

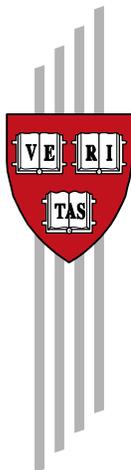
Employment Structure and the Rise of the Modern Tax System

Anders Jensen

CID Faculty Working Paper No. 371

October 2019

© Copyright 2019 Jensen, Anders; and the President and Fellows of
Harvard College



Working Papers

Center for International Development
at Harvard University

Employment Structure and the Rise of the Modern Tax System*

Anders Jensen †

October 2019

Abstract

This paper shows how the increase in information trails through the long-run transition from self-employment to employee-jobs explains the rise of the modern income tax system. I construct a new database which covers 100 household surveys across countries at different income levels and 140 years of historical data within the US (1870-2010). Using these data, I first establish four new stylized facts: 1) within country, the share of employees increases over the income distribution, and increases at all levels of income as a country develops; 2) the income tax exemption threshold moves down the income distribution as a country develops, tracking employee growth; 3) the employee share above the exemption threshold is maximized and remains constantly high; 4) decreases in the exemption threshold are strongly associated with increases in tax collection. These findings are consistent with a model where a high employee share is a necessary condition for effective taxation and where the rise in income covered by information trails through increases in employee share drives expansion of the income tax base. To provide a causal estimate of employee share on income tax systems, I study a state-led US development program implemented in the 1950s-60s which increased the level of employee share. The identification strategy exploits within-state changes in court-litigation status which generate quasi-experimental variation in the effective implementation date of the program. I find that the exogenous increase in employee share is associated with an expansion of the state income tax base and an increase in state income tax revenue.

*I thank Oriana Bandiera, Tim Besley, Henrik Kleven and Emmanuel Saez for relentless support on this project. I am grateful to David Autor, Alan Auerbach, Pierre Bachas, Michael Best, Lucie Gadenne, Francois Gerard, Roger Gordon, Rema Hanna, Michael Keen, Asim Khwaja, Wojciech Kopczuk, Camille Landais, Joana Naritomi, Paul Niehaus, Ben Olken, Torsten Persson, Jim Poterba, Dina Pomeranz, Joel Slemrod, Johannes Spinnewijn, Stefanie Stantcheva, John Van Reenen, Hans-Joachim Voth, Danny Yagan, Gabriel Zucman, as well as seminar participants at Chicago Harris, Columbia, Harvard, HKS, IIES, MIT, Northwestern, Stanford, Toulouse School of Economics, UC Berkeley, UC San Diego, UCLA, University of Michigan, University of Zurich, Warwick, Yale, and ZEW, for helpful comments.

†Harvard Kennedy School and NBER. Contact: anders_jensen@hks.harvard.edu.

1 Introduction

A central question in public finance is understanding how governments can go from raising 5% of GDP in tax collection to 45% over the development path. Recent models suggest that increased availability of information trails can lead to growth in enforcement capacity and tax collection (Gordon and Li, 2009; Kleven, Kreiner and Saez, 2016).¹ In this paper, I provide evidence that the increase in information trails through the long-run transition from self-employment to employee-employment is a determinant of income tax collection along the full development path. Using a unique micro-database, I establish new stylized facts that are consistent with this theory both across and within countries over the long-run. I then provide a direct test of the theory by studying the impacts of plausibly exogenous increases in employee-employment on US states' individually administered income tax collection. In both settings, I find evidence on a new policy mechanism: governments increase collection by expanding the income tax base to capture the growth in enforceable employee-income. The tax base has received little attention in the academic literature, but is considered an important tool to increase tax collection among practitioners (IMF, 2015).

The starting point of this paper is the construction of a new micro-database on employment structure, individual income and tax systems. It is based on nationally representative household surveys from 100 countries at all levels of development today, as well as the U.S. (1870-2010) and Mexico (1960-2010) over the long-run.² Each survey captures all work-types that cover the economically active population,³ as well as all sources of income subject to income tax. In each survey, I calculate the employee-share of income across percentiles of the individual income distribution. The employee-share is a measure of the portion of income that is covered by information trails and is enforceable.⁴ I calculate the income tax base as the share of the economically active population whose income lies above the exemption threshold and is subject to income tax.

