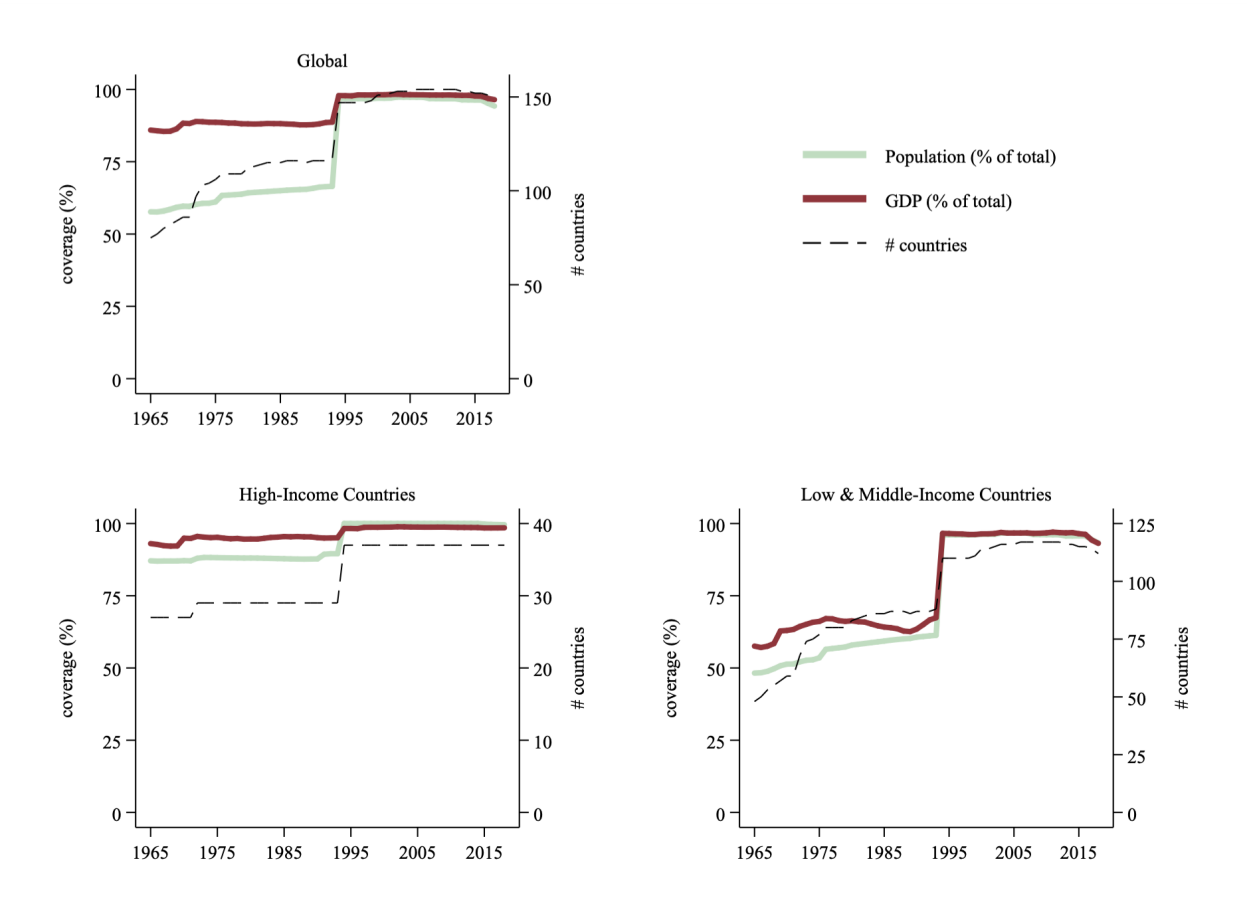
Appendices

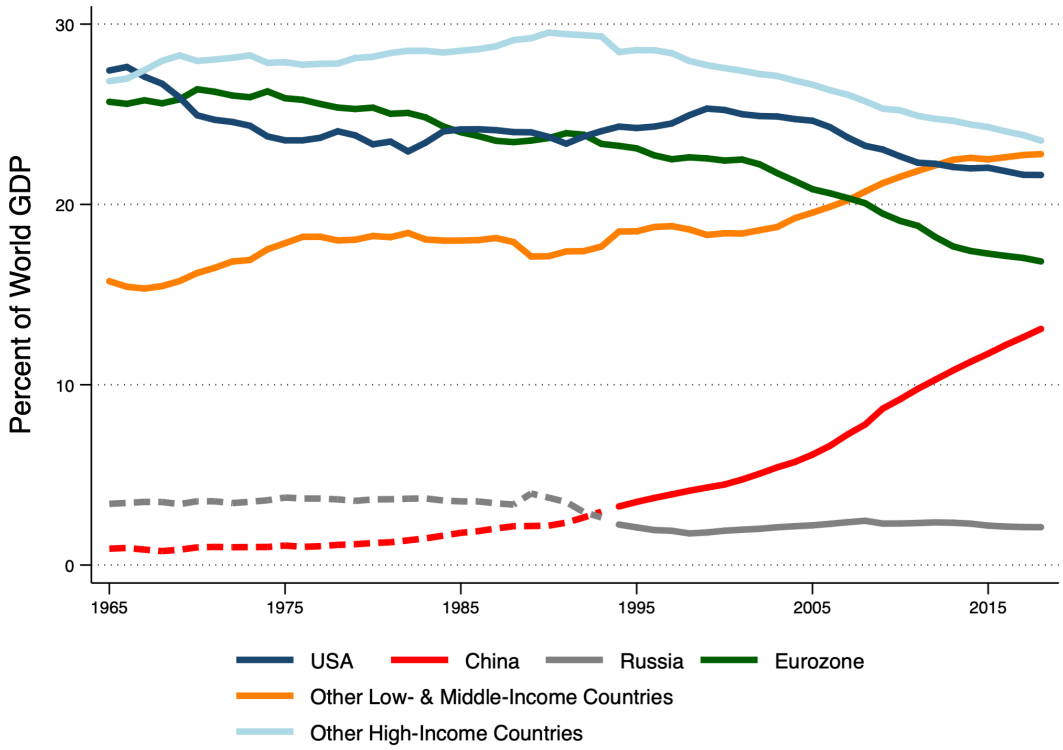
Appendix A Additional Figures and Tables

Figure A1: Data Coverage for Effective Tax Rates



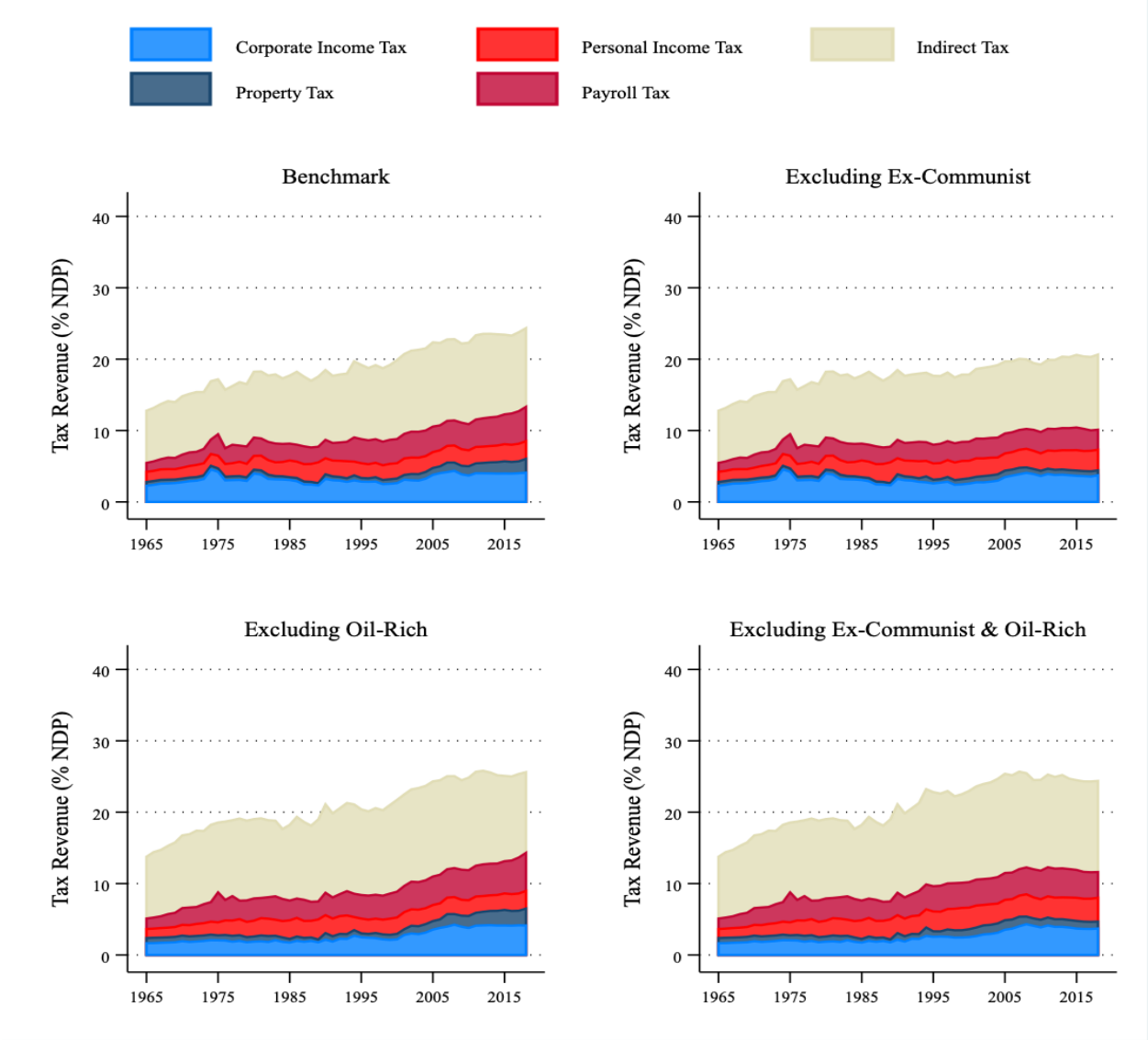
Notes: This figure shows the coverage of our effective tax rate data between 1965 and 2018, globally and for high vs. low- and middle-income countries. The solid lines plot the percentage of total population and GDP that is covered in our data (left axis). The dashed lines show the number of countries 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 B). Other missing country-years correspond to conflicts, independence post 1965, and in a few cases to missing data.

