The Taxing Deed of Globalization[†]

By PETER H. EGGER, SERGEY NIGAI, AND NORA M. STRECKER*

This paper examines the effects of globalization on the distribution of worker-specific labor taxes using a unique set of tax calculators. We find a differential effect of higher trade and factor mobility on relative tax burdens in 1980–1993 versus 1994–2007 in the OECD. Prior to 1994, greater openness meant that higher income earners were taxed progressively more. However, after 1994, we document a globalization-induced rise in the labor income tax burden of the middle class, while the top 1 percent of workers and employees faced a reduction in their tax burden of 0.59–1.45 percentage points. (*JEL* D31, F16, F61, H22, H24)

Over the past decades, global integration of the world economy has risen dramatically—as has inequality across and within countries. That globalization may be a source of inequality is widely accepted by economists; however, the exact channels of causation are still debated. We explore one of them: the influence of globalization on inequality through changes in labor income taxation systems. While progressive income taxation may serve as a means to mitigate adverse effects of globalization, the present paper finds that the cross-border integration of markets since the mid-1990s induced changes in the taxation of labor incomes that exacerbate labor income inequality.

Economies became increasingly integrated in the post-World-War-II era, which was characterized by falling barriers to cross-border flows of goods and production factors. However, the consequences of vanishing barriers to trade and factor mobility are admittedly complex. From a tax perspective, there are two main

channels through which globalization affects workers. On the one hand, globalization has adverse effects on some workers (Autor, Dorn, and Hanson 2013) so that workers demand greater insurance and an expansion of public goods provision (see Rodrik 1998).[1] On the other hand, the greater cross-border mobility of some factors, such as capital (see Persson and Tabellini 1994; Devereux et al. 2002; Devereux, Lockwood, and Redoano 2008) or skilled and high-income workers (see Kleven et al. 2014), limits the opportunities of countries to tax them. Specifically, countries have cut their tax rates on profits and capital over the past two decades, converging to lower rates (see Devereux et al. 2002). Such a convergence can also be documented for labor income tax rates on high-income earners but not on the median earner.

Figure 1 provides a portrait of this development in capital and labor income tax rates (for the top-1 percent and the median income earner) across the 65 biggest economies between 1980 and 2007.[2] The figure confirms that the tax rates levied on relatively mobile bases (corporate profits and top-1 percent incomes) have declined, while taxes on median labor incomes have increased in the recent past. This could be due to the higher sensitivity of capital and highly mobile workers to tax changes (Kleven et al. 2014), while low-mobility workers are less responsive. Hence, governments must, can, and do increasingly rely on fewer and relatively immobile tax bases, which are essentially three: (i) property and wealth, (ii) sales and consumption, and (iii) the labor income of relatively immobile workers.[3] This argument is also raised in the literature on optimal nonlinear taxation in open economies (Simula and Trannoy 2010; Piketty and Saez 2013; Lehmann, Simula, and Trannoy 2014).[4]

The present paper examines the effects of globalization on labor income tax burdens across the income distribution in a host of countries and years and addresses the effects of globalization on the relative size of the revenues from different tax bases.[5] To that end, we compiled the biggest existing dataset on annual labor income tax calculators, between the years 1980 and 2007 for the 65 biggest economies.[6] We combine these tax calculators with information on country-year-specific gross labor income distributions, tax revenues, and measures of globalization.

The key findings are the following. Since the mid-1990s, globalization has caused a decline in the relative tax burden of top income earners, particularly in OECD

---

[1] For the link between globalization and the size of government spending, see also Cameron (1978), Alesina and Wacziarg (1998), and Epifani and Gancia (2009).

[2] Labor income tax rates in this context include the worker-specific labor income tax rate and employee-borne social security contributions.

[3] For the link between globalization and tax design, see Ganghof and Eccleston (2004) and Hines and Summers (2009).

[4] Saez (2001) examined optimal taxation with immobile labor, whereas Simula and Trannoy (2010) and Lehmann, Simula, and Trannoy (2014) extended the analysis to consider migration in a canonical optimal taxation setting along the lines of Mirrlees (1971) and Diamond (1998). Earlier empirical work on the link between migration and labor income taxes is limited to individual countries and location decisions of workers within countries (see Kirchgässner and Pommerehne 1996, Wagner 2000, Schmidheiny 2006). This literature documents higher income tax sensitivity for more skilled (higher earning) and younger individuals than the average.

[5] While revenues from income taxation are the most important source for developed countries, they are less important than value-added taxes in less developed economies.

[6] See Egger and Strecker (2018) for more information on the dataset, which in total covers 252 economies and territories between 1980 and 2012 for 12 household archetypes (distinguished by the types of allowances and deductions granted across all the existing systems).
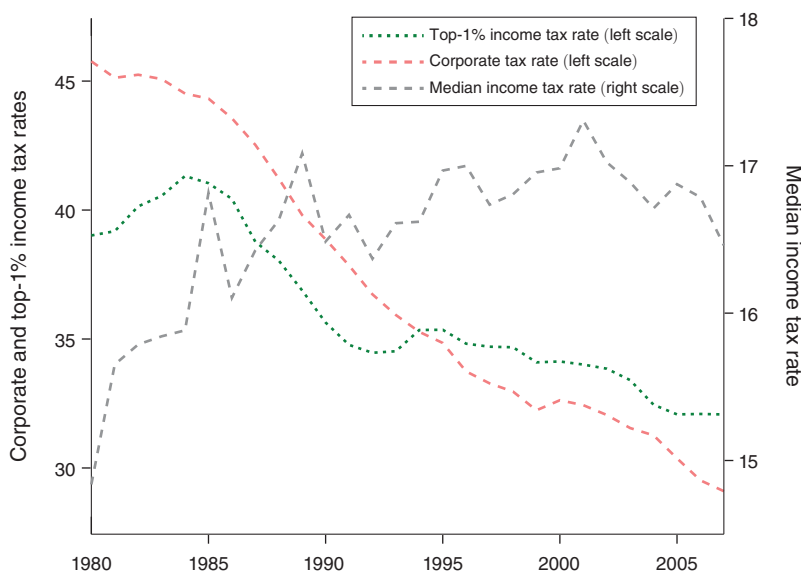
FIGURE 1. CORPORATE TAX RATES AND LABOR INCOME TAX RATES FOR TOP-1 PERCENT AND MEDIAN WORKERS
IN 65 ECONOMIES OVER 1980–2007

countries. The novelty here is that this change was induced inter alia through modifications of countries' income tax laws and codes, which provided for a relatively less aggressive taxation of high-income earners, relative to the center of the labor income distribution. We find that, due to the rise in globalization between 1994 and 2007 alone, the top 1 percent of income earners in the average OECD country faced a globalization-induced reduction in their relative labor income tax burden of 0.59–1.45 percentage points, whereas the tax burden increased by 0.03–0.05 percentage points for the median earner. This suggests that globalization has indeed increased net income inequality through lower tax burdens on top-income earners and higher tax burdens on the middle class. These results are confirmed in other settings: using microdata from the Luxembourg Income Study (LIS) and using microdata at the subnational level and interstate, rather than international, worker mobility in the United States.

In our empirical approach we acknowledge that the ordinary least squares (OLS) results may be biased if the distribution of labor income tax burdens is endogenous to globalization and/or if relevant and correlated variables are omitted. We guard against this potential problem in two ways. First, we define the labor-income-tax-based dependent variables in relative rather than absolute terms which alleviates potential concerns about the influence of worker-specific tax rates on international trade. Second, we apply an instrumental variables strategy by employing two distinct types of instruments for globalization. These two instruments are derived in very different ways and rely on different identifying assumptions. They are both correlated with the measures of globalization but are plausibly exogenous to country-year-worker-specific effective relative tax burdens or rates. More importantly, the instruments confirm the OLS results and lead to very similar qualitative and quantitative predictions regarding the effects of globalization on the

distribution of labor income tax burdens as well as of tax rates and on the relative size of labor income tax revenues in overall tax revenue.

The remainder of the paper is organized as follows. In the next section, we discuss our estimation strategy and examine the link between globalization and relative labor income tax burdens across the labor-income distribution. In Section II, we document the causal link between globalization and the size and composition of tax revenues. Section III is devoted to several robustness tests and extensions. The last section concludes.

I. Globalization and Relative Labor Income Tax Burdens across Earners

It is well documented in the literature that trade grew substantially in all major regions of the world over the last decades; furthermore, the liberalization of discriminatory and nondiscriminatory trade policy measures is partly responsible for this development (see Baier and Bergstrand 2001, Egger and Nigai 2015). Important policy changes in that vein were the conclusion of the GATT (General Agreement on Tariffs and Trade) and WTO (World Trade Organization) rounds, as well as the formation of what Baldwin (2006) called the "spaghetti bowl" of preferential trade agreements. At the same time, labor migration increased substantially and, in some OECD countries, became a major force in counteracting their plummeting fertility rates. Both trade—through the cross-border fragmentation of production processes within firms—and migration, in particular, of high-skilled expatriates, are highly correlated with multinational activity and with each other.[7] However, since trade data are more widely available than migration data, we focus on globalization measured as trade openness in most of the paper and relegate a discussion of migration openness to Section IV. Specifically, we use total manufacturing trade data from the United Nation's COMTRADE (Commodity Trade Statistics) database here and combine them with data for production and sales from UNIDO (United Nations Industrial Development Organization). Further details on the data and their sources are reported in the Appendix.[8]

Globalization is often found to be a source of higher before-tax income inequality (see Goldberg and Pavcnik 2004, 2007). This link may root in a number of different channels.[9] In this section, we focus on one mechanism that connects globalization

---

[7] Data on the 65 biggest economies used in this work suggest that between 1980 and 2007, imports to GDP ratio, inward FDI (foreign direct investment) stocks, and immigrant stocks rose by more than 16 percent, 272 percent, and 37 percent, respectively. While FDI rose most dramatically, its measurement is less precise and less uniform than that of trade and migration. This is due to differences in the definition of FDI across countries and differences in the accounting thresholds imposed by reporting countries which also affects the data provided by the OECD and the UNCTAD.

[8] As indicated above, although we compiled data on income tax calculators for 252 countries and territories for the period 1980–2012, the limiting factors here are the control variables such as (before-tax) income inequality and education levels. The analysis is thus limited to 65 countries between 1980 and 2007 (i.e., 1,820 observations on country-year-specific trade and other variables).

[9] These channels include (i) positive sorting of more productive workers to exporting firms (see Helpman, Itskhoki, and Redding 2010; Burstein and Vogel 2012; Harrigan and Reshef 2015, among others), (ii) firm heterogeneity and a high profitability of exporters combined with wage bargaining (see Egger and Kreickemeier 2009; Amiti and Davis 2012; Egger, Egger, and Kreickemeier 2013), (iii) offshoring of certain skill-specific stages of production (see Feenstra and Hanson 1996, 1999, 2004; Zhu and Trefler 2005), and (iv) capital-skill complementarity and associated skill premia (see Parro 2014). Autor et al. (2014) use individual-level US data and find that a higher exposure to imports from China has had a larger adverse effect on the demand for low-skilled workers relative to that of high-skilled workers, which suggests that globalization imposes heterogeneous costs on workers by skill

and after-tax inequality and the effect of greater openness on the distribution of personal labor income tax burdens across the gross wage distribution. In most of the paper, we present results for a male, single-earner household without spouse or children, which are quite representative. However, as a sensitivity check, we provide more detailed insights in the effects on the labor income taxation across the full range of possible household types.[10]

In this section, we assess changes in the distribution of labor income tax burdens across all percentiles of the wage income distribution before taxes for each country and year. It should be mentioned that some distributional effects on labor income tax burdens could be mechanical with progressive or regressive labor income tax schedules. If globalization had a homogeneous effect on prices, the entire income distribution would shift—with a reduction in prices, arguably to the left. However, as we will illustrate, this was largely not the case. Hence, countries must have adjusted their actual tax schedules to accommodate adverse distributional effects—inter alia of globalization—on gross wages, thereby exacerbating the dispersion of gross wages.

Our analysis requires detailed information on the distribution of gross labor incomes as well as on country-year-specific tax calculators. The LIS provides representative micro-level data on the labor income distributions for some countries in our sample. However, these data cover fewer than 65 countries over 28 years and are only available for 138 country-year pairings. Hence, microdata do not permit the consideration of percentiles of labor income distributions for detailed worker and household archetypes for many countries and years. Therefore, we make use of otherwise available data on Gini coefficients and average wages, which can be obtained for at least the 65 economies covered between 1980 and 2007 (see the Appendix for details). Upon making certain parametric assumptions about the functional form of the wage distribution, the Gini coefficient and the average wage are sufficient statistics to impute wages for all percentiles of the distribution per country and year. In particular, this is possible when assuming that wages follow either (i) a log-normal, (ii) a Pareto, (iii) or a mixture of the two distributions, where the upper tail is Pareto and the rest is log-normally distributed (see Nigai 2017). We provide details on this mixture and describe the imputation and calibration procedures for wage percentiles in the Appendix.

With the mixture distribution (dubbed *Mix* in tables), we first test three alternative weights for the Pareto tail, namely 5 percent, 3 percent, and 1 percent. Hence, we altogether consider five parametric distributions of before-tax labor incomes. Before turning to an analysis based on these data, we document how well the imputations match observed wage distributions. In Table 1, we report the correlation coefficients (and their standard deviations in parentheses) between the predictions of these five

---

and employment status. Autor, Dorn, and Hanson (2015), using local US labor market data, find that rising import competition, depending on industry characteristics, may lead to wage polarization and/or rising unemployment. All these arguments notably pertain to before-tax effects of globalization.

[10]Most tax calculators around the globe take the marital and family status of a worker, the employment status of cohabiting partners and spouses, and many other factors into consideration. Egger and Strecker (2018) provide a detailed description of 12 archetypal household types. However, the present study does not seek to address individual, behavioral responses to changes in income prospects. Rather, we implicitly hold behavior constant and assess how income prospects and particularly the tax calculator itself are affected by globalization. For this purpose, it is sufficient to mainly consider the effects for the single male earner archetype. In the sensitivity analysis, we consider the effects across the complete LIS microdata, which contain a host of household archetypes.

