts about mixed income OS_{PUE} in order to statistically model its missing values. Our stylized facts for OS_{PUE} are as follows:

1. Mixed income (OS_{PUE}) tends to represent a lower share of national income in more developed countries (with a higher income per capita at purchasing power parity, PPP). This is an artifact of the outsized role of the informal economy and household enterprise in developing countries, where formalization, modernization, industrialization and/or corporatization are highly correlated, with each other and with income.

2. Some countries have structurally high (or low) levels of mixed income (as a share of national income)—whether due to long-term economic trends (e.g., the informal economy, as above, or the role of certain industries that that are themselves characterized by informality or household enterprise) or even due to a certain long-term tendency or practice in the national statistical office (e.g., in some countries many household enterprises are closely followed by tax authorities and therefore mapped by statisticians into the corporate sector; or vice-versa for countries where some quasi-corporations are poorly accounted for).

3. Mixed income as a share of national income is persistent and path-dependent, such that the value in year t will be closely correlated to the value in year $t+1$. Exogenous shocks that affect OS_{PUE} without similarly affecting total income per capita are rare, so its decrease is a slow process as income per capita increases.

Therefore, we can model OS_{PUE} as a function of log national income per capita (at PPP), with a random effect to capture constant country characteristics:

$$OS_{PUE_{it}} = \beta_0 + \beta_1 NNIpc_{it} + u_i + \varepsilon_{it}$$

where there is a random effect term u for each country i, and ε is the error term for each country-year it. To account for persistence in OS_{PUE} we model the error term ε_{it} as an AR(1) process:

$$\varepsilon_{it} = \rho\varepsilon_{i,t-1} + \eta_{it}$$

where η_{it} is i.i.d. white noise. As in Blanchet and Chancel (2016), when we know part of the OS_{PUE} series for a given country (observing it in later years), we can estimate the country's random effect u_i and use that in the imputation. When no later value of OS_{PUE} is observed, we assume $u_i = 0$. OS_{PUE} returns to its expected long-run value at a rate of ρ^t.

While mixed income OS_{PUE} is the most frequently missing national income component, others are occasionally missing, as well. When additional national income components are missing, we run the same imputation procedure for those. Therefore, we are able to generate an estimate of factor shares for every country-year in which we have data on total-economy net national income per capita.

The global coverage of factor share data, by source and including imputed values, is shown in Table 2.

A.2 Synthesizing comprehensive long-run tax revenue data

Data As described in Table 2, our tax revenue data draws from four types of sources: the OECD online Government Revenue Statistics (OECD, 2020); the ICTD and UNU-WIDER Government Revenue Dataset (ICTD/UNU-WIDER, 2020); the IMF Government Finance Statistics (IMF GFS, 2005); and, perhaps most importantly, what we can broadly be described as historical archive work. While our use of the first three is more or less self-explanatory, and we retrieve off-the-shelf data from these international organizations,[50] the fourth fountainhead is perhaps the most significant because it fills data gaps from these other three with extensive archival work to stitch together harmonized long-run series for the more than 150 countries.

Our archival work started in the Government Documents section of the Lamont Library archive at Harvard University. For each country in our data-set, we scanned, tabulated, processed, and unified official data from the public budget and national statistical yearbooks, to retrieve the official statistics on revenue collection even in remote historical eras. To complement hard-copy archival data, we also retrieved individual countries' online datasets, usually published by a national statistical office or ministry of finance, or sometimes cross-country datasets from international organizations and scholarly efforts (e.g., CIAT and IDB, 2019; Fisunoglu et al., 2011; Lotz and Morss, 1970). Lastly, we would fill remaining gaps by scouring scholarly sources for individual studies detailing the aggregate and disaggregated levels of (central and decentralized levels of government) tax revenue collection. To corroborate levels and trends of tax revenue across sources, in this way, also helped in our triangulation and harmonization effort discussed below. These sources—and their harmonization (see also 'stitching' section below)—are detailed country-by-country in a forthcoming compendium of country case studies.

Revenue Concepts We have streamlined diverse reporting of tax revenues (across sources, countries, time) according to the OECD Government Revenue Statistics classification of types of taxes (OECD, 2020). In this scheme, there is occasionally a four-digit level of detail to classify types of taxes, but for the purposes of our global long-run analysis of taxes on labor vs. capital — and in light of data limitations in non-OECD countries and in historical eras — we consider the simplified classification of Table 1 above.