Figure A2: Share of Worldwide GDP by Country



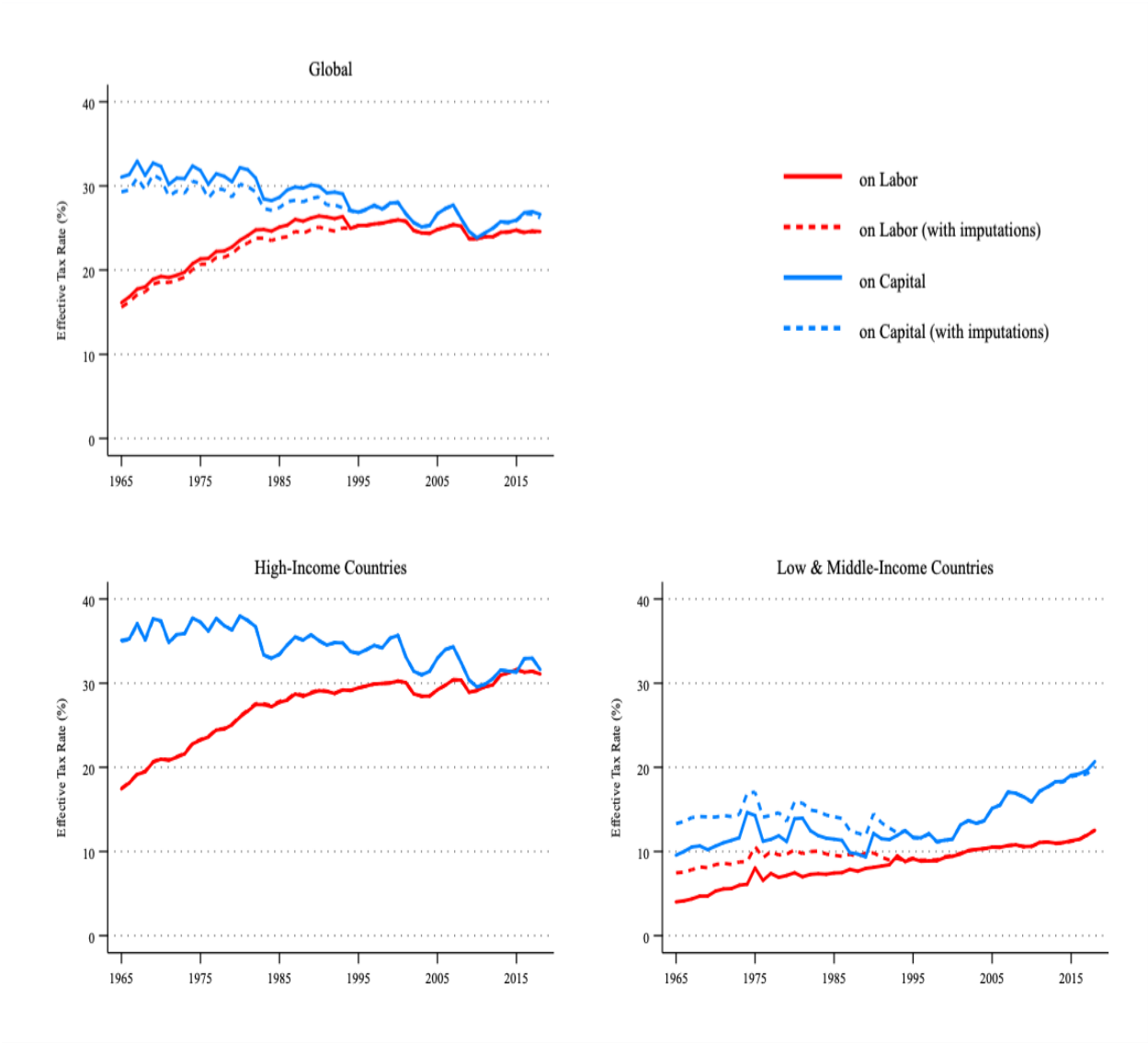
Notes: This figure plots the relative weight in world GDP of important countries. The countries are the USA, Eurozone 19, China, Russia, the remaining high-income countries and the remaining low and middle-income countries (less China and Russia). By construction these lines add to 100% of world GDP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

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



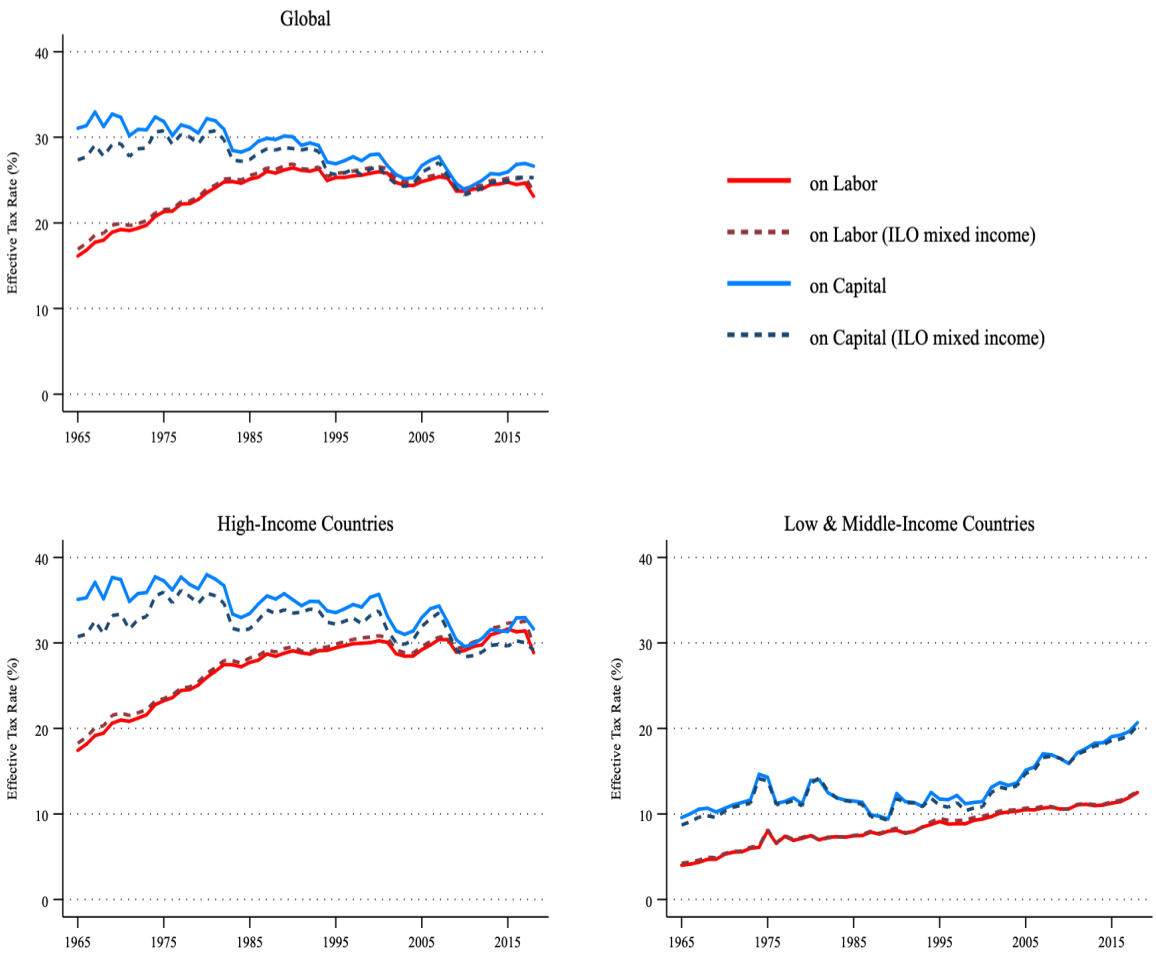
Notes: This figure plots the time series of tax revenue as a share of net domestic product (NDP), separated by revenue source, for the 119 low- and middle-income countries (sample countries that are not high-income OECD countries), weighting country-year observations by their share in that year's total 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 $ETR^K_{CHN,1993} = ETR^K_{CHN,1994} + \hat{\beta}_{1993} - \hat{\beta}_{1994}$.

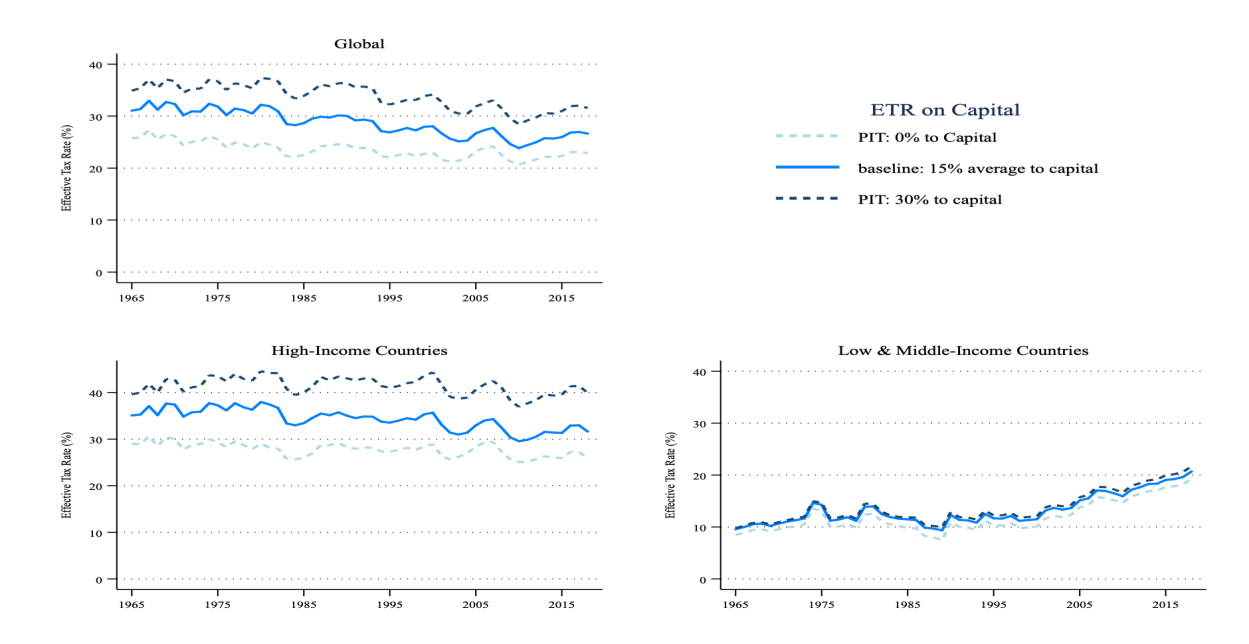
Figure A5: Impact of Mixed Income Allocation on Effective Tax Rates



Notes: This figure plots the time series of effective tax rates for labor and capital varying the method which assigns mixed income in national accounts (self-employment income) to labor vs capital. The benchmark method assigns 70% of mixed-income to labor and 30% to capital. The alternative method (ILO mixed income) imputes to the self-employed a labor income similar to the wage they would have earned in an employer-employee relation, based on observable characteristics. In practice, we extend the estimates in ILO (2019), using ILOSTAT (2021) data on self-employment shares of the workforce, to all countries since 1991, and impute the series backwards to complete it. Further details are in section 3.1.1 and the data Appendix B.