TABLE 1—LIS: IMPUTATION COMPARISON

	LIS	Pareto	log-normal	Mix (5%)	Mix (3%)	Mix (1%)
LIS	1					
Pareto	0.901 (0.058)	1				
log-normal	0.968 (0.033)	0.865 (0.009)	1			
Mix (5%)	0.979 (0.021)	0.935 (0.011)	0.984 (0.009)	1		
Mix (3%)	0.979 (0.022)	0.915 (0.009)	0.993 (0.004)	0.998 (0.001)	1	
Mix (1%)	0.974 (0.028)	0.887 (0.006)	0.999 (0.001)	0.992 (0.004)	0.997 (0.001)	1

parametric imputations of the distribution across percentiles of wage earnings and observed data sourced from the LIS.

The results in Table 1 indicate that the log-normal, the Pareto, and the mixture with a 5 percent Pareto tail align well with the microdata. Accordingly, we use these three distributions in the subsequent analysis for the imputation of average percentile-specific gross wage distributions per country and year. Specifically, we feed the obtained country-year-specific average gross wage of each percentile into the tax calculator for that country and year to obtain the country-year-percentile-specific effective average personal labor income tax rate. This results in two measures of interest:

(i) The nominal wage before personal labor income taxes paid to an average worker located between percentile p and the preceding percentile in the gross wage distribution in country i and year t, which we denote as $w_{i,t}^p$.

(ii) The associated effective employee-borne tax rate, which includes labor income taxes and employee-borne social security contributions, which we denote as $\tau_{i,t}^p$.

The measures of the effective employee-borne and employer-borne tax rates and contributions are defined according to the provisions in countries' tax codes. Hence, in the benchmark specification, we assume that workers and firms follow these provisions when splitting total labor income taxes and contributions.[11]

To that end, let us use $\pi_{ij,t}$ to denote the average of bilateral imports and bilateral exports between countries i and j normalized by the absorption of country i in year t. We consider the sum of $\pi_{ij,t}$ across all partners $j \neq i$ in that year as the measure of openness of country i in this section:

$$(1) \qquad \pi_{i,t} = \sum_{j \neq i} \pi_{ij,t}.$$

---

[11] In the online Appendix, we show that assuming firms are able to shift the burden onto workers does not change the results of interest to this paper, namely the globalization-induced fall in tax burdens for very high-wage earners.

Based on the measures of effective labor income tax rates, we compute the contribution (in percent) of each percentile $p = \{1, \ldots, 100\}$ to total employee-borne tax revenues, $100 \times \dfrac{\tau_{i,t}^{p} w_{i,t}^{p}}{\sum_{k=1}^{100} \tau_{i,t}^{k} w_{i,t}^{k}}$, and determine the associated impact of country i's trade openness in year t on the personal labor income tax burden for each wage percentile p as[12]

$$(2) \qquad 100 \times \frac{\tau_{i,t}^{p} w_{i,t}^{p}}{\sum_{k=1}^{100} \tau_{i,t}^{k} w_{i,t}^{k}} = \gamma^{p} \ln(\pi_{i,t}) + Z_{i,t} \Gamma^{p} + \eta_{t}^{p} + \mu_{i}^{p} + u_{i,t}^{p}$$

$$\text{for } p = \{1, \ldots, 100\},$$

where $Z_{i,t}$ is a vector of controls, $u_{i,t}^{p}$ is the disturbance term, the scalar γ^{p} (which is the quantity of key interest here) and the vector Γ^{p} are wage-percentile-specific regression parameters, and η_{t}^{p} and μ_{i}^{p} are fixed percentile-year and percentile-country effects, respectively. The vector of controls, $Z_{i,t}$, includes nine variables that can be categorized into three groups:

   (i)　Skill composition: three regressors for the share of population with primary, secondary, and tertiary education.

  (ii)　Political regime: three binary indicator variables for democracy and either left- or right-wing majorities in the legislature.

 (iii)　Demography and economy: log population, log real GDP per capita, and an interaction term of the two.[13]

In equation (2) the parameter of interest, γ^{p}, bears a percentile index. Apart from percentiles, the effect of globalization, and hence γ^{p}, may differ across time periods and country groups. For instance, there is a fundamental difference in the nature of globalization prior to the mid-1990s and afterward, and there is a similarly fundamental difference in the sophistication and nature of tax and social security systems between OECD and non-OECD countries. The change in globalization around the mid-1990s is evident in the change in importance of cross-border flows of workers and goods as well as the proliferation of international trade and investment agreements. For example, the Maastricht Treaty allowed European Union nationals to move freely within the European Union from 1992 onward; the Schengen Agreement reduced border barriers between European Union members from 1995 onward; the North American Free Trade Agreement led to a tremendous reduction in trade barriers between Canada, Mexico, and the United States from 1994 onward.[14] All these events relate to the mobility of relevant tax bases and suggest that examining the

---

[12] During 1980–1988, personal labor income taxes in Hungary were zero. For these observations, we set the burdens to one percentage point across all percentiles, so that the burden is equally distributed across percentiles. We also consider alternative approaches: coding the burden in those years as zero or replacing the burden with the wage share. The results remain robust.

[13] We report all data sources in the Appendix.

[14] See Caliendo and Parro (2014) for a quantitative assessment of the NAFTA effects.

effect of openness on income taxation for two separate time periods may be fruitful. To explore this issue, we specifically consider two different equally-sized subperiods, 1980–1993 and 1994–2007, between which we allow regression coefficients to differ and focus on the OECD country group. We also observe a significant change in the behavior of tax authorities between these two periods.[15] Regarding the split into OECD and non-OECD countries, consider the stark heterogeneity between these country groups in the composition of their tax revenues: OECD countries rely heavily on employee-borne taxes to raise revenues, while non-OECD countries lean more heavily on taxes on goods. See also Rodrik (1998) for a motivation to distinguish between OECD and non-OECD countries in this regard. For those reasons, we will consider the parameter estimates in equation (2) to be specific not only to the percentiles of the wage distribution but also to the two mentioned subperiods and country samples each.

A. *Ordinary Least Squares Results*

We start by estimating equation (2) using simple OLS and summarize the findings by way of figures, since there are 100 (percentile-specific) regression coefficients for each subsample. In each figure, we display the point estimates obtained for the contribution to total employee-borne labor income tax revenues obtained under the Pareto, log-normal, and the 5 percent mixture parameterizations of the wage distributions across countries and years along with 90 percent confidence bounds. The standard errors are clustered at the country and year levels.[16] The estimated percentile-specific coefficients $\hat{\gamma}^p$ reflect the response to globalization and are displayed in Figure 2, which is organized horizontally in two panels, with subperiod 1980–1993 in panel A and subperiod 1994–2007 in panel B.[17]

Figure 2 provides the following insights. In the first half of the covered time span, OECD countries collected more-than-proportionately higher labor income tax revenues from earners with higher, relative to earners with lower, wage incomes in response to greater trade openness. This is independent of how wage distributions are parameterized. The relationship was relatively flat and insignificantly different from zero for the lower half of the wage distribution in the early subperiod, while OECD countries leaned significantly more heavily on above-median-income earners in response to greater trade openness. Under the assumptions of the mixed distribution for wages, the coefficient estimate $\hat{\gamma}^p$ (with standard errors in parentheses) is 0.03 (0.04), 0.23 (0.16), and 3.92 (1.87) for the seventy-fifth, ninetieth, and one hundredth percentile, respectively. Between 1980 and 1993, the trade openness measure of globalization increased by about 42 percent in the average OECD country, which suggests a cumulative effect of trade openness on relative labor income

---

[15] We also test alternative cut-off years and the results are very similar for years close to the chosen cut-off, e.g., for 1994/1995, but not for years further way from 1993/1994. We postpone an in-depth discussion and formal evidence of the data-driven choice of the structural break year to Section IE.

[16] The standard errors are clustered by country and year as suggested in Cameron, Gelbach, and Miller (2006). The estimator asymptotically follows a normal distribution. We also follow Wooldridge (2002) in not making small sample adjustments.

[17] We report the coefficients of the control variables in the online Appendix.
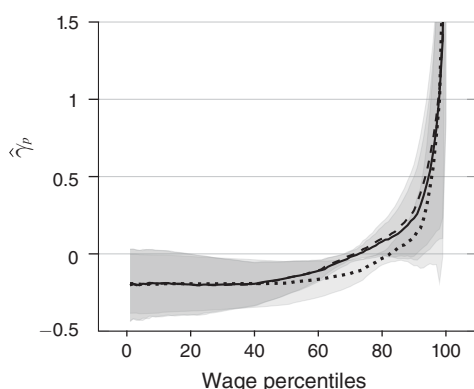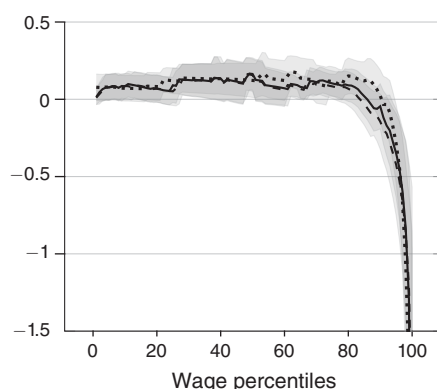
FIGURE 2. OLS COEFFICIENTS: TRADE OPENNESS AND PERCENTILE-SPECIFIC RELATIVE TAX BURDENS

tax burdens of 0.01, 0.10, and 1.65 percentage points for the seventy-fifth, ninetieth, and one hundredth percentile, respectively.

This pattern is consistent with two strands of existing work, as long as globalization did not shift the income distribution to the left. First, the recent trade literature predicts that lowering cross-border barriers to goods transactions raises the inequality among workers by benefiting highly productive workers relatively more (see Egger and Kreickemeier 2009; Helpman, Itskhoki, and Redding 2010; Nigai 2016). With limited cross-border mobility of labor, the optimal taxation literature suggests that an increase in real income inequality should be counteracted by increasing the taxes on high-income earners and redistributing across percentiles (see Saez 2001). If globalization shifted the wage schedule progressively, we would expect to see a response pattern in relative tax burdens consistent with panel A of Figure 2.

However, the pattern is qualitatively different for 1994–2007, as is evident from panel B of Figure 2. Again, the results of the Pareto, log-normal, and mixed wage imputations are very similar. In contrast to the earlier period, the results of the later one suggest an inverse-U-shaped locus for the relationship between the relative labor income tax burdens and wage percentiles in response to globalization. Clearly, the stark change in the relationship between the early and the later period cannot possibly be explained by a differential impact of globalization on the wages behind the percentiles alone but must be related to a change in effective tax rates. In the later half of the covered time period, we observe what we will refer to as the hollowing-out of the middle class: greater openness raised the tax burden on earners around the center of the wage distribution but not in the left or right tail. These results suggest that the top decile of the labor income distribution in fact benefited from trade openness relative to other income percentiles in terms of labor taxes owed, while the lowest quintile of the labor income distribution remained largely unaffected in relative terms. The coefficients (with standard errors in parentheses), which correspond to the seventy-fifth, ninetieth, and one hundredth gross wage percentiles under the

assumptions of the mixed distribution of wages, are 0.10 (0.05), -0.05 (0.09), and -3.67 (1.63), respectively. Between 1994 and 2007, the trade openness measure of globalization increased on average by 16 percent. This suggests that from 1994 to 2007, globalization led to an increase of the labor income tax burden of 0.02 percentage points for the seventy-fifth percentile, while reducing the burden by 0.01 and 0.59 percentage points for earners in the ninetieth and one hundredth percentiles, respectively. These results are consistent with models in the optimal taxation literature, which assume high-income earners to also be highly mobile (see Simula and Trannoy 2010; Lehmann, Simula, and Trannoy 2014).

One may view the estimating equation (2) based on OLS as problematic given the potential endogeneity of globalization and taxes and/or the omitted variable bias. For example, the production of goods depends on (gross-of-tax) factor returns, the consumption of goods depends on (net-of-tax) disposable income, and the location of firms and workers depends on expected net factor returns and the provision of public goods financed by tax revenues. These potential shortcomings of the OLS estimator call for an instrumental variables (IV) approach. We pursue such an approach using two distinct types of instruments for globalization in the next subsection. Before proceeding, let us emphasize that the potential endogeneity concerns between trade and taxes may be less relevant in the context of equation (2), where the dependent variable, $100 \times \frac{\tau_{i,t}^p w_{i,t}^p}{\sum_{k=1}^{100} \tau_{i,t}^k w_{i,t}^k}$, is measured in relative, rather than absolute, terms. Therefore, while unnormalized percentile-specific labor tax revenues, $(\tau_{i,t}^p w_{i,t}^p)$, may affect trade through the aforementioned channels, the bias of the coefficient of interest, $\hat{\gamma}^p$, might be reduced or even eliminated with the normalization of the dependent variable, if the effect of $(\tau_{i,t}^p w_{i,t}^p)$ on $\pi_{i,t}$ was approximately equal across the percentiles of the labor income distribution. Notwithstanding this argument, we will next turn to an instrumental variables approach.

B. *Instrumental Variables* (*IV*) *Results*

In this section, we create two distinct types of identifying instruments for trade openness, $\ln(\pi_{i,t})$. Both of them are designed to capture relative measures of trade costs, but they do so in very different ways. We generally refer to these instruments for country i's openness in year t as $\lambda_{i,t}^\kappa$, where $\kappa = \{I, II\}$ indexes the instrument type.

Instrument I: Exploring the Structure of Quantitative Trade Models.— Instrument I relies on the generic structure of quantitative general equilibrium models of trade, which permits the structural calibration of symmetric country-pair-year-specific trade costs from trade data without error conditional on the model structure (see Eaton and Kortum 2002; Anderson and van Wincoop 2003; Arkolakis, Costinot, and Rodríguez-Clare 2012). Such models impose the following three assumptions:

(i) Seller-producers do not segment markets so that mill prices are independent of the location of customers. Producers are either perfectly competitive and make zero profits or charge a constant, positive mark-up over marginal costs.

(ii) Trade costs are of the iceberg type as defined in Samuelson (1954).

(iii) Aggregate demand and its allocation can be characterized by a two-stage budgeting problem; the problem of total expenditure determination and the allocation of expenditures across consumption items (products) can thus be separated.