The '1000' series of taxes, in the OECD framework, covers taxes on income and profits:

[50] Note, however, that we are using the IMF Historical Government Finance Statistics for the period 1972-89, not the version that is readily available online for the period since then. The historical dataset is available as a CD-ROM with dedicated software.

> This heading covers taxes levied on the net income or profits (i.e. gross income minus allowable tax reliefs) of individuals and enterprises. Also covered are taxes levied on the capital gains of individuals and enterprises. (OECD, 2020).

The most important distinction within the 1000 series is between personal income taxes (PIT, 1100) levied on individuals—which are largely attributable to labor income—and corporate income taxes (CIT, 1200) levied on corporate profits, taxes which are attributable to capital income. As discussed in Section 3.1.2, in non-OECD countries we do not usually directly observe the level of PIT revenue that derives from dividends, capital gains, or even certain types of property taxes (which would be attributed as a tax on capital income), so we apply the two basic assumptions (discussed in Section 3.1.2 above) to retrieve the benchmark measure. First, we estimate a PIT exemption threshold from Jensen (2021), where countries with a higher PIT exemption threshold are assumed to have a higher level of capital income within PIT than countries with a lower PIT exemption threshold; and second, we note countries where a dual-taxation PIT system has set the taxation of capital income within PIT to a measurably lower rate than the taxation of labor income within PIT.

The '2000' series of tax revenues, in our framework, includes both social security contributions and taxes on payroll and workforce:

> Classified here are all compulsory payments to general government that confer an entitlement to receive a (contingent) future social benefit. ... Contributions for the following types of social security benefits would, *inter alia,* be included: unemployment insurance benefits and supplements, accident, injury and sickness benefits, old-age, disability and survivors' pensions, family allowances, reimbursements for medical and hospital expenses or provision of hospital or medical services. Contributions may be levied on both employees and employers. ... Contributions to social insurance schemes which are not institutions of general government... are not considered as social security contributions. (OECD, 2020)

We also include here all payroll taxes, i.e., 'taxes payable by enterprises assessed either as a proportion of the wages or salaries paid or as a fixed amount per person employed.'

We consider all such social contributions and payroll taxes attributable as taxes on labor income, as shown in Table 1.

The '4000' series of tax revenue, in this framework, comprises property and wealth taxes:

> This heading covers recurrent and non-recurrent taxes on the use, ownership or transfer of property. These include taxes on immovable property or net wealth, taxes on the change of ownership of property through inheritance or gift and taxes on financial and capital transactions. (OECD, 2020)

We attribute these revenues as taxes on capital income.

The '5000' series, in the OECD (2020) framework, are indirect taxes on goods and services, including tariffs and trade taxes, value-added taxes, sales taxes, excises, and other forms of consumption taxes and taxes on production. Importantly, we consider these taxes both prior and proportional to factor income returns, such that the attribution of indirect taxes to factor incomes can be considered precisely proportional to factor incomes in the total economy.[51] If we were to include these taxes in our estimates of effective tax rates (see Section 3.1.2), we would only increase the levels of effective tax rates, but we would not change their trends relative to one another.

Finally, note that we exclude other nontax revenues and tax expenditures from consideration.

From these data sources, collected and merged according to the simplified OECD (2020) framework, we proceeded to stitch together long-run tax revenue series for all countries.

Stitching In brief, our comprehensive tax revenue data collection and harmonization process observed a simple set of rules.

When more than one source existed to document levels of tax revenue in a given country-year (with tax revenues disaggregated as in the classification above), we compare statistics in the years of overlap and prioritize sources as follows:

1. OECD revenue statistics

[51] In this sense, we also measure *factor-price* domestic product, i.e., net of indirect taxes.

2. If OECD data is unavailable, then we use HA (historical archive, government documents from Harvard University Library archives and elsewhere [see above]);

3. If remaining country-years are missing, we use complementary sources, including: ICTD/UNU-WIDER (2020) data that begins in 1980; IMF GFS (2005) that covers the period from 1972-1989; or, additional sources as described above. For social contributions in particular, two complementary sources provded helpful: the 'D61' statistic on social contributions in the household sector in UN SNA (1968) and UN SNA (2008) are accurate (to the best of our knowledge) and extensive in its coverage; and, the data from Fisunoglu et al. (2011) provide additional coverage based on manually digitized versions of offline IMF archival documents.