Figure A6: Impact of Personal Income Taxes Allocation on Effective Tax Rates

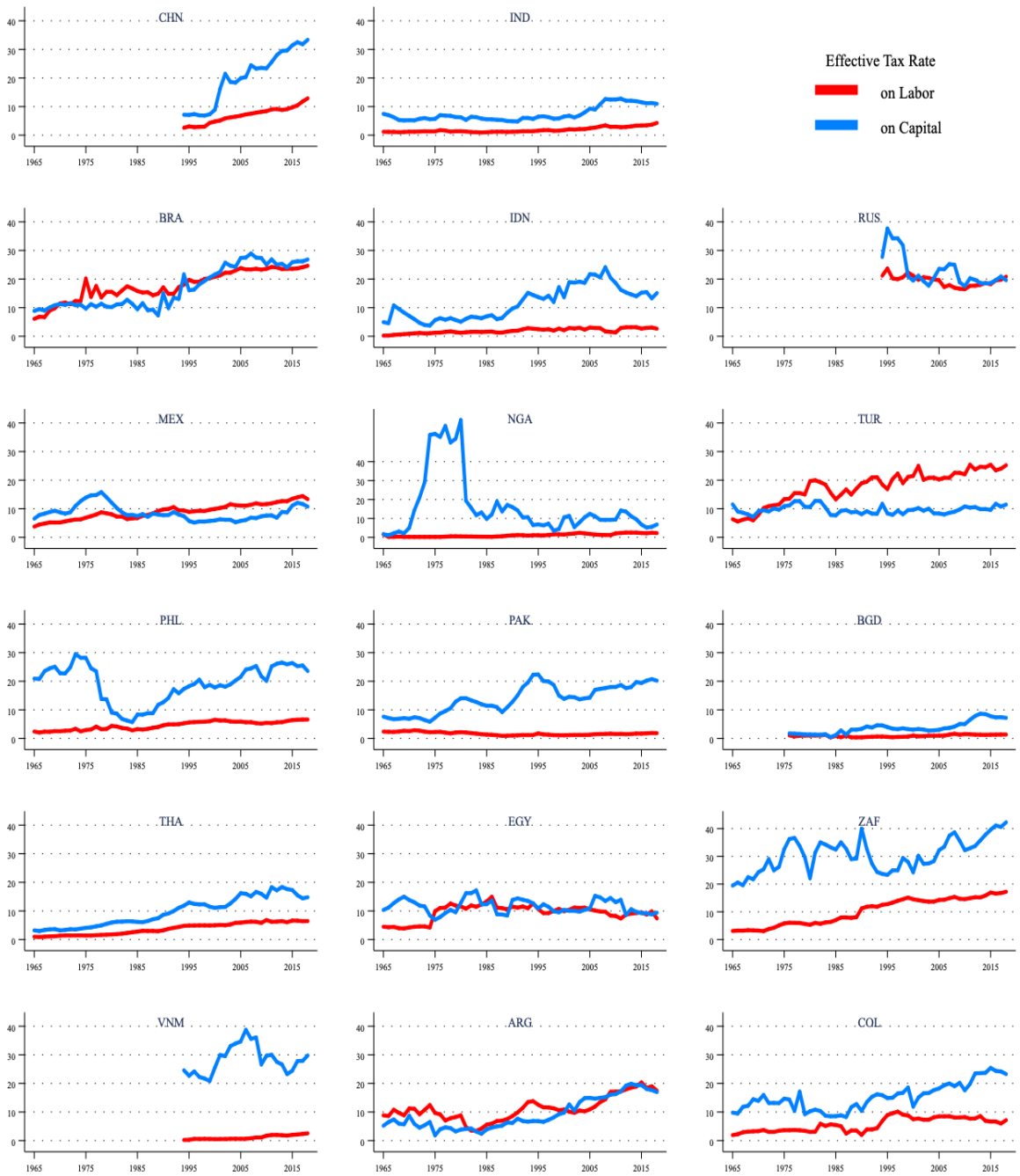
(a) ETR on Capital



(b) ETR on Labor

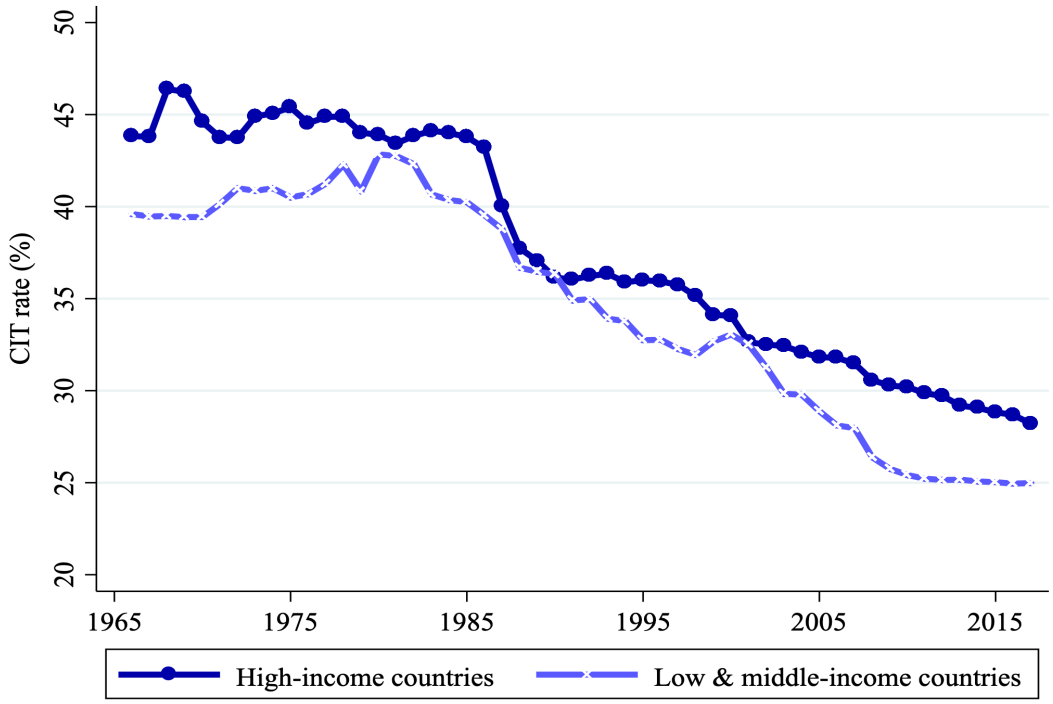
Notes: These figures plot the time series of effective tax rates for capital (panel a) and labor (panel b) as a function of how we allocate personal income tax (PIT) revenue to each factor. Tax revenue from PIT represents a mix of taxes on labor and capital, since it includes salaries, as well as capital income and capital gains. Our baseline, consistent with US data, starts by assuming that 15% of PIT revenue derives from capital and 85% from labor. It is then adjusted to take into account that some forms of capital income face reduced rates compared to labor income, and that in some countries the PIT only covers the very top of the income distribution (see Section 3.1.2). We then show two extreme bound scenarios: as an upper bound for the ETR on capital, we assign 30% of PIT tax revenue to capital and 70% to labor. As a lower bound we assign 0% to capital and 100% to labor.

Figure A7: Effective Tax Rates in Large Developing Countries



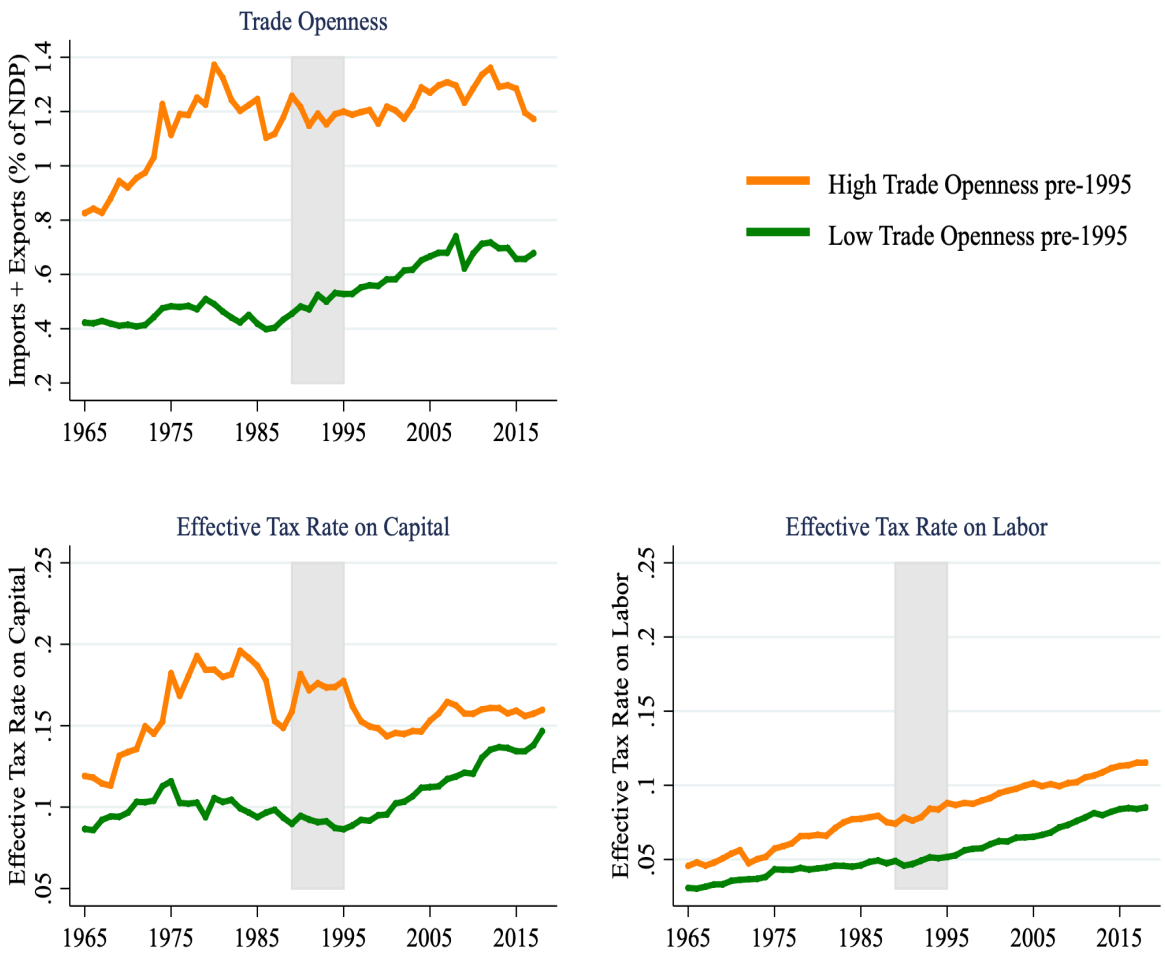
Notes: This figure shows the evolution of effective tax rates on labor and capital for the 17 largest low and middle income countries. Countries are displayed when they rank in the top 20 both in terms of population and GDP, in 2018.