Prominent examples of such models include Eaton and Kortum (2002), Anderson and van Wincoop (2003), and Chaney (2008), and they are characterized in Arkolakis, Costinot, and Rodríguez-Clare (2012).[18] The three assumptions ensure that bilateral consumption shares toward country j by consumers in country i and year t, $\pi_{ij,t}$, are multiplicative in exporter-year-specific, importer-year-specific, and pair-year components:

$$(3) \qquad \pi_{ij,t} = \underbrace{e_{j,t}}_{\text{export-year component}} \times \underbrace{\iota_{i,t}}_{\text{importer-year component}} \times \underbrace{\beta_{ij,t}}_{\text{pair-year component}},$$

where $e_{j,t}$ is proportional to country j's supply potential and $\beta_{ij,t}$ captures the influence of trade costs or preferences in country i against acquiring goods from j in year t. Though the exact interpretation of $e_{j,t}$ depends on the model in mind, it broadly captures production costs and (gross-of-tax) factor income and is influenced by labor income taxation. In the models referred to above, $\iota_{i,t}$ is a function of the consumer price index in country i and year t. Both $e_{j,t}$ and $\iota_{i,t}$ might capture country-year-specific trade costs. However, the pair-specific component $\beta_{ij,t}$ is free of any country-year-specific influence if $e_{j,t}$ and $\iota_{i,t}$ are estimated as fixed effects in empirical models. Within the same class of models as above, $\beta_{ij,t}$ captures trade frictions across country pairs and, eventually, time. Our strategy is to exploit the multiplicative model structure about $\pi_{ij,t}$ in terms of the above three components and recover $\beta_{ij,t}$. For this, we utilize the procedure of Head and Ries (2001), who assume that transaction costs between domestic sellers and domestic customers are zero, so that $\beta_{ii,t} = 1$.[19] In pursuit of this strategy, we first eliminate the importer-year component by normalizing each trade share by the share of the importer's consumption from domestic sellers:

$$(4) \qquad \frac{\pi_{ij,t}}{\pi_{ii,t}} = e_{j,t} \times e_{i,t}^{-1} \times \beta_{ij,t}.$$

Next, we specify the same ratio for i as an exporter and j as an importer in year t and then compute

$$(5) \qquad \frac{\pi_{ij,t}}{\pi_{ii,t}} \frac{\pi_{ji,t}}{\pi_{jj,t}} = \beta_{ij,t} \beta_{ji,t}.$$

---

[18] In trade models with heterogeneous firms (see Melitz 2003) where productivity parameters are drawn from a Pareto distribution as in Chaney (2008), the fixed costs of exporting enter the bilateral trade equation in a log-additive fashion and are isomorphic to iceberg trade costs for our purposes.

[19] The assumption of zero domestic transaction costs is very common in structural quantitative models of bilateral trade; see, e.g., Eaton and Kortum (2002) or Anderson and van Wincoop (2003).

Note that the right-hand side of (5) is free of any exporter-year and importer-year components and only captures average bilateral frictions between countries i and j in year t. We use $\beta_{ij,t}\beta_{ji,t}$ to compute the *average it-specific costs of exporting and importing* as

$$(6) \qquad\qquad\qquad \lambda_{i,t}^I = \sum_{j\neq i} \beta_{ij,t}\beta_{ji,t}.$$

This instrument is valid, if $\beta_{ij,t}\beta_{ji,t}$ is exogenous to exporter-year and importer-year factors. This means that the distribution (not the level) of trade costs among country pairs in a year is not influenced by the level of, e.g., income or tax revenues.

Since all exporter-year and importer-year variables are excluded from (5) by design, it is very unlikely that $\lambda_{i,t}^I$ is influenced by labor income taxes, even less so by the contributions pertaining to individual percentiles of the wage income distribution. Though Instrument I is appealing from a theoretical standpoint, it has certain limitations: in particular, its validity depends on the aforementioned generic structure of customary gravity models of trade. While these models cover a broad range of cases, it is important to check the robustness of the results using an alternative identification strategy.

Instrument II: Exploring the Variation in Oil Prices and Transportation Distances.—Instrument II explores the variation in the average annual import prices of crude oil across the countries and years in our sample interacted with a country-specific measure of access to international markets which we describe in what follows. Crude oil import prices include the cost, insurance, and freight of oil but do not include import duties. These data come from the International Energy Agency Crude Oil Import Register and are described in more detail in the Appendix.

Let us use $p_{i,t}^{oil}$ to denote average crude oil import prices in country i and year t.[20] Moreover, let us use d_i^1, d_i^2, and d_i^3 to denote the distances of the three largest (most populated) cities in i to the respective closest maritime port. We describe the distance data in detail in the Appendix. Specifically, to construct Instrument II we use the interaction between the oil prices and distances as a measure of transportation costs:

$$(7) \qquad\qquad D_{i,t}^k = p_{i,t}^{oil} \times d_i^k \quad \text{for } k = \{1,2,3\},$$

which can be viewed as a proxy for transport costs from city k to the nearest port in country i and year t. Given $D_{i,t}^k$, we define the average measure $\bar{D}_{i,t} = \frac{1}{3-1}\sum_{k=1,2,3} D_{i,t}^k$ and specify Instrument II as

$$(8) \qquad\qquad \lambda_{i,t}^{II} = \frac{1}{3-1}\sum_{k=1}^{3}\left(D_{i,t}^k - \bar{D}_{i,t}\right)^2,$$

which measures the variance of transporting goods from and to the three most important business hubs in a country through sea ports to and from trading partners

---

[20] Oil prices are measured in purchasing power parity adjusted international dollars. We provide more details in the Appendix.

for each country and year. This measure has several advantages. First, it does not rely on bilateral trade data but rather directly measures within-country transport costs of imports and exports. Second, the variance of distance to the closest sea port across several cities in each country also provides a measure of the internal geography of a country which is an important component of transportation costs. This measure is inversely related to various measures of spatial concentration. One could use the average *level* of internal distances to the port instead of the variance within a country to generate the instrument and reach qualitatively similar results. However, unlike the variance, the average distance generally does not capture the level of spatial concentration and it is not as strong in terms of explanatory power in the first stage. For this reason, in our benchmark specification we use the instrument based on the variance rather than the average level.[21]

Instrument II is very different from Instrument I, since it does not rely on quantitative trade theory and is valid under a very different set of assumptions. In particular, Instrument II is valid as long as foreign crude oil import prices are not correlated with the personal income tax burdens in an importing country. We consider crude oil import prices for the average economy to be plausibly exogenous.

C. *First-Stage Results*

In this subsection, we discuss how Instruments I and II relate to the measure of globalization. Formally, in the first stage, we estimate the following regression:

$$(9) \qquad \ln(\pi_{i,t}) = \varphi^\kappa \lambda_{i,t}^\kappa + Z_{i,t}\Xi^\kappa + \eta_t^\kappa + \mu_i^\kappa + \upsilon_{i,t}^\kappa \quad \text{for } \kappa = \{I, II\},$$

where φ^κ denotes the coefficient on the respective identifying instrument type $\kappa \in \{I, II\}$, $Z_{i,t}$ is the same vector of controls as included in the outcome equation (2), Ξ^κ are the parameters on it, and η_t^κ and μ_i^κ denote the year and country fixed effects, respectively. Full regression results for the respective first-stage parameters are available in the online Appendix. Here, we resort to reporting the weak-instrument test F-statistic for the first stage in Table 2. Since we run regressions separately for two different periods, we report the relevant weak-instrument F-statistic separately for 1980–1993 and 1994–2007.

The results in Table 2 suggest that both instruments are strong. However, Instrument I appears to be relatively stronger which is intuitive as it relies on calculating trade costs from bilateral trade data, whereas Instrument II does not rely on any trade data at all. Nevertheless, both instruments pass the weak-instrument test at customary levels and will be used in the subsequent analysis.

For illustration, we also plot trade openness as predicted by $\hat{\lambda}_{i,t}^I$ and $\hat{\lambda}_{i,t}^{II}$ together with the other variables in equation (9) against the data on trade openness in two panels of Figure 3. The data and predictions are plotted in first differences. The R^2 statistics reported in Figure 3 are also calculated using first differences. Each panel pertains to one specific instrument and includes, apart from the

---

[21] One could also employ the variance and the average of the distances to the closest sea port together or population-weighted distances to normalize $D_{i,t}^k$. The results are robust to either of these choices.

TABLE 2—FIRST-STAGE KLEIBERGEN-PAAP WALD F-STATISTIC

	1980–1993	1994–2007
Instrument I	41.66	55.33
Instrument II	16.53	10.28

Note: The reported statistics are robust Kleibergen-Paap Wald rk F-statistics adjusted for country and year clustering.
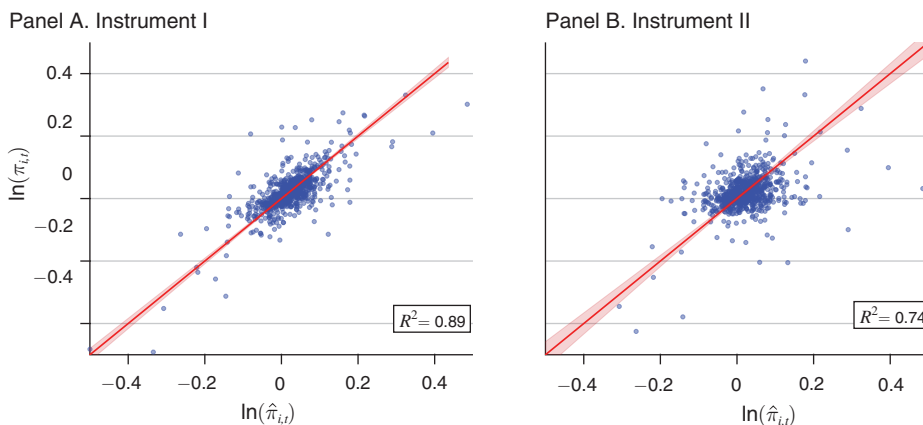

Figure 3. First-Stage Results

Note: The data on trade openness and their predictions are plotted in first differences.

individual, country-year-specific data points, the linear fit as a solid line with 90 percent-confidence bounds.

Figure 3 confirms that the instruments lead to strong first-stage results, even when focusing on first differences.

D. *Second-Stage Results*

Based on the two alternative first-stage regressions, we estimate the second-stage model corresponding to the outcome equation of interest in (2) using percentile-specific relative tax burdens as the dependent variable. Again, we choose to present the results graphically, as only the parameter estimate $\hat{\gamma}^p$ on log openness, $\ln(\pi_{i,t})$, is of key interest here. The results are presented in Figure 4. As before, we plot the loci of the estimated coefficients together with 90 percent-confidence bands around them. We estimate the second-stage regressions using two-stage least squares (2SLS).

Figure 4 is organized in four panels. The upper row plots the results using the two instruments for the first period, 1980–1993, and the lower row reports the same for the second period, 1994–2007. The results in Figure 4 are for OECD countries only. The results for non-OECD countries are available in the online Appendix. As with OLS before, we will discuss the results using the example of the mixture of a log-normal distribution with a 5 percent Pareto tail. The second-stage results are qualitatively consistent with OLS and confirm that higher trade openness increased the progressivity of the relative labor income tax burdens in OECD countries in the first subperiod considered. The respective estimated coefficients (with standard

Panel A. OECD: 1980–1993 Instrument I

Panel B. OECD: 1980–1993 Instrument II

Panel C. OECD: 1994–2007 Instrument I

Panel D. OECD: 1994–2007 Instrument II
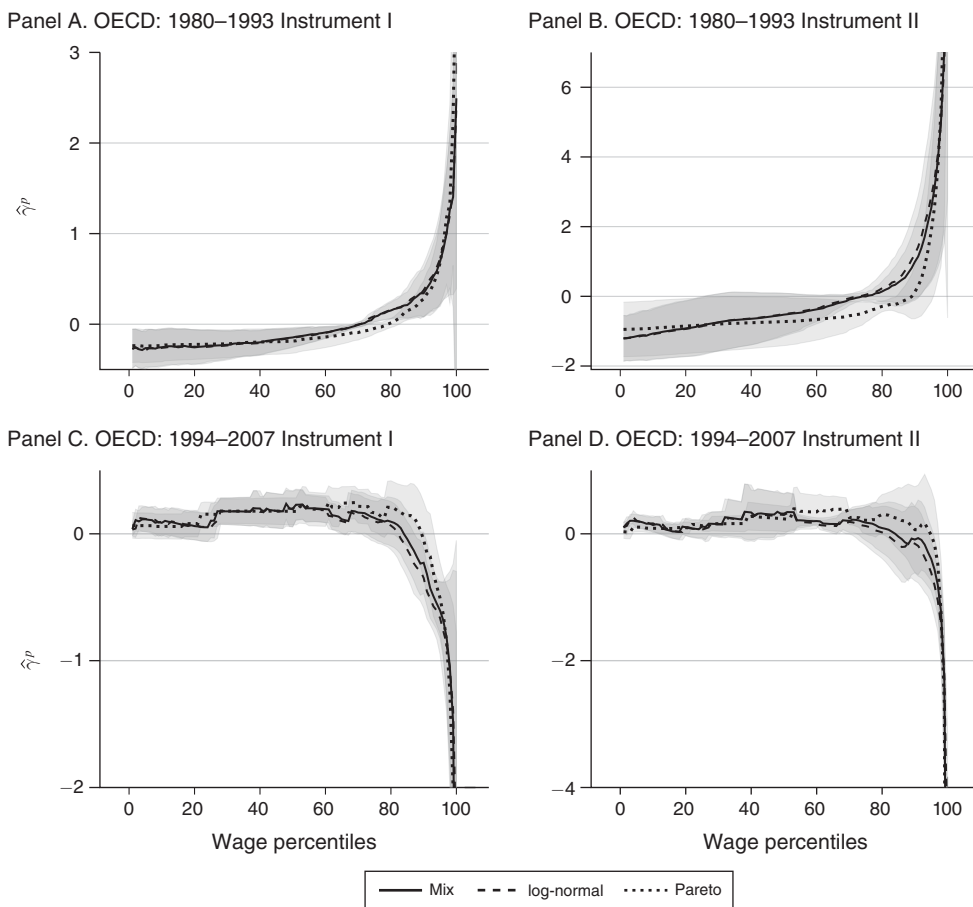
Mix — — — log-normal ⋯⋯⋯ Pareto

Figure 4. IV Coefficients: Trade Openness and Percentile-Specific Relative Tax Burdens

errors in parentheses) based on the first instrument are 0.08 (0.03), 0.34 (0.16), and 2.49 (2.08) for the seventy-fifth, ninetieth, and one hundredth percentile, respectively. The results are qualitatively identical and quantitatively larger for the second instrument. Its estimated coefficients for the three percentiles above are 0.00 (0.07), 0.99 (0.64), and 15.82 (3.98).