We always referred to a variety of scholarly sources for every country, to corroborate against each other and against these sources for the 'stitch' years across overlapping eras with data from multiple sources.

To check the robustness of long-run tax revenue series, country-by-country, and to minimally interpolate where necessary (rarely),[52] we applied the following rules:

1. We do not include a country in our sample if there are fewer than 10 years of observed tax revenue data.

2. We never interpolate more than four years without data during a time series.

3. Only in rare cases do we use more than two data sources per country.

4. If any discrepancies existed during 'stitch' years across data sources, we study the nature of the discrepancy in a review of scholarly sources—to then retrieve statistics on the revenues that are missing from the less-accurate source; and to 'backcast' using year-on-year growth rates of the less-accurate source, if necessary (see UNSD, 2018).

5. To check that we were not missing significant sources of decentralized tax revenues, we use the recent OECD-UCLG (2019) study to find the countries with significant

[52]We interpolated some tax revenue data in only 5% of nearly 7000 observations.

state and local public finance revenue collection, and collected more data for any countries where necessary.

6. We make explicit any assumptions about PIT vs. CIT split during historical eras where the government documents do not clarify the nature and relative weights of individual vs. corporate collection of income tax revenue. In these cases, we rely on local scholarly sources that discuss the legislation.

7. Similarly, in any (rare) cases where the OECD data did not measure the PIT vs. CIT split (and only ascribed certain revenues to a generic total income tax), we often find that actually the split can be made after studying the details of countries' tax legislation. This only occurs rarely in practice.

8. We check for significant regime change, political conflict, inflation episodes or other macro-economic crises that cast significant doubt on the credibility of estimates and the continuity of tax revenue time series. We do not interpolate between years characterized by such events.

9. We exclude all tax revenue observations for territories prior to independence as well as all countries during the Communist era.

Starting from this set of rules, we implement several checks on the quality of our factor income and tax revenue datasets. They are outlined on a country-by-country basis in the online appendix. As an illustration, below we reproduce the checks that were implemented for the tax revenue series in China.

Case Study of China To illustrate these rules above, note that we do not include tax revenue data for China before 1994, so we do not estimate China's effective tax rates on capital and labor income before 1994. Similar to our close examination of long run revenue series for all countries in our data—but with particular attention to the case of China both for its national income (and population) weight and for its relative political economy complexity—we took the decision to ignore pre-1994 revenue data in China after a close reading of the country's revenue statistics, public finance laws, and scholarly sources.

After China's economic reforms began in 1978, one could argue that 1983-85 (but no earlier) marked the transition to a corporate income tax in China (Wong and Bird, 2008). Decree from the State Council in 1983 put a new 55% tax on profits of enterprises, which were still almost entirely state-owned (SOEs run by either the central or local governments).[53] This tax policy replaced the command economy default of 'profit delivery,' in which the state simply spent from profitable SOE revenues and subsidized unprofitable SOEs. We observe this CIT imposition in our data, where there is no revenue that is or can properly be called a CIT until 1985—and then the CIT is massive (as a percentage of domestic product, or of capital income, or of corporate operating surplus). However, not long after this policy was issued, it was reformed into a 'fiscal contracting' system whereby SOEs negotiated with the government to promise certain levels of revenue (regardless of profit) to the local governments, who negotiated to deliver certain levels of revenue to the central government. This is not exactly a tax system, even if some data sources have called it as such.[54] Meanwhile, the overall level of revenue to domestic product bottomed out as China's economy grew but revenue did not.

China's major tax reform in 1994 was prompted in large part by low central government revenues: a low tax-to-GDP ratio and a low central-vs.-decentralized revenue ratio (World Bank, 2008). The reform created their modern tax system: It established a central tax administration to centralize tax collection; reformed the broken 'fiscal contracting' CIT system; unified the PIT; created a VAT; and reduced distortions and 'extra budgetary' (non-tax) revenues.[55]

For the purposes of our analysis and long-run exposition on tax progressivity and factor income taxation, the 'fiscal modernization' reforms of 1994 make that year the most salient entry-point for China into our data-set. These points of reform match most closely our methodological framework, which seeks to exclude non-tax public finance revenues.

[53] Although some revenue data begins in 1950, 1985 is the first year for which a 'corporate income tax' appears in the official public finance statistics available online at https://data.stats.gov.cn/english/index.htm.