Figure A8: Trends in statutory corporate income tax rates by development level



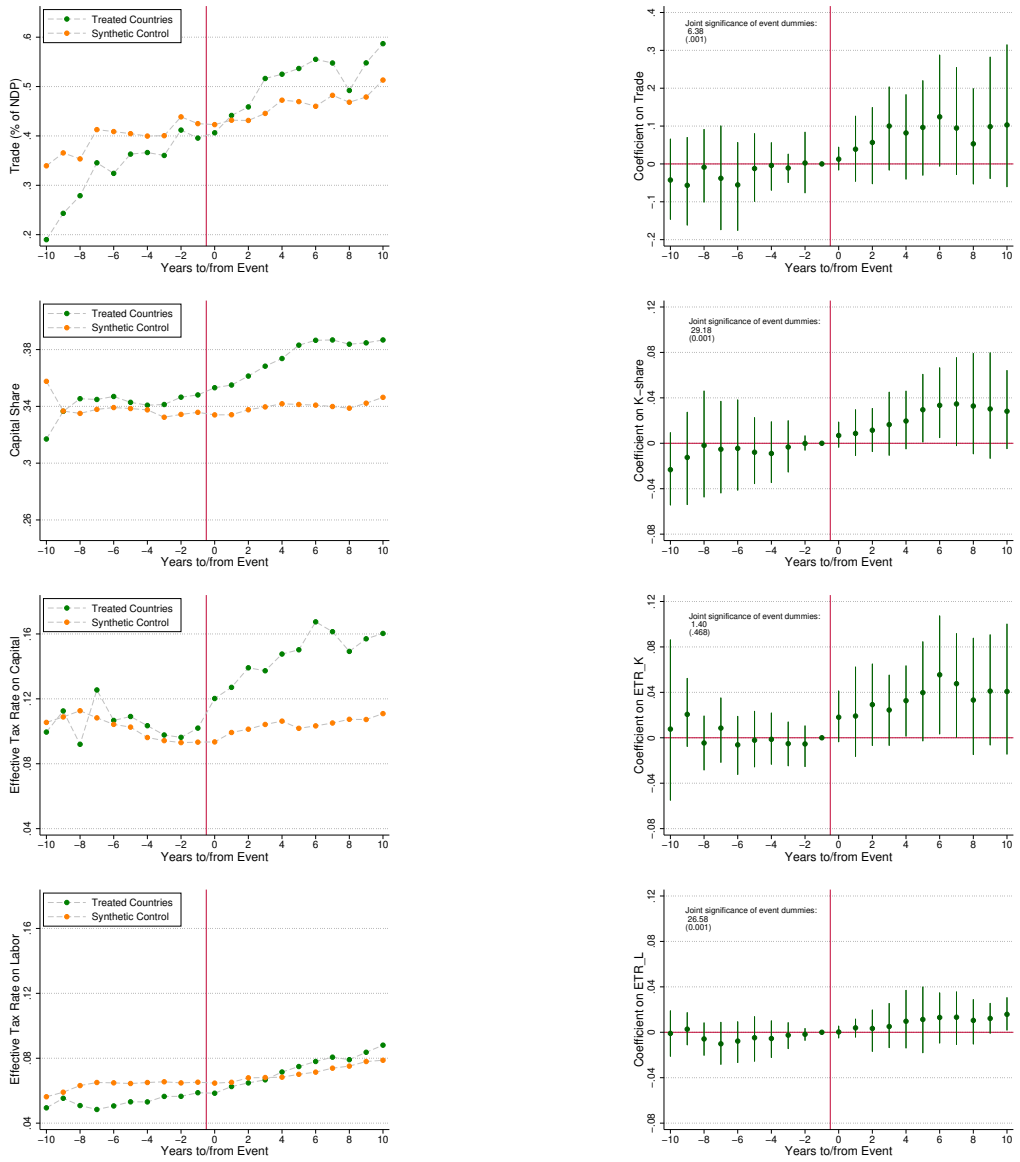
Notes: This figure plot the time series of statutory corporate income tax rates by level of development: the light blue X-line is for middle and low income countries, while the dark blue circle-line is for high income countries. Each line plots the year fixed effects from an unweighted OLS regression, in the relevant sub-sample of countries, 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.

Figure A9: Trends in Factor Taxation by Initial Trade Openness, in Developing Countries



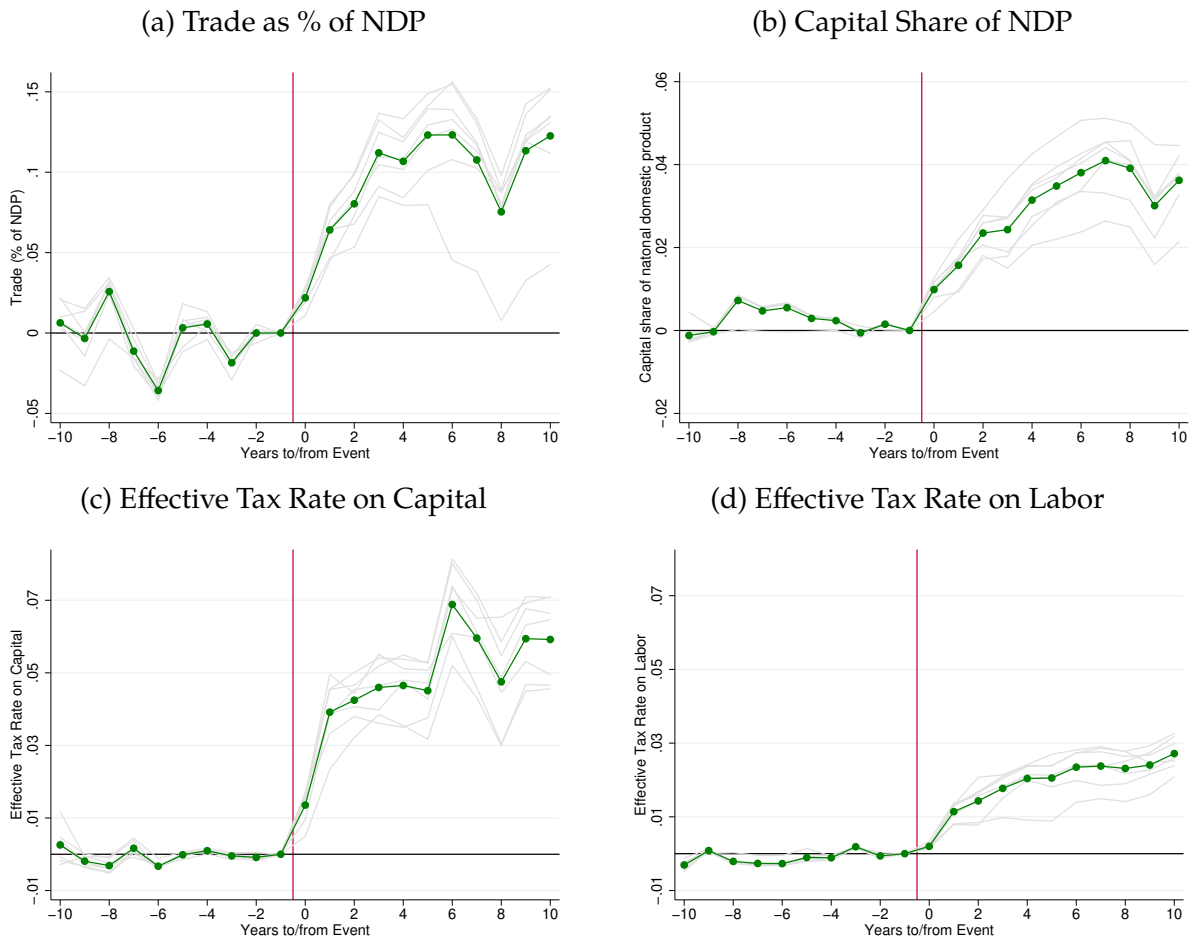
Notes: This figure plot the time series of trade openness (top-left panel), average effective tax rates on capital (top-right panel) and labor (bottom-left panel). The sample is limited to low- and middle-income countries. Within each panel, the green line (orange line) traces the evolution of the group 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 A10: Event Studies of Trade Liberalization, Simultaneous Matching on Outcomes



Notes: These figures show event-studies for trade liberalization reforms in seven countries, over four outcomes: trade (as a percentage of domestic product); capital shares; 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.

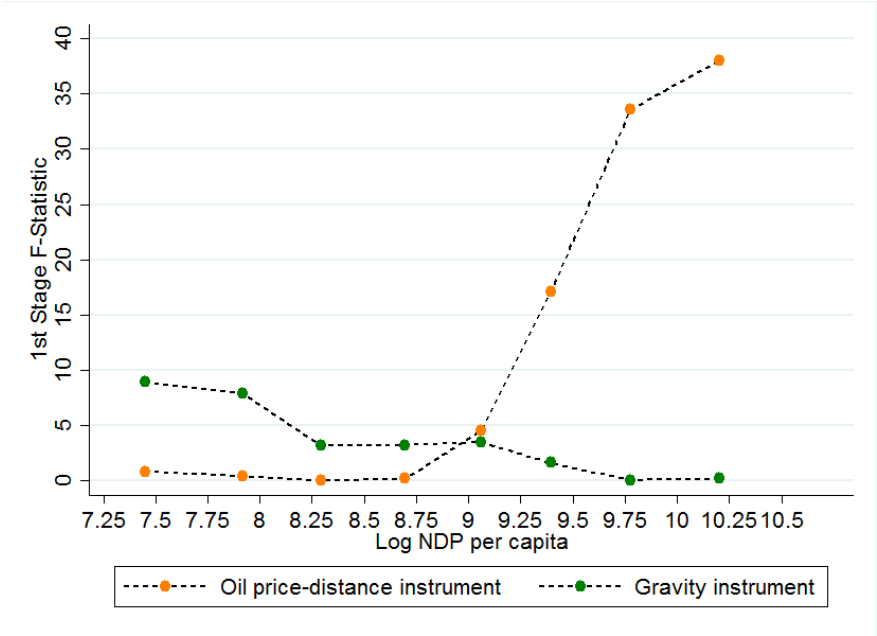
Figure A11: Robustness of Trade Liberalization to Changing Sample of Event Countries



Notes: These figures show event studies for trade liberalization reforms in seven countries, over four outcomes: trade (as a percentage of domestic product); capital share (as a percentage of domestic product); and effective tax rates on capital and labor. In each figure, the solid green 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 A12: Strength of Individual Instruments Across Subsamples