We report detailed results for the seventy-fifth, ninetieth, and one hundredth percentiles for both instruments in Table 3. Further details on the regression results for the representative seventy-fifth percentile are available in the online Appendix. The estimated coefficients translate into a large quantitative impact of globalization in the earlier subperiod. For example, according to the results using Instrument II, the relative tax burden for the one hundredth percentile in the average country increased by 6.64 percentage points between 1980 and 1993 in response to globalization.

In the second subperiod, we see a positive effect of globalization on the relative labor income tax burden of the middle-income class and a negative effect on the relative labor income tax burden of the very-high-income earners. In terms of the absolute value of point estimates, the second-stage results are larger for almost all considered percentiles. For example, under the assumption of the mixed distribution the estimates based on Instrument II are 0.18 (0.17), −0.09 (0.32), and −9.06

Table 3—Trade Openness and Percentile-Specific Relative Tax Burdens

		1980–1993			1994–2007		
		75th	90th	100th	75th	90th	100th
IV-I	$\hat{\gamma}^p$	0.08	0.34	2.49	0.12	−0.23	−4.28
		(0.03)	(0.16)	(2.08)	(0.10)	(0.15)	(2.41)
IV-II	$\hat{\gamma}^p$	0.00	0.99	15.82	0.18	−0.09	−9.06
		(0.07)	(0.64)	(9.93)	(0.17)	(0.32)	(2.13)

Note: Standard errors in parentheses are robust to an unknown form of heteroskedasticity and autocorrelation and clustered by country and year.

(2.13) for the seventy-fifth, ninetieth, and one hundredth, respectively. The results for Instrument I are also reported in Table 3. The estimated coefficients suggest that in the second subperiod the relative tax burden of the highest-income earners in an average OECD country fell by about 1.45 percentage points in response to the total change in trade openness.

Overall, the results of the second stage suggest that between 1980–1993 and 1994–2007 there was a major change in how the progressivity of labor income tax schedules responded to increased globalization. In the first subperiod, OECD countries increased the progressivity of their tax schedules in response to high openness such that the tax burden of very-high-income workers and employees grew relatively more. However, in the second subperiod, those same earners were apparently treated as relatively more footloose, and their relative tax burdens decreased in response to globalization.

E. *On the Choice of the Structural Break Year*

We have presented results for percentile-specific relative labor income tax burdens for two subperiods, 1980–1993 and 1994–2007. At the beginning of Section II, we argued that a structural break year somewhere in the mid-1990s was reasonable considering the far-reaching liberalizations across the OECD in the form of preferential liberalization of trade and factor flows. In this section, we test the choice of the break year against a backdrop of data on liberalization steps and statistical tests for a range of possible structural break years within the panel of country-year data at hand.

First of all, the mid-1990s saw a surge in the number of bilateral agreements signed among OECD countries that addressed preferential trade, bilateral investment, and the double taxation of foreign-earned income. For example, between 1980 and 1993, an average pair of OECD countries in our sample had 1.6 such bilateral agreements in force, whereas between 1994 and 2007 this number increased to 2.6.[22] In the latter period, many such agreements also increased in depth (covering more issues beyond just tariffs).

Second, we conduct formal tests on the choice of the structural break year, where we consider nine alternative potential break years between 1989/1990 and 1997/1998. To make results comparable, we only consider observations within

---

[22] We consider preferential trade agreements, bilateral investment treaties, and double taxation treaties, using data from Egger and Tarlea (2017).
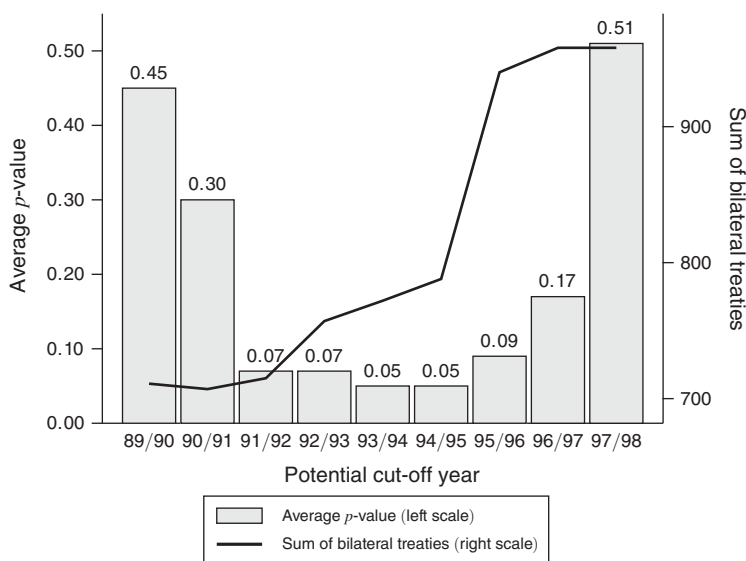
FIGURE 5. AVERAGE p-VALUES OF COEFFICIENTS BY POTENTIAL BREAK YEAR (LEFT AXIS) AND SUM OF ACTIVE BILATERAL TREATIES (RIGHT AXIS)

10-year windows to the left as well as the right of the respective considered break year. For instance, for a possible break in 1989/1990, we examine the period 1980–1989 and 1990–1999. For each time window, we estimate the benchmark equation (2) using our preferred specification which relies on Instrument I and compare the coefficient of interest on trade openness across the percentiles between the ninetieth and one hundredth percentile, as these are the percentiles where we see the largest significant changes. Let us denote the coefficients for one estimation model and break year for the pre-break and post-break period by $\hat{\gamma}^p_{pre}$ and $\hat{\gamma}^p_{post}$. For each percentile $p \in \{90, 100\}$, we then conduct a Wald test under the null hypothesis that the two coefficients are equal, $H_0 : \hat{\gamma}^p_{pre} = \hat{\gamma}^p_{post}$, and also not different from zero, which obtains a p-value for each wage distribution, and percentile.[23]

In Figure 5, we plot two types of information on the left and right vertical axes against different break years between 1989/1990 and 1997/1998 on the horizontal axis. On the right axis, we plot the total number of active bilateral (trade, investment, and double taxation) agreements in the OECD in each year. It is obvious that the number of the considered preferential agreements in force increased dramatically from 1994 onward. On the left axis, we plot the average p-value of the equality-of-pre-post-break-year coefficient tests about γ^p across the top-five percentiles, assumed forms of the wage distribution, and estimation methods per potential year of the break point. The results suggest that, for the average percentile, the structural break years that lead to the rejection of the null with 95 percent confidence are

[23] Because our procedure decreases the number of years available for estimation, we cluster standard errors by country.

1993/1994 and 1994/1995, where the average p-values are below 0.05. These are also the years the surge of preferential agreements begins.[24]

II. The Effect of Globalization on the Size and Composition of Tax Revenues

The results in the previous section suggest that globalization since the mid-1990s led to a significant increase in the labor income tax burden of middle-income earners especially relative to high- and top-income earners. In this section, we show that this shift in labor income tax systems went hand in hand with a change in the volume and general composition of total tax revenues in OECD countries.

This analysis is motivated by a variety of earlier findings. For instance, Rodrik (1997, 1998) suggests that an increase in countries' openness raises citizens' demand for public goods and changes the prices at which governments can provide them. Accordingly, we would expect to see an increase in tax revenues in response to globalization, at least in the long run, in order to finance the increased real consumption of public goods. Moreover, the findings of Devereux, Lockwood, and Redoano (2008) suggest that an increase in capital (firm) mobility across national borders would entail stiffer competition between governments for capital and firms leading to reductions in the tax burden on mobile firms and capital. Altogether, globalization should lead to increasing taxes on less mobile factors (value added, goods consumption, etc.) relative to more mobile ones. Hence, we would expect to see a change in the structure of tax revenues with governments having reduced taxes on firms while having increased taxes on (average) labor, goods, and value added since the mid-1990s.

Here, we are interested in running regressions akin to equation (2) with the exception that we do not consider wage-percentile-specific but tax-revenue-category-specific coefficients on openness. Let us use R^r with $r = \{total, firm, employee, goods, other\}$ to denote government tax revenues of type r, where the categories r refer to *total* tax revenues, revenues from *firm*-borne taxes, *employee*-borne taxes, taxes on *goods* and services such as sales and value-added-type taxes, and *other* tax revenues, respectively. For these categories, we specify the following linear regression model:

$$(10) \qquad 100 \times \frac{R^r_{i,t}}{GDP_{i,t}} = \psi^r \ln(\pi_{i,t}) + Z_{i,t}\Gamma^r + \lambda^r_t + \mu^r_i + u^r_{i,t}.$$

As before, we estimate (10) using OLS and 2SLS for the two alternative instruments. We report results for the two subperiods in two vertical panels of Table 4. Each of the five columns in a panel presents the results of a tax revenue category with each row presenting the results of OLS or 2SLS with the two alternative instruments. We restrict the presentation of the results to the trade openness parameter (with the standard error in parentheses), ψ^r, and the R^2 of the respective second-stage regression. The table focuses again on OECD countries. The results for non-OECD countries and the coefficients of the control variables are relegated to the online Appendix.

---

[24] We find that assuming a break year of 1994/1995 instead of 1993/1994 obtains almost identical globalization-induced effects on labor income taxation to the benchmark analysis, while this is not the case for more distant structural break years.

TABLE 4—TRADE OPENNESS AND TAX REVENUES

	$R_{i,t}^r$	1980–1993				
		Total	Firm	Employee	Goods	Other
OLS	$\hat{\psi}^r$	−0.24	0.37	−0.22	−0.01	−0.38
		(2.75)	(1.12)	(1.54)	(1.46)	(0.36)
	R^2	0.09	0.11	0.02	0.23	0.03
IV-I	$\hat{\psi}^r$	3.19	2.76	−1.20	2.60	−0.97
		(4.23)	(1.09)	(2.44)	(1.55)	(0.73)
	R^2	0.09	0.11	0.02	0.24	0.03
IV-II	$\hat{\psi}^r$	1.46	0.12	−3.31	3.89	0.76
		(5.26)	(2.32)	(3.62)	(2.75)	(1.46)
	R^2	0.08	0.10	0.02	0.23	0.03
Observations		330	330	330	330	330
Countries		25	25	25	25	25
		1994–2007				
OLS	$\hat{\psi}^r$	5.42	2.00	0.87	5.13	−2.58
		(2.85)	(2.22)	(1.51)	(0.97)	(1.19)
	R^2	0.13	0.20	0.05	0.16	0.03
IV-I	$\hat{\psi}^r$	7.17	−1.11	5.40	6.20	−3.33
		(3.53)	(2.59)	(2.14)	(1.21)	(1.38)
	R^2	0.20	0.20	0.10	0.40	0.12
IV-II	$\hat{\psi}^r$	3.13	−5.07	3.53	12.69	−8.01
		(4.04)	(5.09)	(2.38)	(1.55)	(1.35)
	R^2	0.13	0.21	0.05	0.32	0.11
Observations		350	350	350	350	350
Countries		25	25	25	25	25

Notes: Standard errors in parentheses are robust to an unknown form of heteroskedasticity and autocorrelation and clustered by country and year. For brevity, the constant, time, and country fixed effects and coefficients for education shares, democracy, left- and right-wing governments, log population, log real GDP per capita, and the interaction between log population and log real GDP per capita are suppressed. The reported R^2 is adjusted and corrected for two-way fixed effects using the Frisch-Waugh-Lovell theorem.

The findings in the upper panel of Table 4 suggest that in the early (pre-liberalization) period the effect of globalization on the relative size of total tax revenues cannot be precisely estimated for OECD countries. This is the case for all three econometric model types (OLS, IV-I, and IV-II). On the other hand, there is evidence that the late (post-liberalization) period was characterized by increases in the relative size of total revenues from higher employee- and goods-based taxation in response to globalization. In terms of an effect on the relative size of total revenues, the estimated coefficient of interest remains positive across all estimation models; however, it is only statistically significant for the first instrument. The same applies to the share of employee-based tax revenues. Finally, the results for goods-based tax revenues are consistently positive and statistically significant across all three specifications. Overall, the effect of openness on the relative size of tax revenues became more pronounced during the more recent time period for OECD countries. This increase is not due to greater firm-borne taxes but rather due to increases in employee- and goods-based tax revenues. As established in the previous section, the increase in employee-borne labor income taxes was mainly due to higher tax burdens on the upper-middle income class rather than on top-income earners.

III. Robustness Checks

In this section, we assess the robustness of the key results along four lines. We (i) use non-imputed microdata on wage distributions of single-male, single-earner households from the Luxembourg Income Study; (ii) use the full set of Luxembourg Income Study microdata for all household types, taking within-country differences in marital and employment status and the number of children into account; (iii) decompose the total effect of globalization into wage- and tax-rate-based effects; (iv) use migration-stock openness as an alternative measure of globalization; and (v) shed light on the effects of state-to-state migration on state- and federal-level taxes in the United States using microdata.

A. *Using Microdata from the Luxembourg Income Study*

The analysis in the main part of the paper was conducted based on imputed data on gross wages assuming three alternative parametric distribution functions in order to cover as many countries and years as possible. In this subsection, we rely fully on observed, rather than imputed, wage data and abandon any assumption on the functional form of the wage distribution. Instead, we establish country-year-specific wage distributions from Luxembourg Income Study (LIS) microdata. Once again, we first focus on single male earners, computing effective average labor income tax burdens for each percentile of the wage distribution using the tax calculator for each country and year.

However, comparable and representative LIS microdata that can be used in the analysis are only available for 77 OECD-country-year observations in the 1994–2007 subperiod.[25] The purpose of using microdata is to check whether the results in Section I remain consistent and are not an artifact of the imputation.