[54] For these reasons, we exclude this sort of 'pseudo'-CIT revenues dating from 1985 through 1993. More generally, for similar reasons we exclude entirely a long-run tax revenue series for China that could actually start in 1950, but which we are unable to harmonize conceptually with our framework for measuring specifically *tax* revenues, and especially direct taxes.

[55] On these points the state-run media China Daily (2018) agrees with scholarly contributions in World Bank (2008) and elsewhere.

Appendix B Instrumental variables

In this section, we outline the construction of the two instrumental variables. Since both instruments are from Egger, Nigai, and Strecker (2019), more detail can be found in their publication.

B.1 Instrument based on quantitative trade models

The first instrument leverages the general model structure of gravity models in general equilibrium. These models permit the calibration of country pair-year-specific trade costs from trade data. Models that use such models include Eaton and Kortum (2002); Anderson and Wincoop (2003); Arkolakis, Costinot, and Rodriguez-Clare (2012). These models rely on three specific assumptions. First, producers are either perfectly competitive and make zero profits or chrge a constant mark-up; prices are independent of the location of customers. Second, trade costs take the iceberg-form. Third, aggregate expenditure and its allocation across products can be separated through a two-stage budgeting decision. These three assumptions imply that bilateral consumption shares towards country j by consumers in country i in year t, denoted π_{ijt}, are multiplicative components that are exporter-year-specific (e_{jt}), importer-year-specific (ι_{it}) and pair-year specific (β_{ijt}) as follows

$$\pi_{ijt} = \quad e_{jt} \quad \times \quad \iota_{it} \quad \times \quad \beta_{ijt}$$

The component e_{jt} is proportional to country j's suply potential and broadly captures production costs and gross-of-tax factor income and is influenced possibly by both capital and labor taxation. The component ι_{it} is a function of the consumer price index, which varies across years and countries. The key intuition is that both e_{jt} and ι_{it} may capture country-year-specific factors, but the country pair-specific component β_{ijt} is free of any country-year specific factor. Instead, β_{ijt} captures trade frictions across country-pairs and time. Egger, Nigai, and Strecker (2019) exploit the multiplicative model structure about π_{ijt} to recover measures of β_{ijt}. The authors assume that transaction costs between domestic sellers and customers is constant, such that $\beta_{iit} = 1$. Both the importer-year component and exporter-year components can then be eliminated by normalizing import and export trade shares by the importer and exporters' consumption from domestic sellers. In turn, the product of the normalized shares gives the bilateral fractions of importing-exporting country-pairs at a point in time:

$$\frac{\pi_{ijt}}{\pi_{iit}} \cdot \frac{\pi_{jit}}{\pi_{jjt}} = \beta_{ijt} \cdot \beta_{jit}$$

Finally, the sum of these costs $Z_{it}^{gravity} = \sum[\beta_{ijt} \cdot \beta_{jit}]$ measures total trade frictions for country i in year t and constitutes the instrument. Note that all exporter-year and importer-year factors are removed from the instrument. This instrument is valid if the *distribution* of trade costs among country-pairs (not its level) is not influenced by e.g. changes in factor income shares or domestic labor and capital tax revenues. To construct this instrument requires data on country-pair trade flows: we use the UN COMTRADE database to construct as large a sample as possible with non-missing values for bilateral consumption shares.

The first-stage regressions with $Z_{it}^{gravity}$ are reported in Appendix Table A11. The instrument is slightly stronger when using the log-level of trade or the share of goods-trade as the endogenous measure of openness.

B.2 Instrument based on global oil prices and transportation distances

Our second instrument exploits spatial heterogeneity across countries in a way that responds to oil price shocks. To build this instrument we require two parameters: global oil prices over time, and within-country transportation distances. For the former, we collect the long-run world price of oil from the well known 'OPEC Reference Basket' tracking crude oil prices (OPEC, 2021). For the latter, we measure transportation distances from the three largest three cities (from UN DESA, 2019) to their nearest ports,[56] using international shipping logistics calculators at SeaRates (2021). To calculate the distance of each city to

[56]One could measure the distance from city to the nearest *sea* port, or to the nearest container terminal of any kind. We made both measurements and ultimately make use of the latter. However, there is usually little difference when calculating the within-country variance of these measurements across cities, and this difference does not affect our results.

each port, we look at the map of the city and its distance to port, through the lens of these shipping logistics calculators. We manually enter and record the distances by road for each city to its nearest port. These distances vary within a country to the extent that a country is far from a port, and to the extent that cities (far from the port) are also far from one another.