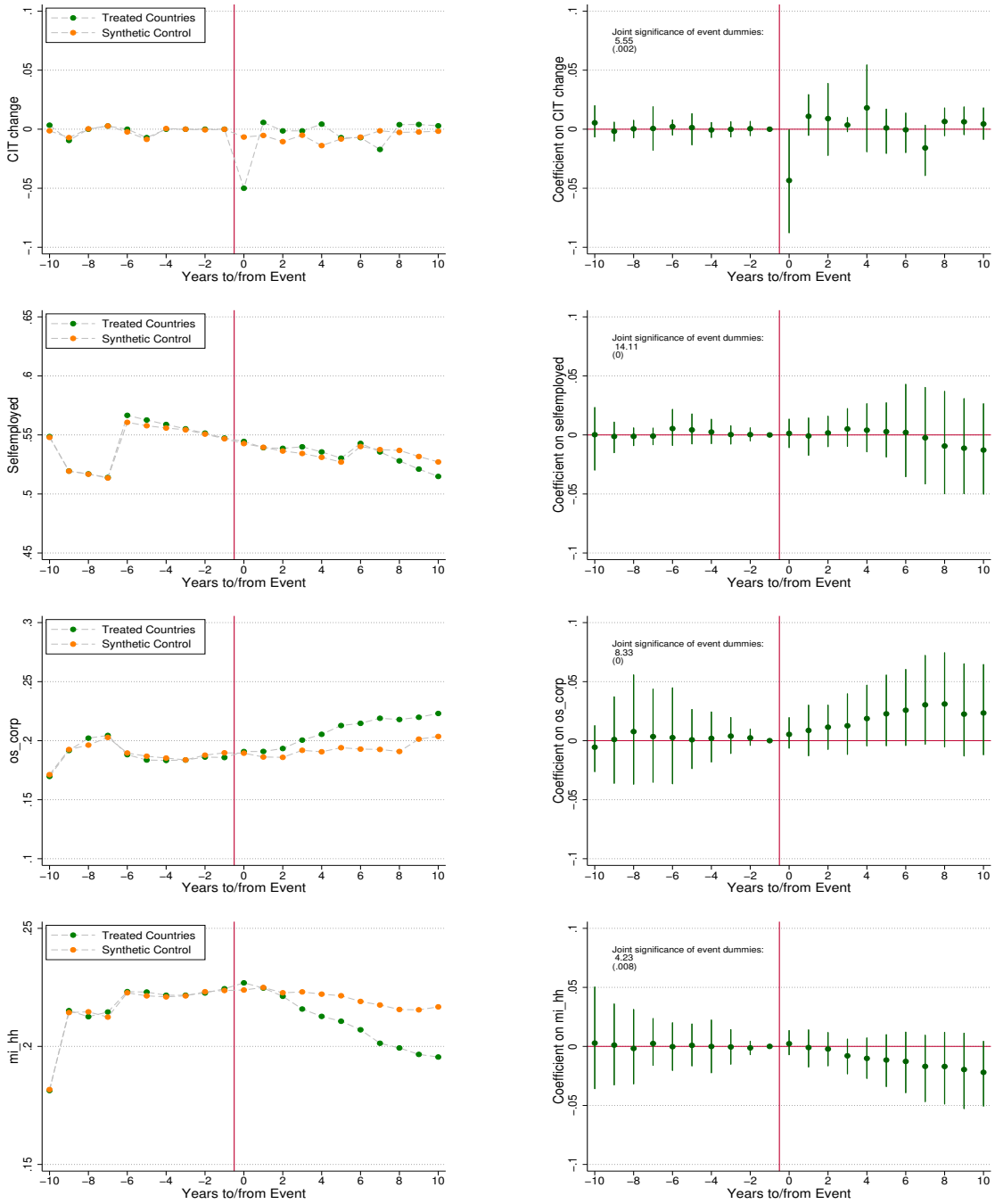
(a) Sub-samples of NDP per capita



(b) Sub-samples of time-periods

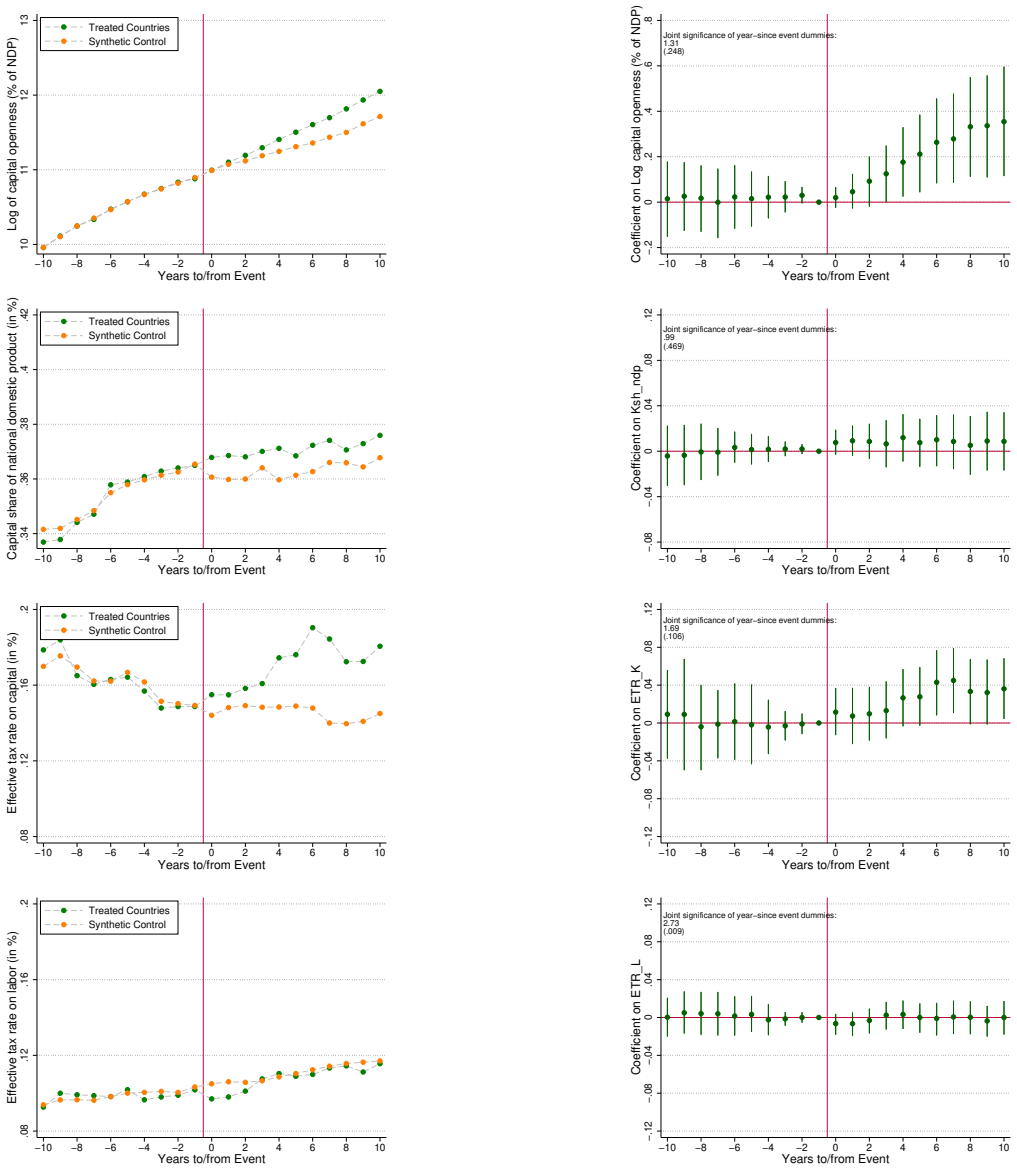
Notes: These figures show the individual statistical strength of the two instruments, denoted 'oil-distance' and 'gravity'. The outcome is the first-stage F-statistic from a regression of trade openness on the individual instruments (see Section 7). The outcome is shown across subsamples of log GDP per 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: Event Study of Trade Liberalization – Mechanism Outcomes



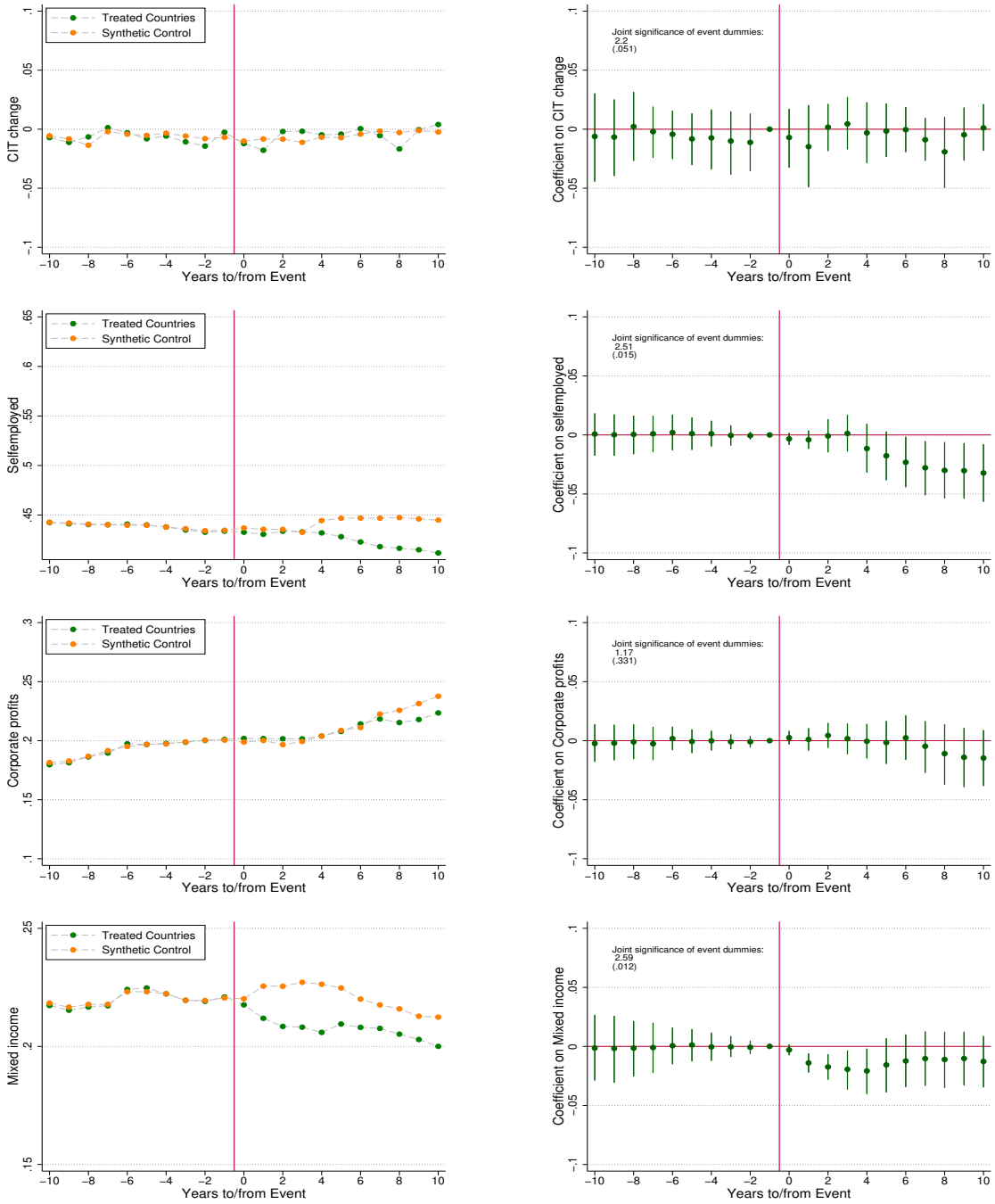
Notes: These figures show event-studies for trade liberalization in seven large developing countries: Argentina, Brazil, China, Colombia, India, Mexico and Vietnam. The panels correspond to different outcomes: corporate income tax rate; self-employed share of workforce; corporate profits share of national income; household mixed income share of national income. 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 control countries. The right-hand graphs show the coefficients on the 'to' ('since') dummies, in a regression with country fixed effects, year 'to' ('since') fixed effects, and calendar year fixed effects. The bars represent the 95% confidence intervals. Standard errors are clustered at the country-reform level and estimated with the wild bootstrap method. The top-left corners report the F-statistic on joint significance of the post-reform dummies, with the p-value in parentheses below. Details on methodology in Section 6.1 and Appendix C.2.

Figure A14: Event Study of Capital Liberalization – Simultaneous Matching on Outcomes



Notes: These figures show event-studies for capital liberalization reforms over four outcomes: capital openness; capital shares; 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. These graphs are constructed similar to Figure 8, with the exception that the synthetic control for each event-country is based on matching simultaneously on the four outcomes. 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.

Figure A15: Event Study of Capital Liberalization – Mechanism Outcomes



Notes: These figures show event-studies for trade capital reforms in the 25 developing countries of Chari, Henry, and Sasson (2012). The panels correspond to different outcomes: corporate income tax rate; self-employed share of workforce; corporate profits share of national income; household mixed income share of national income. 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 control countries. The right-hand graphs show the coefficients on the 'to' ('since') dummies, in a regression with country fixed effects, year 'to' ('since') fixed effects, and calendar year fixed effects. The bars represent the 95% confidence intervals. Standard errors are clustered at the country-reform level and estimated with the wild bootstrap method. The top-left corners report the F-statistic on joint significance of the post-reform dummies, with the p-value in parentheses below. More details are in Section 8 and Appendix C.3.