We repeat the analysis from Section I using the (much scarcer) LIS data on single male earners in Figure 6, where the three panels present the results of OLS and the two IV estimators, respectively. Since statistical agencies are likely to cap reported incomes to protect the earners' privacy and the likelihood that earners in the top-1 percent of the labor income distribution are undersampled, the wages in the top quantile(s) are likely underestimated.[26] Accordingly, we should underestimate (rather than overestimate) the effect of openness on the tax burden for the very top wage quantile(s) in this analysis. However, the three panels in Figure 6 suggest that the general shape and magnitude of the results in the much smaller LIS data sample are consistent with the effects found in the larger, imputed wage dataset.

---

[25] If we include non-OECD countries and all years in our sample, the LIS sample size will contain 151 country-year observations. However, for the estimation in this section we can only use OECD countries in 1994–2007. Even among the 77 country-year units in the OECD, there are ones, where the LIS individual-level data samples are too small to compute 100 percentiles of the wage distribution. In these cases, we linearly interpolate the data between the available quantiles. We are furthermore restricted to our second subperiod as even fewer (only 61) OECD-country-year observations are available between 1980–1993. Finally, there are virtually no non-OECD countries covered for the analysis at hand, illustrating the current limitations of available LIS microdata for cross-country analysis.

[26] For some countries and years, the very right tail of the wage distribution had to be imputed due to data limitations. To that end, one currency unit was added to the previous percentile's wage.
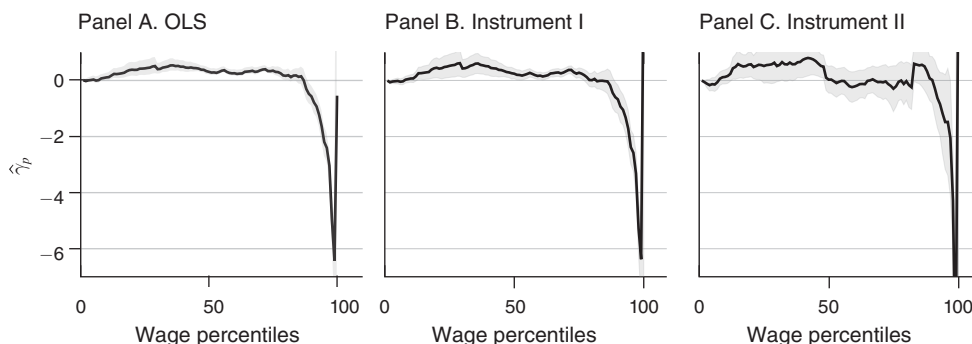
FIGURE 6. REGRESSION COEFFICIENTS: TRADE OPENNESS AND PERCENTILE-SPECIFIC RELATIVE TAX BURDENS IN LIS MICRODATA (SINGLE-MALE, SINGLE-EARNER HOUSEHOLDS) FOR OECD COUNTRIES

This suggests that the results in Section I are not an artifact of the imputation and are in fact supported by available microdata.

B. *Results for Alternative Household Types*

Recall that the benchmark results were obtained for single-earner, single-male households. However, it is important to check the robustness of the main results across all wage-earning household types, because effective tax rates may depend on marital status and the number of children in a household. To that end, we again employ LIS data, this time keeping all households in the sample. We then calculate total household labor income and account for percentile-specific household characteristics (main and second earner incomes, marital status, and the number of children) in a given country and year.[27] We again feed these data into our set of tax calculators to obtain the effective average tax burden for the average household in each percentile. Finally, we use these tax burdens to estimate a version of equation (2) which captures more household types than the benchmark regression. The results are reported in Figure 7.

Given that we do not impose any restrictions on the households included in the regression samples, the results are quite noisy across percentiles, as the frequency of different household types per percentile varies across countries and time. To filter the noise to some extent, we report the response functions across percentiles smoothed via a polynomial regression, along with the original estimates. The corresponding results may be interpreted as to represent smoothed weighted averages.

Figure 7 suggests that globalization generally affected the relative percentile-specific tax burdens of all wage-earning households in the same qualitative way as those of single-male, single-earner households. In 1994–2007, an increase trade openness generally lowered tax burdens for those in the right tail of the wage distribution. Hence, the main insights do not appear limited to single-male, single-earner households but extend to all household types in the income distribution of OECD countries.

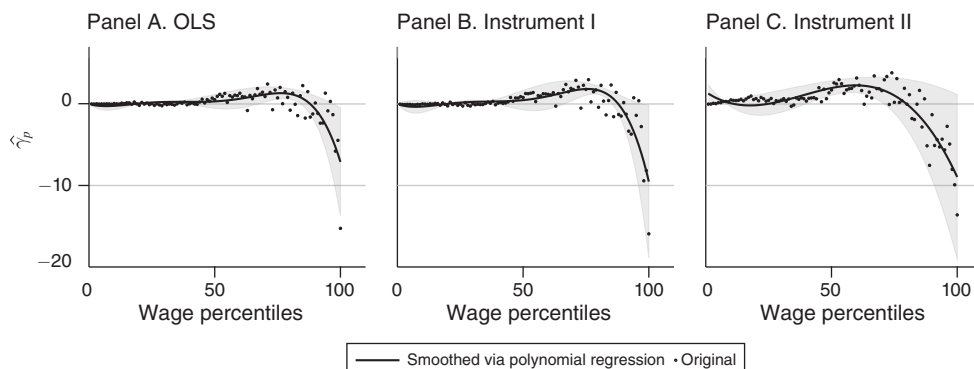[27] As before, we are limited to the years 1994 to 2007 for reasons of data availability.

FIGURE 7. REGRESSION COEFFICIENTS: TRADE OPENNESS AND PERCENTILE-SPECIFIC RELATIVE TAX BURDENS IN
LIS MICRODATA (ALL HOUSEHOLDS) FOR OECD COUNTRIES

C. *Decomposition of the Globalization Effects on Relative Labor Income Tax Burdens*

The relative labor income tax burdens in Section I build on two components: percentile-specific gross wages and the respective labor income tax rates. In this section, we decompose our earlier results and examine the links between globalization and before-tax wages on the one hand and labor income tax rates (or tax calculators) on the other hand. The purpose of this exercise is to document that the results are primarily driven by globalization-induced changes in tax calculators rather than by changes in the distribution of earnings or average wages.

We start by examining the link between globalization and nominal wages before taxation, specifying the nominal gross wage in percentile p, country i, and year t, $w_{i,t}^p$, as a multiplicative function of the average gross wage, $\bar{w}_{i,t}$, and the percentile-specific dispersion parameter, $s_{i,t}^p$:

$$(11) \qquad w_{i,t}^p = \bar{w}_{i,t} \cdot s_{i,t}^p.$$

Effects of globalization on the average wage in a country and year, $\bar{w}_{i,t}$, are standard in virtually all multi-country models of trade and well documented in the literature (see Eaton and Kortum 2002; Arkolakis, Costinot, and Rodríguez-Clare 2013). We run the following regression to isolate the effect of globalization on average wages:

$$(12) \qquad \ln(\bar{w}_{i,t}) = \gamma^c \ln(\pi_{i,t}) + Z_{i,t}\Gamma^c + \lambda_t^c + \mu_i^c + u_{i,t}^c.$$

We use OLS and the two IV specifications and find a positive, in the case of the IV specifications also statistically significant, effect of globalization on the average annual wage.[28]

---

[28] The results are relegated to Table OA.5 in the online Appendix.

We run similar regressions on the wage share, $s_{i,t}^p$, using OLS and our two IV specifications:[29]

$$(13) \quad s_{i,t}^p = \gamma^s \ln(\pi_{i,t}) + Z_{i,t}\Gamma^s + \lambda_t^s + \mu_i^s + u_{i,t}^s \quad \text{for } p = \{1, \ldots, 100\}.$$

However, the regressions in (13) do not point to statistically significant effects of globalization on the dispersion of gross wages around the country-year averages. This may be due to the fact that $s_{i,t}^p$ is not perfectly aligned with skill levels (see Feenstra and Hanson 1996, 1999) or the distribution of firms in terms of their productivity across gross wage percentiles (see Egger and Kreickemeier 2009, Amiti and Davis 2012).

However, effective labor income tax rates across percentiles depend on the average wage, which itself depends on trade openness and the tax calculator. To isolate the effect of globalization on the tax calculator, we compute a counterfactual wage distribution for each country and year that removes any change in wages related to trade openness according to the specification in (12). While there are multiple approaches to this isolation, our preferred approach computes percentile-specific counterfactual wages for each estimation model assuming a hypothetical as-of-1980 level of trade openness:

$$(14) \quad \tilde{w}_{i,t}^p = \bar{w}_{i,t} \exp\left(\hat{\gamma}^c\left[\ln(\pi_{i,1980}) - \ln(\pi_{i,t})\right]\right) \cdot s_{i,t}^p.$$

In logs, this subtracts $\gamma^s \ln(\pi_{i,t})$ from $\ln(\bar{w}_{i,t})$ and adds $\gamma^s \ln(\pi_{i,1980})$ to produce $\ln(\tilde{w}_{i,t})$. After exponentiation, this counterfactual average wage is then multiplied by the percentile-specific wage share, $s_{i,t}^p$, to produce the counterfactual gross (before-tax), percentile-specific wage, $\tilde{w}_{i,t}^p$. We subsequently feed $\tilde{w}_{i,t}^p$, rather than the observed gross wages, $w_{i,t}^p$, into the tax calculators and obtain counterfactual tax rates per percentile, $\tilde{\tau}_{i,t}^p$. The effect of changes in trade openness on the wage distribution is thus netted out of the new effective tax rates and we can produce relative counterfactual labor income tax burdens to estimate the following regression:

$$(15) \quad 100 \times \frac{\tilde{\tau}_{i,t}^p \tilde{w}_{i,t}^p}{\sum_k \tilde{\tau}_{i,t}^k \tilde{w}_{i,t}^k} = \gamma^{p,\tilde{\tau}} \ln(\pi_{i,t}) + Z_{i,t}\Gamma^{p,\tilde{\tau}} + \lambda_t^{p,\tilde{\tau}} + \mu_i^{p,\tilde{\tau}} + u_{i,t}^{p,\tilde{\tau}}.$$

As before, we estimate (15) in three ways, OLS and the two alternative IV models.[30] First, we present OLS results for the counterfactual relative tax burdens in Figure 8. They point to a similar effects pattern as for the original tax burdens in Section I: in 1980–1993, an increase in trade openness raised the progressivity of labor income tax rates, whereas in 1994–2007 this was only the case for medium-high incomes. During this last period, globalization resulted in regressive taxes for very high wages. Under the assumption of the mixed distribution of earnings, the OLS coefficients for the seventy-fifth, ninetieth, and one hundredth percentiles

---

[29] The results for the seventy-fifth, ninetieth, and one hundredth percentiles are presented in Table OA.6 in the online Appendix.

[30] Since γ^c is specification-specific in (12), we produce specification-specific counterfactual relative tax burdens in (14), which are subsequently tested under the applicable specification in (15).

Panel A. OECD: 1980–1993 OLS
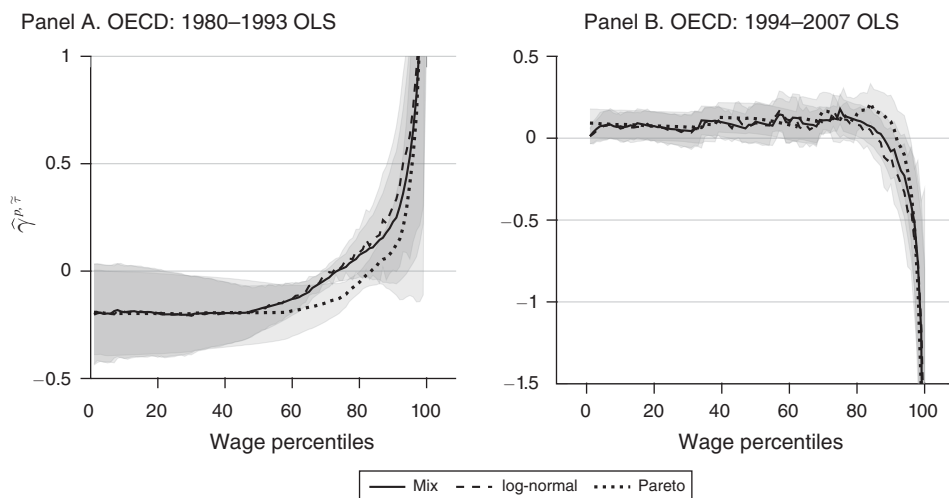
Panel B. OECD: 1994–2007 OLS

Figure 8. OLS Coefficients: Trade Openness and Percentile-Specific Counterfactual Relative Tax Burdens

are 0.01 (0.03), 0.23 (0.17), and 4.48 (1.92) for 1980–1993 and 0.18 (0.05), −0.08 (0.10), and −3.48 (1.65) for 1994–2007.

As before, the IV coefficients using the two alternative instruments are larger in absolute value than the OLS estimates, resulting in coefficients for the seventy-fifth, ninetieth, and one hundredth percentiles of −0.01 (0.09), 1.07 (0.65), and 19.72 (11.89) in the first half and −0.10 (0.10), −0.58 (0.52), and −5.72 (2.06) for the second half, using Instrument II. However, qualitatively the results remain very similar. We report IV results in Figure 9, which consists of four panels: the results for the pre-liberalization period are in the upper two panels and those for the post-liberalization period are in the lower two panels. The results are similar to those in Figure 4.

We also considered more conservative alternatives to isolate the effect by keeping the wage distribution (i) constant at the 1980 level and (ii) constant between 1980 and 1993 at its 1993 level and constant between 1994 and 2007 at its 1994 level. The results for the OLS estimates of these two procedures are reported in the online Appendix. These results and the ones in Figures 8 and 9 suggest that the main driver behind the changes in the effect of globalization on the distribution of tax burdens between 1980–1993 and 1994–2007 is indeed the response in labor income tax calculators (rather than changes in gross wages). We can thus conclude that the main force behind the results in Section I are changes in the tax calculators rather than ones in gross wages or their distribution. Furthermore, we can conclude that governments and tax authorities in the OECD were largely responsible for the increases in after-tax inequality in response to greater openness since the mid-1990s.