We then take the variance of the oil price p_t^{oil} $\times$ distance d_i^k for each city k in country i and year t:

$$Z_{it}^{pricedist} = \frac{1}{2}\sum_{k=1}^{3}[(p_t d_i^k - p_t \overline{d_i})^2]$$

This variance is increasing in countries whose principal population centers are far from the nearest port *and*, more importantly, from each other, which implies a larger shock to transportation costs in spread-out (and far from the port) countries than in countries with concentrated populations (near to port).[57] It is this shock to trade that our instrument captures. Alternatively, we can measure the variance in distance and then multiply it by the global price. The distribution of the variance instrument $Z_{it}^{pricedist}$ across country-years would not change; the only impact would be a level-shift in factor *p*. While we consider the main approach to more closely capture the sensitivity of spatial concentration to shocks in transportation costs, results remain similar using this alternative measurement approach (results available upon request).

This second instrument is very different from the first instrument since it does hinge on any theoretical assumptions and is valid under very different assumptions. Specifically, it relies on the assumption that the distribution of trade-costs induced by global oil price shocks across countries with different domestic trade networks is not correlated with contemporaneous changes in factor shares and tax revenues. Importantly, we verify that the main results are robust to allowing major oil producing countries to be on a separate, non-parametric time-path throughout the sample period.

[57]In this sense, our measure of the variance here does not here penalize a country for having its nearest port outside the national borders, or even thousands of kilometers away. The $Z_{it}^{pricedist}$ variance increases when distance-to-port varies within a country, not when the average distance-to-port increases. One could perhaps imagine an even stronger trade instrument that accounts for cross-country variance in the average distance-to-port with an additional parameter. However, in the first-stage results discussed here, we find sufficient evidence of a trade effect merely on the within-country variance of distance. An additional trade outcome not discussed here, of course, is on domestic trade between cities. We only measure an international trade effect, but one would similarly expect domestic trade to decrease as transportation costs increase.

Constructing the instrument $Z_{it}^{pricedist}$ as described above, the first stage results are presented in Appendix Table A11. The instrument is strongly associated with changes in trade within-country over time and is robust to using various measures of trade openness.

Appendix C Event studies

C.1 Description of events

Our selection of trade episodes is determined by three criteria: (i) that the event can unambiguously be related to *measurable* policy reforms; (ii) that the policy reforms induced *large* changes in trade barriers; and (iii) that the event has been studied in peer-reviewed academic publications. The first criterion improves the transparency of the event-study design, which rely on changes in outcomes around an explicitly defined event in time which can be attributed to a policy-reform. The second criterion increases the likelihood that we will be able to graphically observe sharp breaks in trend in our macroeconomic outcomes which coincide with the timing of the policy event. The third criterion increases the external validity of our results, since these prior papers have already established positive effects of the reforms on cross-border trade (and other economic outcomes).

Selection of events These criteria lead us to focus on the six trade liberalization events referenced in review articles by Goldberg and Pavcnik (2007) and Goldberg and Pavcnik (2016), GP henceforth, to which we add China's WTO accession event (studied in Brandt et al., 2017). These liberalization events all feature reductions in tariff barriers, the most commonly studied component of globalization, in part because they are easier to measure on a consistent basis across space and time compared to other forms of globalization (Goldberg and Pavcnik, 2007). These events also features reductions in non-tariff barriers which are harder to measure; for example the number of products subject to import licences and quotas fell. Fortunately, tariff and non-tariff barrier reductions seem highly correlated (Goldberg and Pavcnik, 2007).

All selected events feature reductions in policy-induced trade costs, primarily through the reduction in tariff rates. Most of the selected countries did not participate in the early

GATT/WTO negotiation rounds; consequently, tariffs remained high in these countries at the onset of the events, such that reductions in tariff rates remained a policy lever that was available to the governments.[58] These policy-induced liberalization events were drastic: Brazil cut tariff rates from 58.8 percent to 15.4 percent; India reduced rates from 80 percent to 39 percent; China reduced tariffs from 48% on average to 20%; Mexico reduced tariff rates from 23.5% to 11.8%, while import licence requirements went from covering 92.5% of national production to 25.4%; Colombia's tariffs were reduced from 27% to 10% and import requirements dropped from 72% of national production coverage to 1.1%. In the selected countries, "tariff reductions constitute a 'big part' of the globalization process" (Goldberg and Pavcnik, 2016). The timing of these events, as well as references to additional papers which study the events in detail is provided in Appendix Table A2.