Table A1: Allocation of Taxes to Factor Incomes, by Type of Tax (Benchmark Estimates)

Type of tax	Series	Allocation to labor (λ_τ)
Panel A: Direct Taxes		
Personal income tax (PIT)	1100	Share λ [70,100%] to L (See 3.1.2)
Corporate income tax (CIT)	1200	0% to L $\Rightarrow$ all CIT to K
Other (unallocable) income tax	1300	50% to L, 50% to K; rare
Social security & payroll taxes	2000	100% to L
Property & wealth taxes	4000	0% to L $\Rightarrow$ all asset taxes to K
Panel B: Indirect Taxes & Other Revenue		
Indirect taxes	5000	excluded, assumed proportional
Other taxes	6000	excluded ; minor
Non-tax revenue	7000	excluded; non-tax

Notes: This table shows our benchmark assignment of statutory tax incidence λ_τ on labor (where the assignment to 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, 2019b). 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 A2: Synthetic Difference-in-Difference of Trade Liberalization

	Trade	Capital Share of Income	Capital Share Corp. Income	ETR Capital	ETR Labor
Panel A Synthetic control for each outcome separately					
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)
Panel B Synthetic control for all outcomes jointly					
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	292	282	290

Notes: This table shows the result from the difference in difference regression of our outcomes on interest in event countries (treated), compared to synthetic control countries. Panel A shows the results when the synthetic control matching is done for each event-country and outcome separately. Panel B shows the results when the synthetic control matching is done jointly on all outcomes (but still separately for each event-country). In practice we run the following regression: $Y_{it} = \beta^{DiD} * \mathbb{1}(j \geq 0)_t * D_i + \theta_t + \kappa_i + \pi_{Year(it)} + \epsilon_{it}$ Where, the β^{DiD} coefficient is the difference-in-difference estimate, representing the average treatment effect from period 0 through 10 post the trade liberalization event. We also 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*. * p<0.10 ** p<0.05 *** p<0.01.

Table A3: Reduced Form Regressions

	Trade (1)	Imports (2)	Exports (3)	ETR_K (4)	ETR_L (5)
$Z^{Gravity}$	0.040**	0.034***	0.023***	0.012**	0.005**
	(.017)	(.009)	(0.007)	(0.004)	(0.002)
$Z^{Oil-Distance}$	-0.012***	-0.005***	-.004***	-0.002***	-.001***
	(0.003)	(0.001)	(0.001)	(0.000)	(0.000)
N	*4518*	*4518*	*4518*	*4518*	*4518*

Notes: This table shows additional impacts of the two instruments $Z^{Gravity}$ and $Z^{Oil-Distance}$. Both the gravity and the oil-distance instrument is from Egger, Nigai, and Strecker (2019). For ease of comparison, in this table we use standardized transformations of each instrument. Across columns, we regress both instruments simultaneously on: trade, measured as the sum of imports and exports (% of GDP); imports (% of GDP); exports (% of GDP); the effective tax rate on capital, ETR_K; and, the effective tax rate on labor, ETR_L. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses are clustered at the country level.

Table A4: Robustness of IV 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 A5: 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 1, 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 A6: 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
Mendoza et al. (1994)	-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 = 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)
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)
ILO factor shares	0.00126	0.0214**	0.231*	0.194***	0.237**	0.264***	0.234*	0.0865
	(0.0332)	(0.0103)	(0.137)	(0.0652)	(0.0990)	(0.0622)	(0.126)	(0.0591)
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 to three alternative measures of effective tax rates: (1) with the Mendoza, Razin, and Tesar (1994) ETR definition. (2) with the capital share of PIT revenues bounded at levels that are either unrealistically low (0%), at an intermediate fixed parameter of (15%), or unrealistically high (30%). Finally, with the ILO (2019) factor share method, which adjusts the labor share of mixed-income a based on observable characteristics of the self-employed. 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 % of NDP); (log) population; (log) GDP 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. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Table A7: 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 D for further discussion of these instruments. For ease of interpretation, these estimates are unweighted (therefore, these estimates match Panel B of Table 1 as well as Table **??**). 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 A8: Association between statutory CIT rate and capital tax collection

	ETR_K			CIT revenue (% GDP)		
	All	High inc	Mid-low inc	All	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 A9: Impacts of trade on sectoral composition

	Industry (1)	Services (2)	Agriculture (3)
Panel A: OLS	0.041**	-0.018	-0.007
	(0.016)	(0.014)	(0.012)
Panel B: IV weighted	0.152**	0.089	-0.189
	(0.064)	(0.102)	(0.181)
First-stage F-statistic	33.40	33.40	33.40
Panel C: IV with controls	0.094***	0.069	-0.104
	(0.032)	(0.063)	(0.093)
First-stage F-statistic	22.86	22.86	22.86
Panel D: IV unweighted	0.012	0.088	-0.074
	(0.095)	(0.094)	(0.075)
First-stage F-statistic	7.75	7.75	7.75
N	*3849*	*3849*	*3849*

Notes: This table shows the impacts of trade on sectoral composition. Across columns, the outcome variable is (expressed as a % of national value-added): industry (manufacturing and non-manufacturing); services; agriculture. Across panels, we present: OLS estimates (Panel A); benchmark IV (Panel B); IV with controls (Panel C); IV without weights (Panel D). The weights used in all panels except D are net domestic product in constant 2019 USD at PPP. In Panel C, 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, 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). At the bottom of each panel, we report the 1st stage F-statistic on the joint significance of the two instruments. All estimates include country and year fixed effects, with errors clustered at country level. * p<0.10 ** p<0.05 *** p<0.01. Standard errors in parentheses.

Appendix B Data construction

B.1 National income data

Data sources We used two datasets to compute factor shares of net domestic product. Both come from the United Nations Statistics Division. The first is the SNA2008 (UN SNA, 2008) online data repository.[44] The second is SNA1968 archival material (UN SNA, 1968). While the SNA2008 data has extensive coverage, we were able to extend further its coverage using SNA1968 observations. As shown in Table B1, we nearly double our 'complete' coverage observations by adding the historical data. What we 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. We briefly describe these national income accounting concepts here, and how we attribute them to labor or capital factors of production, before returning to discuss missing data and coverage.

Income Concepts Net domestic product at factor prices is (equation (1) in text):

$$Y = CE + OS_{CORP} + OS_{HH} + OS_{PUE}$$

We detail below the national accounting concepts

- '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, house-

[44] We accessed this data via the World Inequality Database (WID, 2020).

hold 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).

All operating surpluses—of corporations, of unincorporated enterprises, and of the household sector—are estimated net of depreciation (consumption of fixed capital, CFC). This explains why we obtain factor shares of Net Domestic Product instead of Gross Domestic Product (below we provide further detail on the data on depreciation).

These concepts are sufficient for the decomposition of factor income into labor and capital income flows, and our definition of factor shares follows that in most of the literature. For completeness, we explain some of the other concepts that do not figure in our calculations. We explain here why we decided not to include them and the role they would play if included.

(1) *net indirect taxes* (indirect taxes net of subsidies, NIT) are excluded from consideration. In our data these comprise 8-15% of national income. These indirect taxes are assumed to be factor-neutral, i.e., levied on the returns to capital and labor proportional to each factor's share in the affected production process. (2) net foreign income—the net flow of labor and capital incomes (usually compensation of employees or corporate profits) from abroad— is ignored. One could move from 'net domestic product' to 'net national income' as our total income concept. In practice, this seems to be a less relevant concept for our purposes (most countries tax income earned domestically regardless of citizenship, whereas net foreign income is taxed only with difficulty), but, this adjustment is minor for most countries. In national accounting concepts following the 'income approach' to domestic product,

government surplus is zero by construction and thus does need to be distributed to labor or capital.[45] Nonprofit institutions are included in the household sector.[46] (3

Missing data and imputations When values are still missing after we have appended the two UN data sources, we follow a simple a transparent imputation procedure. In the description here, we focus on mixed income, as this is the concept whose values are most frequently missing in UN SNA data—in both UN SNA (2008) and especially UN SNA (1968)—but the same procedure applies to other income concepts whose values are occasionally missing in UN SNA data, as well.

We follow the imputation procedure from Blanchet and Chancel (2016), which presents a simple and transparent method to impute consumption of fixed capital (depreciation), used in the World Inequality Database (WID) series on national income (contrarily to us the WID does not systematically decompose income into factor shares). To explain our imputation of mixed income OS_{PUE}, we present the following stylized facts:

1. Mixed income (OS_{PUE}) represents a lower share of national income in rich countries. This is due to the outsized role of the informal economy and household enterprises in developing countries.

2. Some countries have structurally high (or low) levels of mixed income as a share of national income. These specific long-term country effects represent either (i) a 'real' economic trend (e.g., the informal economy, as above, or the role of certain industries that that are characterized by self-employment); or (ii) a specific 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).

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

[46] 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).

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 B1.

Mixed Income Adjustment As a robustness check (discussed above in Section 3.1, with results in Appendix Table A6), we implement the method developed in Guerriero (2019) and ILO (2019), to adjust for the share of labor income in mixed income. For each of the three classes of self-employed workers for which data is available in ILO labor force survey data (*viz.* self-employed employers, 'own-account' workers, and 'contributing

family members'),[47] we estimate a country-year coefficient, computed as the ratio of the average employee's wage to that of the self-employed worker.

Total labor income for a given country-year is 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 national accounts; s_{emp} is 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 worker. Each self-employment category i corresponds to a specific earnings coefficient γ_i relative to the average employee compensation.[48] Capital income Y_K' is $Y - Y_L'$.

Note, in Appendix Table A6, that the results from this robustness check do not affect our estimates, as the factor shares estimated from this method do not systematically differ from the factor shares calculated with a benchmark $\phi = 70\%$ labor share of mixed income.

For the United States, we use estimates of factor shares in net domestic product from Piketty, Saez, and Zucman (2018) that incorporate a granular treatment of mixed income reflecting specificities of the US non-corporate business sector. Specifically in the United States, a number of large businesses (including listed firms) are organized as partnerships (as opposed to corporations) and as such are classified as non-corporate businesses, while they would be treated as corporations in other countries. Their income is counted as mixed income in the US National Income and Product Accounts, while it would typically be recorded as corporate profits in other countries. Since 1986, a growing number of businesses have been organized as partnerships for tax reasons. The series we use take these facts into account by: (1) assuming a higher capital share of income for partnerships vs. other non-corporate businesses (sole proprietorships); and (2) factoring in the rising capital intensity of partnerships since the 1980s.

[47] This is according to the ILO's International Classification of Status in Employment, ICSE-93, with data from 1991 to present.

[48] For country-years missing these concepts s_i and γ_i in raw ILOSTAT (2021) data—particularly for the era prior to 1991—we impute observations using the same procedure as above.

B.2 Tax revenue data

Data As described in Table B1, our tax revenue data draws from four types of sources:

- OECD online Government Revenue Statistics (OECD, 2020)
- ICTD/UNU-WIDER Government Revenue Dataset (ICTD/UNU-WIDER, 2020)
- Historical archive work, principally from the Harvard Lamont Library.
- IMF Government Finance Statistics (IMF GFS, 2005)

While our use of the first three is more or less self-explanatory, and we retrieve off-the-shelf data from these international organizations,[49] 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.

[49] 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.

Revenue Concepts We classify each type of tax revenue following the OECD Government Revenue Statistics (OECD, 2020). We consider the following simplified classification (see also Table A1).

- The '1000' series of taxes, covers taxes on income and profits:

  > 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).

  We always separate the 100 series between personal income taxes (PIT, 1100) and corporate income taxes (CIT, 1200).