D. *Migration Openness as a Measure of Globalization*

The results above were based on trade openness as a measure of globalization. The literature has frequently documented that other forms of globalization,
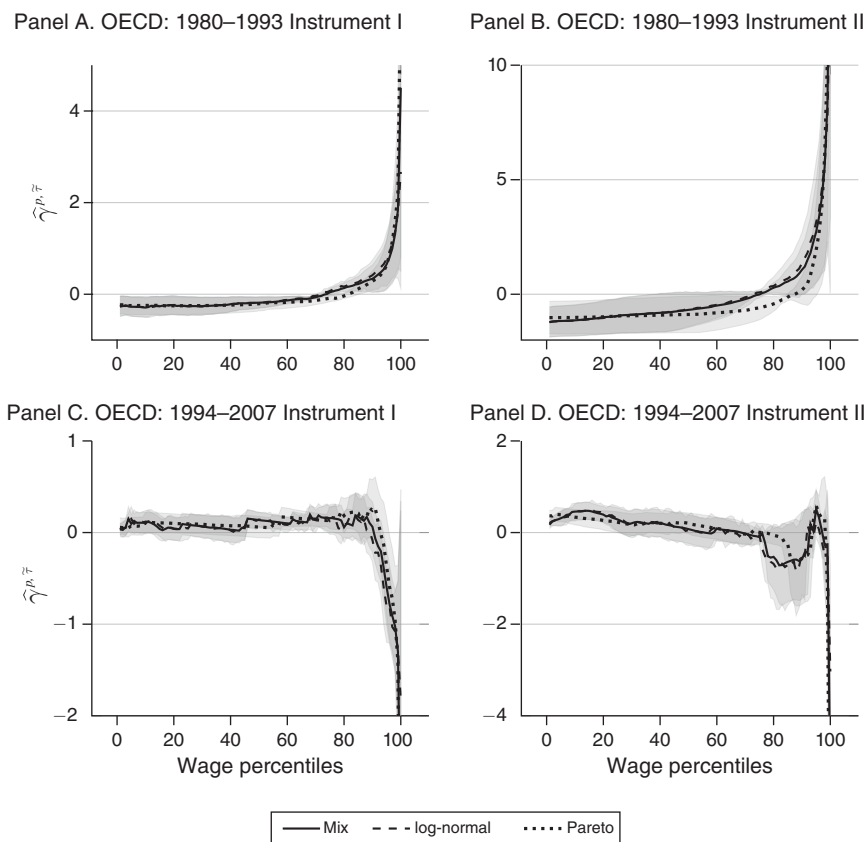
FIGURE 9. IV COEFFICIENTS: TRADE OPENNESS AND PERCENTILE-SPECIFIC COUNTERFACTUAL
RELATIVE TAX BURDENS

such as migration and foreign direct investment, are highly correlated with trade. However, data on these alternative forms of globalization are much scarcer and the relative magnitude of those activities in terms of flows is small in comparison to trade. Nevertheless, we are able to provide some evidence on the nexus between migration and labor income taxation in this subsection.

There are specific shortcomings of migration data. First, annual data on bilateral cross-border migration are generally sparse, if not missing, for a large number of countries and country pairs—for migrant flows more so than for migrant stocks. In pursuit of our goal, we compile as much information on bilateral migrant stocks and flows as possible across different sources and combine, and to a certain extent must impute, the respective data (see the Appendix for details). This leads to a higher degree of imprecision in the estimation. Second, the measure of migration encompasses all types of migration (high- and low-skilled, politically- and economically-driven, etc.) which may mask the key, purely economic relationship of interest. For these reasons, we view the results based on trade openness as more precise and reliable and consider those based on migration openness as secondary.

Let us use $\pi_{i,t}^{mig}$ to denote migration openness and define it as the average immigrant/emigrant stock relative to the native population. With migration openness, we

TABLE 5—FIRST-STAGE KLEIBERGEN-PAAP WALD F-STATISTIC MIGRATION

	1980–1993	1994–2007
Instrument migration	17.04	8.81

Note: The reported statistics are robust Kleibergen-Paap Wald rk F-statistics adjusted for country and year clustering.

obtain a migration-cost-reflecting instrument inspired by recent general equilibrium models of migration:[31]

$$(16) \qquad \lambda_{i,t}^{mig} = \sum_{j \neq i} \frac{\pi_{ij,t}^{mig}}{\pi_{ii,t}^{mig}} \frac{\pi_{ji,t}^{mig}}{\pi_{jj,t}^{mig}},$$

where $\pi_{ij,t}^{mig}$ is the observed share of natives from j that have chosen to migrate to i. This instrument, $\lambda_{i,t}^{mig}$, is the direct analogue to $\lambda_{i,t}^{I}$ for trade openness; however, there is no direct migration-based analogue to the transportation costs of Instrument II. We report the first-stage results for this instrument in Table 5.

As with trade openness, the instrument is of high relevance, reflected in the associated partial F-statistic. However, the instrument is generally weaker in the second period.[32]

To demonstrate that our main results are robust to using migration instead of trade openness, we run the following regression:

$$(17) \quad 100 \times \frac{\tau_{i,t}^{p} w_{i,t}^{p}}{\sum_{k=1}^{100} \tau_{i,t}^{k} w_{i,t}^{k}} = \gamma^{p,mig} \ln(\pi_{i,t}^{mig}) + Z_{i,t} \Gamma^{p,mig} + \eta_{t}^{p,mig} + \mu_{i}^{p,mig} + u_{i,t}^{p,mig},$$

for $p = \{1, \ldots, 100\}$. As before, we report OLS and 2SLS results for the three income distributions, plotting $\hat{\gamma}^{p,mig}$ against wage percentiles in Figure 10.

The results in Figure 10 are qualitatively consistent with those in Section I. Between 1980 and 1993, migration openness resulted in increasing tax burdens for the earners of higher incomes, whereas this was reversed between 1994 and 2007. This result holds for both OLS and 2SLS estimations. However, the results are quantitatively smaller than for trade openness, which is plausible since migration is measured in stocks rather than flows over time.

E. *Interstate Migration and Changes in Effective Labor Income Taxation in the United States*

The results above indicate that globalization affects the net wage inequality through changes in labor income tax systems (apart from its effects on gross wages). We found that governments in developed countries reduced relative labor taxes for high-income earners in response to globalization, while increasing them for earners around the center of the wage distribution to retain labor tax revenues. In this

---

[31] See Artuç, Chaudhuri, and McLaren (2010); Anderson (2011); and Dix-Carneiro (2014).

[32] As both $\pi_{ij,t}^{mig}$ and $\pi_{ij,t}$ as well as $\lambda_{i,t}^{mig}$ and $\lambda_{i,t}^{I}$ are highly correlated, it is not possible to include both trade and migration openness in the estimating equation and discern their impact at sufficiently high precision.
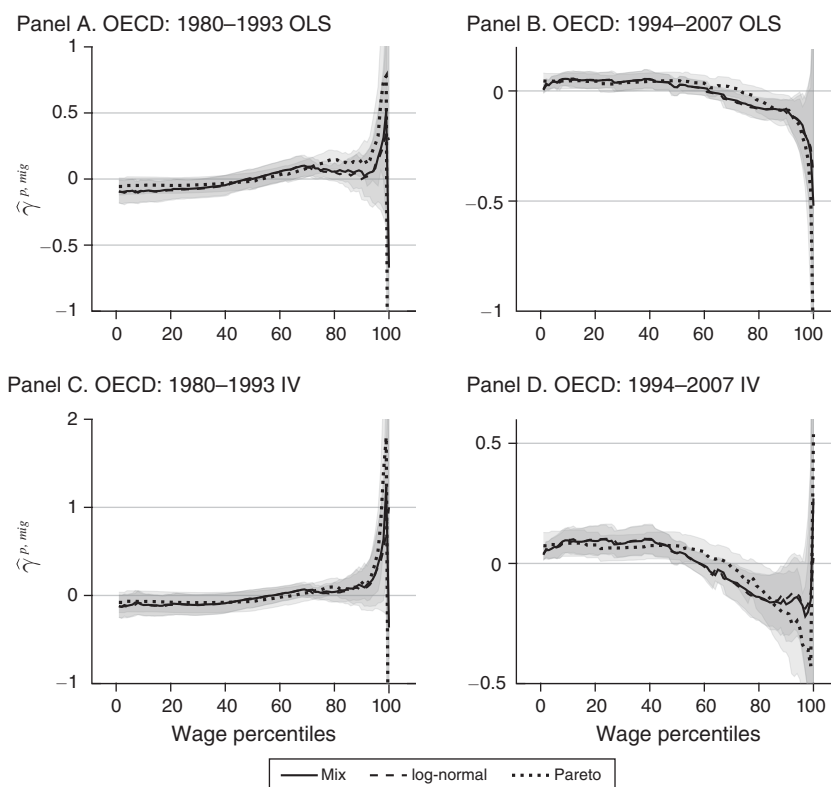
FIGURE 10. REGRESSION COEFFICIENTS: MIGRATION OPENNESS AND PERCENTILE-SPECIFIC RELATIVE TAX BURDENS

subsection, we examine whether similar mechanisms can be observed at a subnational level, using the example of the United States.

US microdata allow us to gauge the mobility and wages of workers across state borders, a level of detail that is unmatched at the international scale for many countries.[33]

The US tax system provides an interesting case, since goods trade is largely and migration is completely free within the country. Moreover, various levels of government may levy their own labor income taxes on individuals (and firms) within their respective jurisdictions: at the federal, state, and even local, substate levels. We use this variation across state-level income taxes to examine whether the mobility of high-skilled workers drives subnational tax competition.

For the corresponding analysis, we use the following data. First, we employ microdata from the Annual Social and Economic (ASEC) Supplement to the Current Population Survey (CPS) provided by the Integrated Public Use Microdata Series (IPUMS). These data include the taxable income of each individual surveyed, as well as their federal and state-level tax payments in a year. Since some of the state

---

[33] Internationally, the wealth of trade data is much greater than that of migration data. Subnationally, the opposite is true. US interstate trade is assessed (indeed, estimated) differently from international trade and only surveyed on a decennial basis. We thus limit our analysis to worker mobility rather than trade in this subsection.

samples are relatively small, we increase the granularity of the quantiles in this subsection to avoid an excessive imputation of the data. Specifically, we generate 50 taxable income quantiles and calculate their relative income tax burden at the state and federal level in each year. Second, we complement these data with ones on state-to-state migration flows provided by the Internal Revenue Service (IRS). The IRS annually records the location of each tax filer,[34] which we then use to compute overall state-to-state flows. We use the relative tax burden for single individuals with positive income (to match the international archetype as closely as possible) and measure migration flows in terms of the number of tax returns. We generate the instrument for interstate migration in the same way as in Section IIID, using normalized state-to-state flows to compute relative mobility costs for a state s in year t. Using these data, we cover all 50 states for 8 years (2000–2007), yielding 400 observations in at least a part of the later period used in Section I.

We start our analysis with the impact of mobility on the relative tax burden across income quantiles:

$$(18) \quad 100 \times \frac{\tau_{s,t}^q w_{s,t}^q}{\sum_{k=1}^{50} \tau_{s,t}^k w_{s,t}^k} = \gamma^{q,mig} \ln(\pi_{s,t}^{mig}) + Z_{st} \Gamma^{q,mig} + \lambda_t^{q,mig} + \mu_s^{q,mig} + u_{s,t}^{q,mig},$$

for $q = \{1, \ldots, 50\}$, where we include state-level controls in vector Z_{st}. We report the data sources in the Appendix. There are minor differences to the approach in Section I. In addition to only using fifty rather than hundred quantiles with each q capturing two percentiles of state s's population, we cannot differentiate between taxes on labor versus other forms of income, $w_{s,t}^p$ therefore denotes total taxable income and $\tau_{s,t}^q$ denotes the total income taxes paid on that income in this subsection. Finally, due to the anonymity requirement, IPUMS caps reported income(s), such that we likely underestimate taxable income for the highest quantile(s) in some states/years, leading us to also underestimate the effect of migration on tax outcomes for those top quantiles.[35]

We report the results for the relative state-level income tax burden in panel A of Figure 11, where for ease of exposition we again plot the estimated coefficients $\hat{\gamma}^{q,mig}$ against quantiles q with 90 percent-confidence bands. In qualitative terms, the shape of the estimated response locus is strikingly similar to Figure 4 between 1994–2007. The middle and upper-middle classes experienced an increase in their relative state-level income tax burden in response to higher interstate mobility, whereas earners in the top six percentiles experienced significant reductions in their relative state-level tax burden. The estimated coefficients suggest that a one-percent increase in interstate mobility (into and out of) a state led to a decline of 0.15 percentage points in

---

[34] Specifically, the IRS provides the total taxable income, the total number of returns, and the total number of exemptions filed within states by the state in which they were filed in the prior year. To account for the international component of migration, we include one additional residual destination and origin for in- and outward migration from and to abroad, which is provided by the IRS.

[35] For some states/years, the state-year samples are too small to produce the top quantile(s). In those cases, we linearly interpolate and extrapolate the missing observations from adjacent quantiles, similar to the imputation applied to the missing observations in the LIS data.
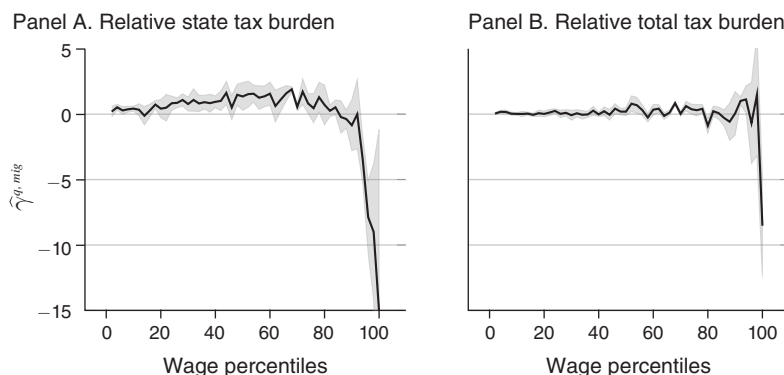
FIGURE 11. IV COEFFICIENTS: STATE-TO-STATE MIGRATION OPENNESS AND QUANTILE-SPECIFIC RELATIVE TAX BURDENS

the relative state-level tax burden of the top-income percentile. This is in line with our findings on top incomes using cross-country data.