Discussion of reform-timing and post-reform periods Most studies provide detailed discussions of the context surrounding the events. We repeat below the rationale cited for the liberalization events we use, and discuss why these events are plausibly exogenous to the country's economic circumstances at the time.

The Brazilian liberalization event of 1988 is detailed in Dix-Carneiro and Kovak (2017). The authors note that the high pre-reform average level of tariffs was driven by large cross-industry variation in protectionism:

> "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 licences and removed nearly all of the remaining special customs regimes." (Dix-Carneiro and Kovak, 2017)

Similarly, tariff reductions in Colombia in 1985 were driven by the country's commitment to impose uniform rates under the negotiation commitments to the WTO. In Colombia's case, Goldberg and Pavcnik (2007) note that the reform objective was to reduce cross-

[58]Some countries were not GATT members by the time of the event, such as Mexico; others (Brazil, Colombia, India) were nominal GATT members, but were not forced to reciprocate tariff concessions negotiated with GATT until the Uruguay Round (Goldberg and Pavcnik, 2016).

industry dispersion under WTO negotiations, thereby making "the endogeneity of trade policy changes less pronounced here [in Colombia] than in other studies."

The 2001 reform in Vietnam was implemented as a broad trade agreement that did not involve negotiations over specific tariffs (McCaig and Pavcnik, 2018). The reform was driven by the American government's decision to reclassify Vietnam from 'Column 2' of the US tariff schedule to the 'Normal Trade Relations' schedule. 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. McCaig and Pavcnik (2018) show that there are no differential trends between Vietnamese exports to the US relative to exports to other high-income countries. The exogeneity of Vietnam's case is also compelling, as the liberalization even was triggered by a foreign party, rather than by its own government.

India's 1991 event occurred as a result of an IMF intervention that dictated the pace and scope of the liberalization reforms. Under the IMF program, the 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) provide an extensive discussion of the Indian reform, arguing that it "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 that forced the reform was in response to India's balance of payment crisis, which was triggered by"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" (Topalova and Khandelwal, 2011).

In China, Brandt et al. (2017) note that trade openness reforms had gradually been implemented in China prior to the country's WTO accession event, but that the tariff reductions implemented in the immediate post-accession period were large, "less voluntary" and largely complied with the fixed WTO accession agreements. Importantly, the potential accession to WTO contributed to timing of the privatization initiatives in the pre-WTO years, in which the Chinese government restructured and reduced its ownership in state-owned enterprises. While the privatization efforts began in 1995 and were also gradual (Jefferson, 2016), given their importance in the national economy, it is possible that additional sell-offs in the immediate post-WTO years contribute to the observed break in trends in our main outcomes. This is particularly true for the capital share of corporate income and the effective tax rate on capital.

These descriptions of reform context do not argue that liberalization events were triggered by trends in taxation needs or changes in relative growth of capital versus labor income. Thus, this narrative analysis complements the absence of a pre-trends result (Figure 7), to help guard against endogeity concerns in the timing of the events.

Even if the reform timing is uncorrelated with confounding trends, the interpretation of the event studies still depends on whether other reforms and macroeconomic shocks occurred in the immediate post-reform years. The detailed review provided in (Goldberg and Pavcnik, 2007) is very helpful, as it notes liberalization events which followed the initial important events: Mexico's 1985 trade event is followed by a removal of capital inflow restrictions in 1989 and accession to NAFTA in 1994; Argentina's event in 1989 is followed by accession to Mercosur in 1991; Brazil's event in 1988 is followed by accession to Mercosur in 1991 and the currency crisis in 1998; and India's event in 1991 is followed by foreign direct investment liberalization in 1993. Their discussion, then, suggests that the majority of other reforms occurred with a few years lag relative to the trade liberalization event; and, with the exception of the Chinese domestic privatization efforts, these other reforms served to reduce other non-tariff barriers to cross-border flows. As such, while the immediate post-event impacts may more likely be attributed to trade liberalization, the medium- and long-run impacts should be interpreted as the reduced-form effects of globalization more generally, which includes an the international flow of goods, services, and capitalflows.