  While all corporate income taxes is allocated to capital taxation, PIT revenue is a mix of labor and capital income. Thus, within PIT, an important parameter is the share of revenue assigned to capital, denoted λ_{PIT} (as in Table A1). As discussed in Section 3.1.2, we rarely observe directly the level of PIT revenue that derives from capital income and gains versus labor income.[50] In the United States, Piketty, Saez, and Zucman (2018) find that about 15% of the PIT revenue is from capital and about 85% from labor. To construct country-year specific λ_{PIT}, we start from the US benchmark ($\lambda_{PIT} = 15\%$) to which we make two adjustments:

  1. First, how high the PIT exemption threshold is located in the income distribution, will impact λ_{PIT} since capital income is concentrated at the top of the income distribution. we retrieve PIT exemption thresholds from Jensen (2022) and extend these estimates to all country-years in our data. Countries with a higher PIT exemption threshold are assumed to have a higher λ_{PIT} than countries with a lower PIT exemption threshold. Since the US has a low exemption threshold

[50] The subtotal of PIT revenue from capital income includes: taxes on dividends; on the capital share of self-employment income; and on capital gains. OECD data on statutory income tax rates tell us about dual-income tax systems where the rate on dividends differs from the rate on salaries (used in our adjustments below). OECD (2020) revenue data occasionally reports on PIT revenue from capital gains. From 2010-18, among 27 countries that report some capital gains tax revenue within PIT revenue, this revenue averages just under 4% of PIT revenue. In the US, that number is 7.5%. Of course, many countries do not tax capital gains as such, and most countries tax other forms of capital income within the PIT system.

and has $\lambda_{PIT} = 15\%$, we similarly assign a 15% value to countries for which the PIT covers half or more of the population (usually high-income countries). For countries where the PIT covers 1% or less of the population, we assign a maximum capital share of 30%. For PIT thresholds with a coverage between 1 to 50% of the population we do a linear extrapolation assigning a λ parameter between 15 and 30%.

2. Second, we reduce λ_{PIT} in countries where a dual-PIT system sets capital income taxation to a lower— often flat— rate, while labor income continues to be taxed with (higher) progressive marginal tax rates. In these dual systems, we measure the percent difference between the capital tax rate (dividends) and the labor tax rate (this percent difference will be negative). Data on capital and labor rates are taken from OECD Revenue Statistics and country-specific tax code documents. We assume that 50% of capital in PIT comes from these sources that are taxed preferentially, reflecting the fact that self-employment income, and occasionally capital gains, are likely not to be taxed differently from labor. For this 50%, we multiply λ by the percent difference in capital versus labor tax rates.

- The '2000' and '3000' series of tax revenues, include respectively 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 A1.

- The '4000' series 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, 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 non-tax revenues and tax expenditures from consideration. Once we classified each type of tax according to the simplified OECD (2020) framework, we proceeded to stitch together long-run tax revenue series for all countries.

Appending and Stitching Data Sources 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:

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

1. OECD revenue statistics take precedence.

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.

B.3 Case study: China's establishment of a modern tax system in 1994

Based on the rules we established, we only include ex-command economies into our data starting in 1994. Given China's importance for the global economy, we discuss here the reason for this choice.

The tax revenue data for China covers most of our sample period, although it improves a in quality in the 1980s.[53] Prior to the 1980s, China had a command economy's model

[53]Official public finance statistics are available online at https://data.stats.gov.cn/english/index.htm.

of 'profit delivery,' in which the state spent directly the revenues of profitable SOEs and subsidized unprofitable ones, an example of soft budget constraints (Kornai, Maskin, and Roland, 2003). As part of the decade of economic reforms beginning in 1978, a corporate income tax appears in China around 1983-84 (Wong and Bird, 2008). A decree from the State Council in 1983 put a new 55% tax on the profits of enterprises, which were still almost entirely state-owned. Starting in 1985, we observe CIT revenue in the data, although it appears implausibly high (as a percentage of NDP, or of capital income). In addition, not long after, the tax system was further reformed into a 'fiscal contracting' system whereby firms negotiated a fixed tax revenue level (regardless of economic outcomes) to local governments, who in turn delivered a share to the central government. This system resembles a firm "poll" tax, and subsumed all types of tax revenues, including CIT, VAT and payroll, making it impossible to assign taxes to capital or labor. For these reasons, we exclude this sort of 'pseudo'-CIT revenues dating from 1985 through 1993.

China's modern tax system arguably begins in 1994. A valuable resource to study the introduction of a modern tax system is the book *Public Finance in China* (World Bank, 2008). The book argues that prompted by low central government revenues, China established in 1994 a central tax administration; reformed the broken 'fiscal contracting' system; unified the PIT; created a VAT; and reduced 'extra budgetary' (non-tax) revenues.[54] Thus from 1994 onward we can categorize tax revenue by sources, assign them to capital or labor, and thus construct ETRs. It should be noted, however, that to the extent that they were measurable, tax revenue appeared higher in the 1980s than in the 1990s in China.

[54]The state-run media summarizes these policy reforms succinctly (China Daily, 2018), and agrees with scholarly work (e.g., World Bank (2008) see esp. pp.13-22; or Wong and Bird (2008) pp.430-440)

Table B1: Main 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	40.2%
ICTD	1276	18.5%
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, online sources, and IMF historical data. (See Appendix B.)

Appendix C Event studies

C.1 Description of trade events

Our selection of trade events is determined by three criteria: (i) the event can unambiguously be related to measurable policy reforms; (ii) the policy reforms induced large changes in trade barriers; and (iii) 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 policy event. The second criterion increases the likelihood that we are able to observe sharp breaks in trend in our macroeconomic outcomes which coincide with the timing of the policy event. The third criterion allows for our results to be compared to previous work, and for anyone to study the narrative surrounding these reforms. 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 it is easier to measure on a consistent basis across space and time, as compared to non-tariff trade barriers and other forms of globalization. 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 large reductions 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 prior to the events, such that reductions in tariff rates remained an available policy lever.[55] These trade liberalization events were drastic: Brazil

[55]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).

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 and references to additional papers which study the events in detail are provided in Table C1.

Timing of events Most studies provide detailed discussions of the context surrounding the events. We repeat the rationale cited for the liberalization events and discuss why these events are plausibly exogenous to the country's economic circumstances at the time.

- **Brazil** The 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 and that the reform was unexpected:

  > "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)

- **Columbia** Similarly to Brazil, tariff reductions in Colombia in 1985 were driven by the country's commitment to impose uniform rates across products and industries under the negotiation commitments to the WTO. In Colombia's case, Goldberg and Pavcnik (2007) note that the reform objective was to reduce cross-industry dispersion under WTO negotiations, thereby making "the endogeneity of trade policy changes less pronounced here [in Colombia] than in other studies."

- **China** Brandt et al. (2017) note that trade openness reforms had gradually been implemented in China prior to the country's WTO accession event in 2001, 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 outcomes.

- **India** The 1991 event in India occurred as a result of an IMF intervention that dictated the pace and scope of the liberalization reforms. Under the IMF program, 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 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).

- **Vietnam** 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. Vietnam's case is compelling since the liberalization even was triggered by a foreign party, rather than by its own government.

These descriptions of reform context do not argue that liberalization events were triggered by trends in taxation needs. Thus, this narrative analysis complements the absence of a pre-trends result (Figure 7), to help alleviate endogeity concerns in the timing of the events.

Post-event reforms Yet, even if the reform timing is uncorrelated with confounding trends, the interpretation of the event studies 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 all major further events which followed the initial liberalization events. 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; India's event in 1991 is followed by foreign direct investment liberalization in 1993; and Mexico's 1985 WTO accession was followed by a removal of capital inflow restrictions in 1989 and accession to NAFTA in 1994.

Their discussion suggests that other reforms did occur, often a few years after the trade liberalization event; and, these 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-run impacts might be better interpreted as the reduced-form effects of globalization more generally, which includes an increase in the flow of goods, services, and capital, as well as further policy responses.

C.2 Event study methodology

Sample Construction Our sample is constructed by applying a synthetic matching procedure to every treated country, for each outcome of interest. The donor pool (the set of all control countries from which to chose the synthetic control group) has to be fully balanced in all pre-event periods. Thus, we discard all countries with data gaps before 1976. This gives us a sample of 103 countries for each outcome.[56] We have ten 'pre' event

[56]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.

years for every country, except for Mexico and Colombia, where we have nine pre-period years.[57] We then pool together all seven treated countries and their synthetic control units.

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

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

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, β_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:

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

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 β^{DiD}, in Table A2.

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

[57] Moreover, China and Vietnam only have tax revenue data from 1994 onward, but its event happens in 2001. We therefore only match on seven pre-period years.

1. 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.

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

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

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

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 group per treated country. The resulting weights for the three most prominent countries in each control group are reported in Table **??**. 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 A10. 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 **??**, 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.

C.3 Capital liberalization: Events and results

In this subsection, we provide details on the event-study analysis of capital liberalization in Section 8. We use the 25 liberalization events in developing countries measured by Chari, Henry, and Sasson (2012). The authors identify liberalization dates using the point in time when countries first permitted foreigners to purchase shares of companies listed on the domestic stock market. At first glance, this event may appear to be a narrow way to define capital liberalization, but the authors provide several arguments that the dates in fact serve as indicators for larger move towards open capital markets.

First, net equity inflows soared in the periods where developing countries liberalized their stock markets. While such flows were practically non-existent in the 1970s and early 1980s, they accelerated sharply around the end of the 1980s when most countries implemented their liberalization reforms. It is important to note that stock market liberalizations account for a significant fraction of foreign direct investment (FDI). Indeed, for purposes of official statistics, FDI includes any stock transaction (that is, a cross-border merger or acquisition) that results in the purchaser owning 10 percent or more of the voting shares. Chari, Henry, and Sasson (2012) quote data from Latin American and East Asia which show that cross-border mergers and acquisitions account for 40% to 60% of FDI. Second, by facilitating cross-border financial flows and ownership, stock market liberalization induces large inflows of capital goods. The authors quote other studies

which show that, following capital liberalization, the share of capital goods in total imports rises by 9%. Through this channel, the liberalization events contribute to raising aggregate investment. Third, policymakers often worry that increased net equity inflow can easily convert into net outflow at the early signs of macro-economic instability. In turn, the increased outflow can compound the instability and introduce volatility. But the authors note that the liberalization events considered here were not marked by any reversal of the liberalization policies or the equity inflows in the post-reform years. This allows us to study the effects of liberalization reforms more cleanly and over longer time-horizons. Fourth, there are difficulties in determining precise liberalization dates and the literature has mainly chosen to ignore this measurement problem (Eichengreen, 2001). By focusing on the initial date where stock markets were first opened to foreign capital, the events are more cleanly measured than other candidate de-jure policy changes which may be harder to pinpoint. For further details about the complexities of determining liberalization dates, see Henry (2007).

For these reasons, the authors consider the stock market liberalization events to be meaningful indicators of countries' progression towards freer capital inflows. The authors select 25 liberalization events which: are recorded in official sources (Standard and Poor's Emerging Markets Database); are verified from primary sources; and, have been used elsewhere in the literature. Table C2 lists the 25 countries with their liberalization dates.

Our sample is constructed by applying a synthetic matching procedure to every treated country for each outcome of interest. We follow the methodology and estimation techniques described in detail in Appendix Section C2. Table C2 provides details on the main countries that form the synthetic control group for each treated country and each outcome. Our main measure of capital openness is the sum of foreign assets and liabilities. This stock is expressed as a percent of GDP; given the strong right-skew, we use the log of this ratio in the analysis. We find similar results when using alternative measures of capital openness, including portfolio equity assets and liabilities (% of GDP) and the KOF financial globalization index (Gygli et al., 2019). Moreover, we also find that the liberalization events led to positive changes in the de jure index of capital openness produced by Chinn and

Ito (2006). These results are available upon request. Finally, we note that due to limited data availability, we were not able to study the impact on foreign direct investment.