Although state governments set their own state-level income taxes, each individual is also subject to federal income taxes, irrespective of their resident state. Hence, the US federal government superimposes an additional income tax layer that cannot be avoided through interstate mobility. The impact of interstate mobility on the overall (state-plus-federal) income tax burden should thus be less pronounced relative to the singular state-level component. To assess this hypothesis, we calculate the total relative tax burden combining the state and federal income tax for each income quantile in the distribution and rerun the regressions in (18). The results are reported in panel B of Figure 11 and suggest that higher interstate mobility indeed has no significant impact on the overall (state-plus-federal) income tax burden apart from the two highest income percentiles. The coefficients are largely insignificant and become negative only at the very right tail of the income distribution, which suggests that the federal tax layer operates as a tax policy coordination device in income taxation, capable of mitigating the adverse effect of increasing interstate mobility on state-level income tax policy for most of the US income distribution.

IV. Conclusions

The labor income tax system is widely considered to be one of the instruments capable of mitigating the inequality-fostering effects of, inter alia, greater openness across countries. This paper provides evidence that this is indeed how income tax systems were used in an average OECD country up until the mid-1990s but not so afterward. In fact, we demonstrate that the average OECD government raised the labor income tax burden of earners in the upper-middle of the gross wage distribution while easing the tax burden significantly for very-high-income earners between 1994 and 2007. This effect does not result from cold progression, i.e., inflation mechanically moving incomes into higher tax brackets of progressive tax systems, but from actual changes in tax calculators. The associated change could be justified by the higher mobility among high-income earners relative to low- and middle-income earners in the later years of our sample. Interestingly, this phenomenon is

more prevalent among developed countries (i.e., the OECD) than among less developed ones.

While we establish most of the results using trade openness (which is relatively precisely recorded) as a measure of globalization, we find a similar pattern for migration (for which data are much scarcer) using cross-national migration stocks. The results are established for imputed country-year-percentile-specific tax burdens and extensive information on respective country-year-specific labor income tax calculators. The findings are confirmed using a much smaller sample of representative microdata on wage incomes with fewer countries and years. A similar pattern is found for interstate migration within the United States, where greater cross-state mobility appears to induce heavier taxation of the middle class relative to high-income earners. A key insight from this analysis is that these effects pertain to the behavioral responses of tax authorities rather than to the indirect effects of globalization on the level of dispersion of pre-tax incomes.

Estimating these novel effects of globalization on labor income tax schedules and calculators was at the heart of the present paper. Accordingly, we resorted to descriptive analysis and reduced-form empirical work. In the future, it will be interesting to consider quantitative effects on a wider range of outcomes along the lines of structural models.

Appendix A. List of Countries and International Data Sources

A. *List of Countries*

OECD: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Great Britain, Greece, Hungary, Iceland, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Portugal, Spain, South Korea, Sweden, Switzerland, Turkey, and the United States.

Non-OECD: Argentina, Bangladesh, Barbados, Bolivia, Cameroon, Chile, China, Colombia, Costa Rica, Cyprus, Ecuador, Fiji, Ghana, Guatemala, Honduras, India, Indonesia, Israel, Jamaica, Jordan, Kenya, Kuwait, Malaysia, Malta, Mauritius, Morocco, Nepal, Pakistan, Peru, Philippines, Senegal, Singapore, South Africa, Sri Lanka, Thailand, Trinidad, Tunisia, Uruguay, and Venezuela.

B. *Data Sources*

In this section, we describe the data sources of the variables used in the analysis. We can categorize them in three main groups: (i) tax burdens, (ii) tax revenues, and (iii) control variables. We defer a discussion of the imputation of percentile-specific wages to Appendix B.

(*i*) *Tax Burdens.*—We calculate the final tax rate for each income percentile by aggregating all labor taxes, all employee-borne social security contributions, and all other taxes payable on wage income by employees and subtracting all relevant deductions and credits for a percentile's income level. The final tax must include social security, as the decision to alter the structure of the income tax schedule is

co-determined with the structure of the social security contribution schedule. The data on income tax and social security codes were collected by the authors and are described in greater detail in Egger and Strecker (2018).

(*ii*) *Tax Revenues.*—The components of interest are total tax revenue, $R_{i,t}^{total}$, as well as revenue from firm-borne, $R_{i,t}^{firm}$, employee-borne, $R_{i,t}^{employee}$, goods- and services-based, $R_{i,t}^{goods}$, and all other taxes, $R_{i,t}^{other}$. We combine data from the IMF's Government Finance Statistics and the OECD's Tax Statistics to collect the respective variables. Since the two databases have different definitions of *goods*-based taxes, we use the IMF's definition (taxes levied on the production, extraction, sale, transfer, leasing, and delivery of goods and services) and harmonize the OECD data by excluding taxes on imports, exports, and cross-border transactions. Further, for some country/year combinations disaggregated data on employer- and employee-borne social security contributions are not available and the respective shares must be estimated based on country-specific tax codes in Egger and Strecker (2018).

(*iii*) *Endogenous Explanatory Variables.*—In our analysis, we also employ a vector of endogenous explanatory variables:

(*a*) *Trade:* As previously stated, we associate openness with average exports and imports in total consumption of manufacturing goods, which we calculate using the aggregate volume of manufacturing trade flows and total manufacturing production. These data are taken from the World Bank's World Integrated Trade Solution Database. We classify manufacturing goods according to the SITC 1 classification. The domestic sale shares are calculated using data on manufacturing production from the OECD's Structural Analysis Database and the United Nations Industrial Development Organization's Industrial Statistics Database (when available). Otherwise, we predict production log-linearly using manufacturing value added.

(*b*) *Migration*: To compute the measure of migration openness we combine several datasets starting with the World Bank's Global Bilateral Migration Database available for 1980, 1990, and 2000, complemented with data from Adserà and Pytliková (2015). Next, we use data on migration flows (when available) from several sources (EUROSTAT, United Nations Global Migration Database, and International Labor Organization) to compute annual migration stocks. When flow data were unavailable we assumed zero flows and held the migration stock constant relative to the previous year.

(*iv*) *Exogenous Control Variables*.—We also employ a vector of exogenous control variables:

(*a*) *Primary*, *Secondary*, *and Tertiary Education*: The shares of the population with primary, secondary, and tertiary education are based on Barro and Lee's (2010) data on educational attainment between the years 1970 and 2000, available in 5-year intervals. Intermittent observations were interpolated via regression on a polynomial of the year variable.

(*b*) *Democracy*, *Left-Wing*, *and Right-Wing Legislative Majorities*: We include a democracy index as a binary indicator (as opposed to autocratic regimes), as well as binary indicators for left- and right-wing majorities in the legislature (center being the excluded variable) from the *Quality of Government Basic Dataset*.

(*c*) *Population*, *Real GDP per Capita*, *and Interaction Term*: We control for country size, level of development, and their interaction by including these three normalized measures in logs obtained from the World Bank's World Development Indicators database.

(*v*) *Components of the Oil Price Instrument.*—In order to generate Instrument II, we rely on the following data:

(*a*) *Oil Prices:* As stated in Section IB, Instrument II exploits the variation across time and space in crude oil prices. To that end, we use data on energy prices and taxes from the International Energy Agency, which collects information on the value and volume of crude oil as recorded by customs administration at the time of import in each importer country. The average annual price is then derived as the ratio between the two recorded measures. The recorded values include insurance and freight but not import duties.[36] We convert these prices using PPP-adjusted exchange rates from the Penn World Tables.

(*b*) *Distances:* As stated in Section IB, d_i^k denotes the (log) road distance between the three largest cities in country *i* in year 2007 and their nearest port. Distances were collected from SeaRates, a web platform that provides information on the logistics of international container shipping. The largest cities were defined following the OECD's metropolitan area statistics (see OECD 2012).

APPENDIX B. IMPUTATION OF WAGES AND COMPARISON TO AVAILABLE DATA

The analysis requires percentile-specific measures of nominal gross labor income for a panel of countries. Unfortunately, microdata sources only cover a handful of countries for a very small number of years. To overcome these limitations, we parameterize wage income distributions using moments observed in the data and produce country-year-percentile-specific gross wages. We employ three alternative assumptions about the underlying income distribution: Pareto, log-normal, and the 5 percent mixture of the two. To calibrate the parameters of the distributions, we use information on two moments: the gross wage Gini coefficient, $Gini_{i,t}$, and the average gross wage, $\bar{w}_{i,t}$.[37] Pareto and log-normal distributions

---

[36] Whenever oil price data were missing for country/year we replaced the values with the corresponding regional averages provided by the EIA.

[37] That household incomes and wealth follow Pareto-type power laws was the very insight of Pareto (1896). Recent evidence by, among others, Felbermayr, Hauptmann, and Schmerer (2014) and Egger, Egger, and Kreickemeier (2013) for worker data supports this fact.

require two parameters to be precisely identified. For the mixed log-normal-Pareto distribution, we assume different values for the parameter that divides the two tails and calibrate the remaining two (we generally use a mixture of 95 percent log-normal and 5 percent Pareto). We use the following cumulative distribution functions (CDF) for the Pareto and the log-normal distributions:

$$(19) \qquad F_{i,t}(w) = 1 - \underline{w}_{i,t}^{\phi_{i,t}} w^{-\phi_{i,t}} \quad \text{and} \quad F_{i,t}(w) = \Phi\left(\frac{\ln w - \mu_{i,t}}{\sigma_{i,t}}\right).$$

We calibrate $\underline{w}_{i,t}$ and $\phi_{i,t}$ for the Pareto and $\mu_{i,t}$ and $\sigma_{i,t}$ for the log-normal distribution. The mixed distribution (see Nigai 2017) has the following CDF:

$$(20) \quad F_{i,t}(w)$$

$$= \begin{cases} \dfrac{\rho_{i,t}}{\Phi(\alpha_{i,t} s(\alpha_{i,t}, \rho_{i,t}))} \Phi\left(\alpha_{i,t} s(\alpha_{i,t}, \rho_{i,t}) + \dfrac{\ln w - \ln \theta_{i,t}}{s(\alpha_{i,t}, \rho_{i,t})}\right) & \text{for } w \in (0, \theta_{i,t}] \\[3mm] 1 - (1 - \rho_{i,t}) \dfrac{\theta_{i,t}^{\alpha_{i,t}}}{w^{\alpha_{i,t}}} & \text{for } w \in [\theta_{i,t}, \infty) \end{cases},$$

where $\Phi(\alpha_{i,t} s(\alpha_{i,t}, \rho_{i,t})) \sqrt{2\pi} \, (\alpha_{i,t} s(\alpha_{i,t}, \rho_{i,t})) e^{\frac{1}{2}[\alpha_{i,t} s(\alpha_{i,t}, \rho_{i,t})]^2} = \rho_{i,t}(1 - \rho_{i,t})^{-1}$. We fix $\rho_{i,t}$, the share of the left (log-normal) tail, and calibrate $\theta_{i,t}$ and $\alpha_{i,t}$. The following two identities relate the CDF to average wages, $\bar{w}_{i,t}$, and to the Gini coefficients:

$$(21) \quad \bar{w}_{i,t} = \int_{\Omega_w} w \, dF_{i,t}(w) \quad \text{and} \quad Gini_{i,t} = (\bar{w}_{i,t})^{-1} \int_{\Omega_w} F_{i,t}(w)(1 - F_{i,t}(w)) \, dw.$$

For the mixture distribution, we solve the equations numerically. Once calibrated, we calculate an average income within each of the hundred percentiles per country and year.

Average labor income levels as well as gross wage Gini coefficients are obtained/calculated from the International Labor Organization's ILOSTAT database. For the OECD sample between 1980–2007, we obtain 180 Gini observations from ILOSTAT directly, which cover parts of the OECD in different years. Where specific years or countries were not available, we imputed missing gross wage Gini coefficients via linear regressions using gross *income* Gini coefficients from the Standardized World Income Inequality Database (Solt 2014), the average wage, the distribution of education levels in the population, and total capital stock as predictors. We base these regressions on 300 observations (OECD and non-OECD countries combined) to predict the rest. The R^2 of this regression is 0.71.

To ensure that our results are robust to using only direct observations of gross wage Gini coefficients, we re-run the benchmark regression on tax burden using only those specific observations of OECD countries. The sample is insufficient to estimate the effect in the earlier subperiod; however, the results of the OLS and two IV regressions for the second half of our sample are in Figure 12. The results are nearly identical to those in Section I. We consider this further evidence that our imputation procedure for Gini coefficients is accurate and that, conversely, the results in Section I are not driven by the imputation procedure for the Gini coefficient.
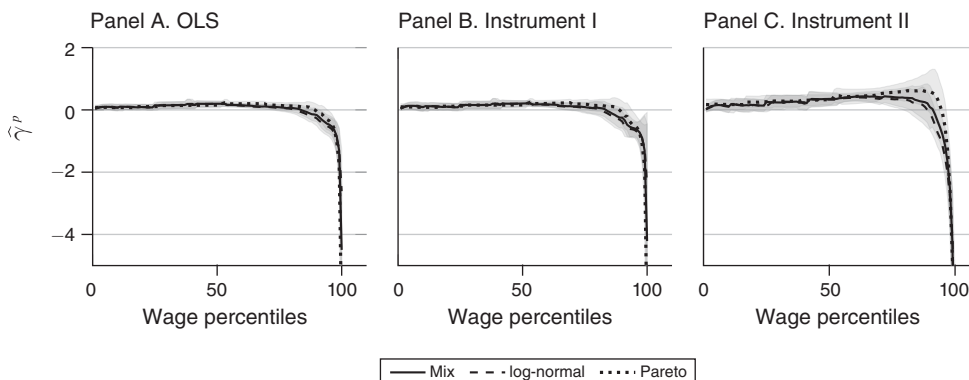
FIGURE 12. REGRESSION COEFFICIENTS: TRADE OPENNESS AND PERCENTILE-SPECIFIC RELATIVE TAX BURDEN
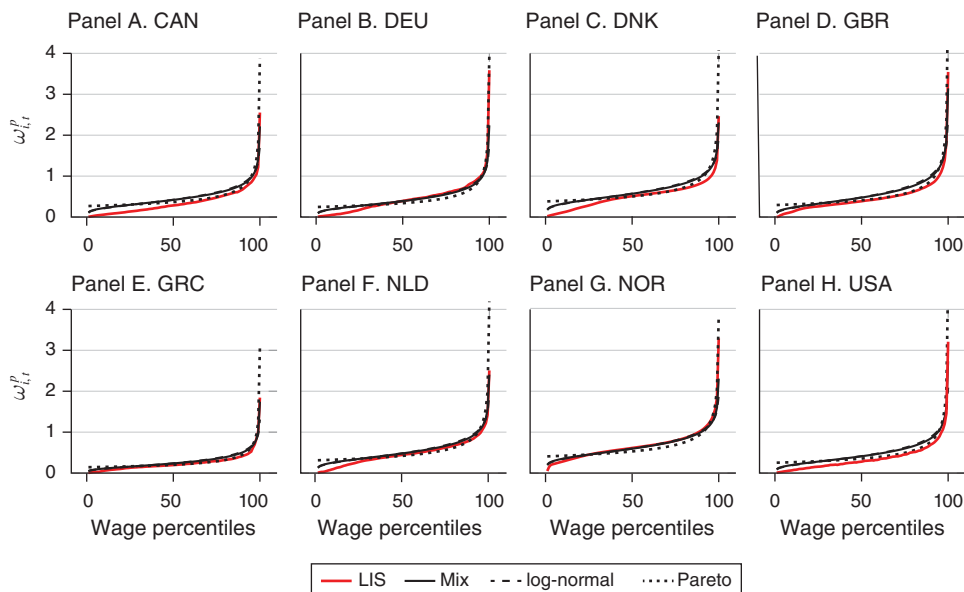(NON-IMPUTED)

FIGURE 13. LIS AND IMPUTED INCOME PERCENTILES: 2007

To check the soundness of the imputed wage distributions based on Gini coefficients and average wages, we compare the predictions against the percentile-specific wages underlying the results in Section IIIA for a selection of countries in 2007 in Figure 13 (in hundred thousands). It is apparent that the imputed percentiles match the data very well.[38]

---

[38] Additionally, we show that the imputed wages produce tax revenues that align well with the data on total labor income tax revenues in the online Appendix.