C.2 Methodology

Sample Construction Our sample is constructed by first applying a synthetic matching procedure to every treated country, for every individual outcome of interest. The donor pool (the set of all control countries from which to chose the synthetic control group) has to be fully balanced in all pre-event periods. Thus, we discard all countries with small data gaps before 1976. This gives us a sample of 103 countries for each outcome.[59] We have ten 'pre' event years for every country, except for Mexico and Colombia, where we have nine pre-period years.[60] We then pool together all seven treated countries and their seven synthetic control units.

Empirical Strategy Using this panel, we estimate the following event study regression:

$$Y_{it} = \sum_{j=-10, j\neq -1}^{10} \mu_j * \mathbb{1}(j=t)_t * D_i + \theta_t + \kappa_i + \pi_{Year(it)} + \epsilon_{it}$$

where θ_t and κ_i represent, respectively, time relative to/from the event and country fixed effects, and where $\pi_{Year(it)}$ is a set of fixed effects for calendar years. D_i is a dummy equal to one if observation i is a treated country. Hence, μ_j capture the difference between the treated countries and the group of synthetic controls across event time, with year t-1 as the reference period.

In addition to the event study regressions, we also use this setting to estimate a simple difference-in-difference coefficient:

$$Y_{it} = \mu^{DiD} * \mathbb{1}(j \geq 0)_t * D_i + \theta_t + \kappa_i + \pi_{Year(it)} + \epsilon_{it}$$

Here, μ^{DiD} can be interpreted as an average treatment effect over the first 10 years post treatment. We run both regressions—the event study and the difference-in-difference regression—on a set of nine outcomes, and cluster standard errors at the country level. We note that statistical inference based a small sample size should be approached with caution (Abadie, Diamond, and Hainmueller, 2010). We therefore also report standard errors based on the wild bootstrap method (Cameron, Gelbach, and Miller, 2008) for μ^{DiD}, in Table A3.

Moreover, we use the imputation method developed by Borusyak, Jaravel, and Spiess (2021) to report average treatment effects comparable to μ^{DiD} with a technique that deals with issues with two-way fixed effects and heterogeneous timing of events (such as our

[59]The exception to this rule is the trade variable, where we have more data gaps and consequently must drop more countries from the donor pool. Here, the donor pool consists of 90 countries.

[60]Moreover, Vietnam only has tax revenue data from 1994 onward, but its event happens in 2001. We therefore only match on six pre-period years.

setting). The approach provides a transparent alternative method to the difference-in-difference equation specified above. The average treatment effect τ is calculated in three steps, detailed below.

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

$$Y_{it} = \theta_t + \kappa_i + \pi_{Year(it)} + \epsilon_{it}$$

if $t < 0$ or $D_i = 0$. We note that to bring the approach developed by Borusyak, Jaravel, and Spiess (2021) closer to our estimation strategy, we include fixed effects for the years in relative time.

Second, equipped with the fitted values $\hat{\theta}_t$ and $\hat{\kappa}_i$, we can now impute the unit specific treatment effect as:

$$\hat{\tau}_{it} = Y_{it} - \hat{\theta}_t - \hat{\kappa}_i - \hat{\pi}_{Year(it)}$$

Finally, the third 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.

Simultaneously Matching on Main Outcomes As we perform the synthetic matching procedure for each event and outcome based on the outcome, we have a different 'synthetic' control for each country in every outcome. This means that while we use the same group of treated countries in every regression, the set of control countries that feeds into the synthetic control group varies across outcomes. We want to test that our results hold up in a more rigid setting.

To ensure robustness to a more restrictive synthetic control, we repeat all of the above analysis, using a slightly different approach of constructing the synthetic control countries. Specifically, we use our four main outcomes—trade (as a percentage of domestic product); capital share of domestic product; and effective tax rate on labor and capital—to predict one synthetic control country per treated country. The resulting weights for the three most prominent countries in each control group are reported in Table A1. This exercise allows us to still run separate regressions for each outcome, but with the exact same composition of the control group in each regression.

Level and event study graphs for each outcome are reported in Figure A9. While it is obvious that for some outcomes, the levels do not line up exactly in the 'pre' periods, the graphs confirm the conclusions from Figure 7. In Table A4, we also report all estimated coefficients (including the simple difference-in-difference estimate) of this alternative approach. While we prefer the approach presented in the main body of this paper in Section 6.1, this robustness exercise does demonstrate that the results there do not hinge on the flexible nature of our synthetic control design.