The level and event-study graphs for each outcome (capital openness; capital share of national income; capital share of corporate income; ETR_K; ETR_L) are reported in Figure 8. The statistic in the top-left corner reports the F-statistic and p-value for the joint significance of all post-liberalization dummies. It is important to note that, when jointly considered, the post-liberalization coefficients for ETR_K are not statistically significant. The panels make clear this is because the statistically significant effect on capital taxation materializes with a few years' lag to the liberalization event. Appendix Figure A14 shows that the results are robust to creating synthetic controls for each event-country based on matching the four outcomes simultaneously (capital openness; capital share; ETR_K; ETR_L). In particular, the events continue to be associated with a large increase in capital openness, a positive impact on capital taxation and a muted (null) impact on labor taxation.

Table C1: Weights in Synthetic Control for Trade Liberalization Events

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

Notes: This table shows the seven treated countries and the three countries with the largest weight in the synthetic control group for each outcome and each treated country. 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 C2: Weights in Synthetic Control for Capital Liberalization Events

Country	Event Year	Capital share (NDP)	Weight	ETR on capital	Weight	ETR on labor	Weight	Capital Openness	Weight
Argentina	1989	Uruguay	19.7 %	Bangladesh	39.7 %	Chad	35.8 %	United States	42.3 %
		Qatar	16.0 %	Paraguay	29.8 %	Togo	35.8 %	Papua New Guinea	33.3 %
		Bolivia	11.8 %	Nepal	7.5 %	Uruguay	28.2 %	Tanzania	11.8 %
		...	...	...	...	...	...	...	...
Brazil	1988	Malawi	9.6 %	Sudan	34.0 %	Guyana	30.3 %	United States	54.1 %
		Swaziland	9.6 %	Mali	10.9 %	Congo	18.8 %	El Salvador	15.5 %
		Mauritius	7.8 %	Zambia	10.8 %	Togo	17.2 %	Iran	13.1 %
		...	...	...	...	...	...	...	...
Chile	1987	Zambia	46.0 %	Bolivia	45.8 %	Madagascar	49.9 %	Kuwait	35.8 %
		Botswana	15.8 %	Uganda	25.2 %	Uruguay	42.3 %	Lebanon	16.7 %
		Kuwait	13.0 %	Canada	19.0 %	Norway	7.6 %	Ireland	12.3 %
		...	...	...	...	...	...	...	...
Colombia	1991	Fiji	29.6 %	Bangladesh	59.2 %	Sudan	60.1 %	Kuwait	31.7 %
		Bolivia	17.1 %	Iceland	20.9 %	Trinidad and T.	24.5 %	Gabon	19.7 %
		Ireland	8.6 %	Papua New Guinea	11.2 %	Paraguay	14.9 %	Austria	13.0 %
		...	...	...	...	...	...	...	...
Egypt	1991	Qatar	33.1 %	Sierra Leone	43.3 %	Nicaragua	50.7 %	Bolivia	45.6 %
		Paraguay	18.2 %	Uruguay	24.3 %	Bahrain	23.1 %	United States	31.9 %
		Mauritius	10.9 %	Zambia	18.5 %	Luxembourg	13.0 %	Syria	18.0 %
		...	...	...	...	...	...	...	...
Spain	1993	Canada	22.6 %	Swaziland	22.1 %	Lesotho	30.3 %	Italy	44.6 %
		Swaziland	18.0 %	Jamaica	16.5 %	Norway	29.8 %	Singapore	30.5 %
		Albania	8.3 %	New Zealand	8.9 %	Italy	18.7 %	Sweden	12.9 %
		...	...	...	...	...	...	...	...
Greece	1994	Costa Rica	50.5 %	Myanmar	41.8 %	Panama	30.3 %	Botswana	13.5 %
		Mozambique	11.4 %	Bahamas	21.4 %	Austria	25.9 %	United States	9.4 %
		Albania	10.4 %	Malawi	19.2 %	Nicaragua	20.6 %	Saudi Arabia	8.6 %

		...	...	...	...	...	...	...	...
Indonesia	1989	Mozambique	20.4 %	Benin	17.1 %	Lebanon	36.0 %	Ireland	30.3 %
		Sweden	5.7 %	Uganda	16.3 %	Bahamas	19.5 %	Oman	21.1 %
		Gabon	4.8 %	Bangladesh	14.3 %	Myanmar	11.5 %	United States	19.3 %
		...	...	...	...	...	...	...	...
India	1986	Mauritius	7.4 %	Uganda	29.3 %	Lebanon	64.2 %	Peru	49.7 %
		Costa Rica	7.0 %	Bahrain	14.2 %	Chad	15.7 %	Denmark	19.9 %
		Finland	6.1 %	Peru	10.8 %	Peru	7.1 %	United States	14.6 %
		...	...	...	...	...	...	...	...
Israel	1989	Denmark	34.9 %	Sweden	39.1 %	Sweden	78.8 %	United States	20.9 %
		Algeria	27.5 %	Ecuador	17.8 %	Guyana	21.0 %	Mauritania	20.2 %
		Bahrain	17.5 %	Barbados	16.3 %	...	...	Saudi Arabia	19.4 %
		...	...	...	...	...	...	...	...
Jordan	1995	Haiti	23.8 %	Lebanon	29.4 %	Nepal	70.2 %	Gabon	21.9 %
		Afghanistan	19.7 %	Ghana	25.4 %	Ghana	16.4 %	Cameroon	19.7 %
		Saudi Arabia	19.2 %	Honduras	22.3 %	Honduras	13.4 %	Botswana	19.4 %
		...	...	...	...	...	...	...	...
Korea	1987	Lebanon	31.6 %	Uganda	38.2 %	Kuwait	14.3 %	United States	21.0 %
		Cote d'Ivoire	28.4 %	Iceland	17.1 %	Madagascar	11.6 %	Qatar	19.5 %
		Panama	11.9 %	Ethiopia	10.3 %	Bangladesh	10.0 %	Tanzania	15.2 %
		...	...	...	...	...	...	...	...
Mexico	1989	New Zealand	35.1 %	Niger	41.8 %	Chad	68.0 %	Saudi Arabia	43.5 %
		Oman	22.3 %	Sierra Leone	31.2 %	Uruguay	15.3 %	United States	20.7 %
		Panama	19.4 %	Central African Republic	10.4 %	Norway	8.9 %	Bolivia	20.2 %
		...	...	...	...	...	...	...	...
Malaysia	1987	Paraguay	14.8 %	Denmark	19.3 %	Sudan	25. %	Syria	31.4 %
		Kuwait	10.3 %	Ecuador	18.4 %	Mauritius	15.8 %	Guatemala	19.4 %
		Luxembourg	5.9 %	Switzerland	15.8 %	Togo	11.2 %	United States	18.0 %
		...	...	...	...	...	...	...	...
Nigeria	1995	Mozambique	40.0 %	Rwanda	60.3 %	Chad	26.4 %	Paraguay	36.1 %
		Zambia	20.1 %	Sweden	17.2 %	Myanmar	22.3 %	Denmark	27.0 %

		Saudi Arabia	17.9 %	Papua New Guinea	13.8 %	Nepal	21.2 %	Oman	13.4 %
		...	...	...	...	...	...	...	...
Pakistan	1991	Nepal	7.5 %	Lebanon	39.3 %	Sierra Leone	40.9 %	Australia	18.2 %
		Mozambique	4.8 %	Bahamas	9.7 %	Bangladesh	14.8 %	Iran	14.2 %
		Tunisia	4.6 %	Malawi	8.3 %	Sri Lanka	14.0 %	Bolivia	8.7 %
		...	...	...	...	...	...	...	...
Philippines	1986	Congo	67.1 %	Mauritius	52.4 %	Peru	29.1 %	Cote d'Ivoire	21.7 %
		Iceland	15.5 %	Central African Republic	47.4 %	Chad	22.8 %	Kuwait	16.6 %
		Mozambique	11.1 %	...	...	Bahrain	21.6 %	Saudi Arabia	16.2 %
		...	...	...	...	...	...	...	...
Portugal	1993	France	94.3 %	Mozambique	40.0 %	Ecuador	32.6 %	Singapore	27.7 %
		Zambia	4.8 %	Guyana	26.8 %	Iceland	23.1 %	Iran	22.2 %
		Mozambique	0.8 %	Singapore	14.3 %	Finland	22.6 %	Canada	21.6 %
		...	...	...	...	...	...	...	...
Thailand	1987	Cote d'Ivoire	38.4 %	Bahamas	34.0 %	Lebanon	34.7 %	El Salvador	17.3 %
		Qatar	22.4 %	Gambia	20.2 %	Oman	24.3 %	Qatar	16.8 %
		Ireland	8.5 %	Haiti	11.8 %	Jamaica	9.7 %	Japan	12.7 %
		...	...	...	...	...	...	...	...
Turkey	1989	Oman	54.1 %	Honduras	65.4 %	Peru	65.6 %	Niger	23.9 %
		Sierra Leone	12.3 %	Benin	19.4 %	Norway	33.3 %	Germany	22.3 %
		United Kingdom	9.8 %	Canada	9.8 %	Iceland	0.9 %	France	15.8 %
		...	...	...	...	...	...	...	...
Taiwan	1986	Paraguay	24.6 %	New Zealand	26.8 %	Bangladesh	35.1 %	Saudi Arabia	52.4 %
		Congo	9.0 %	Bahrain	22.6 %	Haiti	23.6 %	Qatar	16.6 %
		Niger	7.1 %	Ethiopia	18.8 %	Central African Republic	11.2 %	Nicaragua	13.6 %
		...	...	...	...	...	...	...	...
Venezuela	1990	Sierra Leone	56.1 %	Gabon	48.6 %	Jamaica	51.2 %	Tanzania	31.4 %
		Kuwait	16.4 %	Oman	39.1 %	Lesotho	34.2 %	Panama	19.9 %
		Botswana	14.5 %	Mozambique	12.0 %	Barbados	14.5 %	United States	17.3 %

		...	...	...	...	...	...	...	...
South Africa	1995	Nepal	16.0 %	Sweden	57.3 %	Honduras	52.2 %	United States	24.0 %
		Honduras	15.1 %	Congo	16.5 %	Iceland	39.8 %	Panama	21.9 %
		Jamaica	14.5 %	Zambia	9.3 %	Barbados	5.4 %	Senegal	16.4 %
		...	...	...	...	...	...	...	...
Zimbabwe	1993	Belgium	35.6 %	Togo	52.9 %	Dominican Republic	42.3 %	Honduras	31.0 %
		Austria	31.9 %	Barbados	20.2 %	Iceland	35.6 %	Benin	30.5 %
		Mozambique	17.2 %	Trinidad and T.	12.3 %	Lesotho	13.4 %	Iran	10.5 %
		...	...	...	...	...	...	...	...

Appendix D Instrumental variables for trade

In this section, we outline the construction of the two instrumental variables. Both instruments are from Egger, Nigai, and Strecker (2019), which provide further details.

D.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 A7. The instrument is slightly stronger when using the log-level of trade or the share of goods-trade as the endogenous measure of openness.

D.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,[58] using international shipping logistics calculators at SeaRates (2021). To calculate the distance of each city to each port, we look at the map of the city and its distance to port, through the lens of these

[58]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.

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).[59] 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 not 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.

[59] 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 A7. The instrument is strongly associated with changes in trade within-country over time and is robust to using various measures of trade openness.