APPENDIX C: UNITED STATES CASE STUDY DATA SOURCES

(*i*) *Endogenous Explanatory Variable*.—In our US case study, we employ the following endogenous explanatory variable:

(*a*) *Migration*: We use the number of migrated tax returns (inflows and outflows) from the *IRS's US Population Migration Data* for the years 2000–2007.

(*ii*) *Exogenous Control Variables*.—We also use a vector of control variables that closely correspond to those in the international benchmark regression:

(*a*) *Primary*, *Secondary*, *and Tertiary Education*: The shares of the population with primary, secondary, and tertiary education are based on the *US Census Current Population Survey*'s data on educational attainment between 1990 and 2010, where missing observations were interpolated via regression on a polynomial of the year.

(*b*) *Left-Wing and Right-Wing Legislative Majorities*: We include concurrent, binary indicators for left- and right-wing majorities in state legislatures from the *National Conference on State Legislatures*.

(*c*) *Population*, *Real GDP per Capita*, *and Interaction Term*: We control for state size, level of development, and their interaction by including these three normalized measures from the *Bureau of Economic Analysis* database on *Regional Economic Accounts*.

REFERENCES

Adserà, Alícia, and Mariola Pytliková. 2015. "The Role of Language in Shaping International Migration." *Economic Journal* 125 (586): F49–81.

Alesina, Alberto, and Romain Wacziarg. 1998. "Openness, Country Size and Government." *Journal of Public Economics* 69 (3): 305–21.

Amiti, Mary, and Donald R. Davis. 2012. "Trade, Firms, and Wages: Theory and Evidence." *Review of Economic Studies* 79 (1): 1–36.

Anderson, James E. 2011. "The Gravity Model." *Annual Review of Economics* 3: 133–60.

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

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

Artuç, Erhan, Shubham Chaudhuri, and John McLaren. 2010. "Trade Shocks and Labor Adjustment: A Structural Empirical Approach." *American Economic Review* 100 (3): 1008–45.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103 (6): 2121–68.

Autor, David H., David Dorn, and Gordon H. Hanson. 2015. "Untangling Trade and Technology: Evidence from Local Labour Markets." *Economic Journal* 125 (584): 621–46.

Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song. 2014. "Trade Adjustment: Worker-Level Evidence." *Quarterly Journal of Economics* 129 (4): 1799–1860.

Baier, Scott L. and Jeffrey H. Bergstrand. 2001. "The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity." *Journal of International Economics* 53 (1): 1–27.

Baldwin, Richard E. 2006. "Multilateralising Regionalism: Spaghetti Bowls as Building Blocs on the Path to Global Free Trade." *World Economy* 29 (11): 1451–1518.

Barro, Robert J., and Jong Wha Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics* 104: 184–98.

Bureau of Economic Analysis. Various years. Regional Economic Accounts (accessed September 21, 2015).

Burstein, Ariel, and Jonathan Vogel. 2012. "International Trade, Technology and the Skill Premium." Unpublished.

Caliendo, Lorenzo, and Fernando Parro. 2014. "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies* 82 (1): 1–44.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2006. "Robust Inference with Multi-Way Clustering." National Bureau of Economic Research Working Paper 327.

Cameron, David R. 1978. "The Expansion of the Public Economy: A Comparative Analysis." *American Political Science Review* 72 (4): 1243–61.

Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review* 98 (4): 1707–21.

Dahlberg, Stefan, Sören Holmberg, Bo Rothstein, Felix Hartmann, and Richard Svensson. 2015. The Quality of Government Basic Dataset, Version Jan15. University of Gothenburg: The Quality of Government Institute. (Accessed April 2, 2015.)

Devereux, Michael P., Rachel Griffith, Alexander Klemm, Marcel Thum, and Marco Ottaviani. 2002. "Corporate Income Tax Reforms and International Tax Competition." *Economic Policy* 17 (35): 449–95.

Devereux, Michael P., Ben Lockwood, and Michaela Redoano. 2008. "Do Countries Compete over Corporate Tax Rates?" *Journal of Public Economics* 92 (5–6): 1210–35.

Diamond, Peter A. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review* 88 (1): 83–95.

Dix-Carneiro, Rafael. 2014. "Trade Liberalization and Labor Market Dynamics." *Econometrica* 82 (3): 825–85.

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

Egger, Hartmut, Peter H. Egger, and Udo Kreickemeier. 2013. "Trade, Wages, and Profits." *European Economic Review* 64: 332–50.

Egger, Hartmut, and Udo Kreickemeier. 2009. "Firm Heterogeneity and the Labor Market Effects of Trade Liberalization." *International Economic Review* 50 (1): 187–216.

Egger, Peter H., and Sergey Nigai. 2015. "Structural Gravity with Dummies Only: Constrained ANOVA-Type Estimation of Gravity Panel Data Models." *Journal of International Economics* 97 (1): 86–99.

Egger, Peter H., Sergey Nigai, and Nora M. Strecker. 2019. "The Taxing Deed of Globalization: Dataset." *American Economic Review.* https://doi.org/10.1257/aer.20160600.

Egger, Peter H., and Nora M. Strecker. 2018. "A Tour of Labor Income Tax in the World, 1980–2012." Unpublished.

Egger, Peter H., and Filip Tarlea. 2017. "Comparing Apples to Apples: Estimating Consistent Partial Effects of Preferential Economic Integration Agreements." CEPR Discussion Paper 11894.

Epifani, Paolo, and Gino Gancia. 2009. "Openness, Government Size and the Terms of Trade." *Review of Economic Studies* 76 (2): 629–88.

EUROSTAT. 2015. Migration and Citizenship Data. (Accessed September 16, 2015.)

Feenstra, Robert C., and Gordon H. Hanson. 1996. "Globalization, Outsourcing, and Wages Inequality." *American Economic Review* 86 (2): 240–45.

Feenstra, Robert C., and Gordon H. Hanson. 1999. "The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979–1990." *Quarterly Journal of Economics* 114 (3): 907–40.

Feenstra, Robert C., and Gordon H. Hanson. 2004. "Global Production Sharing and Rising Inequality: A Survey of Trade and Wages." In *Handbook of International Trade,* Vol. 1, edited by E. Kwan Choi and James C. Harrigan, 146–85. New York: Wiley.

Felbermayr, Gabriel, Andreas Hauptmann, and Hans-Jörg Schmerer. 2014. "International Trade and Collective Bargaining Outcomes: Evidence from German Employer-Employee Data." *Scandinavian Journal of Economics* 116 (3): 820–37.

Flood, Sarah, Miriam King, Steven Ruggles, and J. Robert Warren. 2015. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0 [dataset]. Minneapolis: University of Minnesota (accessed November 12, 2015).

Ganghof, Steffen, and Richard Eccleston. 2004. "Globalisation and the Dilemmas of Income Taxation in Australia." *Australian Journal of Political Science* 39 (4): 519–34.

Goldberg, Pinelopi K., and Nina Pavcnik. 2004. "Trade Inequality, and Poverty: What Do We Know? Evidence from Recent Trade Liberalization Episodes in Developing Countries." National Bureau of Economic Research Working Paper 10593.

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

Harrigan, James, and Ariell Reshef. 2015. "Skill-Biased Heterogeneous Firms, Trade Liberalization, and the Skill Premium." *Canadian Journal of Economics* 48 (3): 1024–66.

Head, Keith, and John Ries. 2001. "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade." *American Economic Review* 91 (4): 858–76.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding. 2010. "Inequality and Unemployment in a Global Economy." *Econometrica* 78 (4): 1239–83.

Heston, Alan, Robert Summers, and Bettina Aten. 2012. Penn World Table Version 7.1. University of Pennsylvania: Center for International Comparisons of Production, Income and Prices. November 2012.

Hines, James R., and Lawrence H. Summers. 2009. "How Globalization Affects Tax Design." National Bureau of Economic Research Working Paper 14664.

Internal Revenue Service (IRS). Various years. SOI Tax Stats: Migration Data. Web (accessed October 8, 2015).

International Energy Agency. Various years. Energy Prices and Taxes. Web (accessed October 23, 2017).

International Monetary Fund (IMF). Various years. Government Financial Statistics. Web (accessed January 13, 2015).

International Labor Organization. Various years. ILOSTAT. Web (accessed December 20, 2012).

Kirchgässner, Gebhard, and Werner W. Pommerehne. 1996. "Tax Harmonization and Tax Competition in the European Union: Lessons from Switzerland." *Journal of Public Economics* 60 (3): 351–71.

Kleven, Henrik J., Camille Landais, Emmanuel Saez, and Esben Schultz. 2014. "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark." *Quarterly Journal of Economics* 129 (1): 333–78.

Lehmann, Etienne, Laurent Simula, and Alain Trannoy. 2014. "Tax Me If You Can! Optimal Nonlinear Income Tax between Competing Governments." *Quarterly Journal of Economics* 129 (4): 1995–2030.

Luxembourg Income Study Database (LIS). Dataset. Multiple countries. Luxembourg: LIS. www.lisdatacenter.org (data run in July 2015).

Melitz, Mark J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–725.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (2): 175–208.

National Conference of State Legislatures. 2015. Partisan Composition of State Legislatures. Web (accessed September 28, 2015).

Nigai, Sergey. 2016. "On Measuring the Welfare Gains from Trade under Consumer Heterogeneity." *Economic Journal* 126 (593): 1193–237.

Nigai, Sergey. 2017. "A Tale of Two Tails: Productivity Distribution and the Gains from Trade." *Journal of International Economics* 104: 44–62.

Organisation for Economic Co-operation and Development (OECD). 2012. *Redefining "Urban": A New Way to Measure Metropolitan Areas.* Paris: OECD.

Organisation for Economic Co-operation and Development (OECD). 2014. STAN Database for Structural Analysis. Web (accessed December 22, 2014).

Organisation for Economic Co-operation and Development (OECD). 2015. Revenue Statistics. Web (accessed July 22, 2015).

Pareto, Vilfredo F. D. 1997 (1896). "La courbe de la répartition de la richesse." *Rivista di Politica Economica* 87: 647–700.

Parro, Fernando. 2014. "Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade." *American Economic Journal: Macroeconomics* 5 (2): 72–117.

Persson, Torsten, and Guido Tabellini. 1994. "Representative Democracy and Capital Taxation." *Journal of Public Economics* 55 (1): 53–70.

Piketty, Thomas, and Emmanuel Saez. 2013. "Optimal Labor Income Taxation." In *Handbook of Public Economics*, Vol. 5, edited by Alan J. Auerbach et al., 391–474. New York: Elsevier.

Rodrik, Dani. 1997. "Trade, Social Insurance, and the Limits to Globalization." National Bureau of Economic Research Working Paper 5905.

Rodrik, Dani. 1998. "Why Do More Open Economies Have Bigger Governments?" *Journal of Political Economy* 106 (5): 997–1032.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1): 205–29.

Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *Review of Economics and Statistics* 36 (4): 387–89.

Schmidheiny, Kurt. 2006. "Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland." *Journal of Public Economics* 90 (3): 429–58.

Simula, Laurent, and Alain Trannoy. 2010. "Optimal Income Tax under the Threat of Migration by Top-Income Earners." *Journal of Public Economics* 94 (1–2): 163–73.

Solt, Frederick. 2014. The Standardized World Income Inequality Database, Version 14c. Web (accessed December 22, 2014).

United Nations Conference on Trade and Development (UNCTAD). FDI/TNC Database. Web (accessed December 22, 2014).

United Nations Department of Economic and Social Affairs, Population Division. 2011. International Migration Flows to and from Selected Countries: The 2010 Revision. Web (accessed July 22, 2015).

United Nations Industrial Development Organization (UNIDO). Industrial Statistics Database. Web (accessed December 22, 2014).

United Nations Statistics Division. Commodity Trade Statistics Database. Web (accessed December 22, 2014).

United States Census. Current Population Survey Data on Educational Attainment. Web (accessed September 28, 2015).

Wagner, Don. 2000. "Do Tax Differences Cause the Brain Drain?" *Policy Options* (December): 33–41.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

World Bank. Global Bilateral Migration Database. Web (accessed September 15, 2015).

World Bank. World Development Indicators. Web (accessed December 22, 2014).

World Bank. World Integrated Trade Solution (WITS) Database. Web (accessed July 22, 2013).

Zhu, Susan C., and Daniel Trefler. 2005. "Trade and Inequality in Developing Countries: An Empirical Assessment." Society for Economic Dynamics 2004 Meeting Papers 535.