Using the micro database, I begin by establishing four new stylized facts on employ-

¹A trail is created when information on taxable transactions is shared with a third-party. Third-party information sharing reduces the ability to mis-report activities for tax purposes, and increases enforceability.

²All surveys have been chosen using multiple criteria to make comparability across countries and over time as high as possible. These criteria led me to source most surveys in low and middle-income countries directly from national statistical agencies rather than from international public repositories.

³In addition to own-account workers, the quality of the surveys allows me to include as self-employed several other work-types that are particularly important in developing countries: family workers, casual wage-laborers, and unincorporated household enterprises.

⁴The enforceability distinction between employees and self-employed is well-known among practitioners (Bird and Oldman, 1964; OECD, 2006). It is due to the existence of information trails on employees, which deters evasion of their income taxes. The empirical evidence and theoretical foundations for the compliance success of information trails are discussed in Section 2.

ment structure and tax systems across space and time. Stylized fact #1 shows that at lowest levels of development, the employee-share is concentrated at the top of the country's income distribution; over development, the employee-share rises gradually further down the income distribution. Stylized fact #2 shows that at lowest levels of development, the exemption threshold is also located at the top of the income distribution; over development, the threshold gradually decreases, by tracking growth in employee-share. Fact #3 shows that over development, the exemption threshold is systematically located such that it maximizes the employee-share on the tax base. These facts are illustrated in Figure 1, using four countries at different levels of development today and the US over time (1870-1960).⁵ Decreases in the threshold lead to increases in the size of the tax base. Fact #4 documents a strong positive association across countries between size of tax base and level of income tax collection.⁶

These facts are consistent with a theory in which the rise in enforceable income through increases in employee-share occur gradually further down the income distribution (fact #1). In turn, the government increases the size of the base to capture the growth in enforceable income (fact #2), while ensuring that the base remains highly enforceable (fact #3). The expansion of an enforceable base leads to higher tax collection (fact #4). I show that this theory can explain the observed rise of a modern income tax system that enjoys high levels of collection from a broad base both across countries at all levels of development today and within countries over the long-run. I argue that this contrasts with pre-existing theories of income tax growth, which cannot account for both the cross-sectional and time-series variation.

In a second part of the paper, I provide a direct test of the theory by generating plausibly exogenous variation in employee-share and studying its impacts on both tax collection and the base. I focus on U.S. states, where each state determines its own income tax system and there exists rich historical variation across and within states in taxes and employment structure. I first document that the four stylized facts also hold within the average US state between 1950 and 2010. To generate variation in employment-structure, I study a previously unexploited state-led development program, Industrial Development Bonds (IDB), implemented across states in 1950s-1960s. The IDB program sought to expand industry employee-jobs to areas characterized by under-employment among

⁵I show in Section 3 that facts #1-#3 are remarkably similar *in levels* when comparing historical distributional profiles of the US to profiles of current countries at similar levels of development.

⁶Consistent with fact #4, Piketty and Qian (2009) show that China's income tax expansion since 1980 was largely achieved through increases in the base, while India's tax base and collection have stagnated. In addition, the exemption threshold is related to the extensive margin of compliance of moving workers into the tax base. As such, movements in the size of the base directly address this extensive margin elasticity which is considered an important constraint on tax collection in developing countries (Keen, 2012).

self-employed, thereby acting as a positive level-shifter in employee-share.

The identification strategy exploits the institutional features of IDB. The lack of historical precedent for the unusual IDB funding scheme led to a lag in time between the state House voting to begin the program and the state supreme court upholding the constitutional legality of the program (Abbey, 1965; Cobb, 1993).⁷ I show the upholding event was strictly necessary to effectively start the program. I therefore use changes in court litigation status as the source of identifying variation, comparing changes in outcomes before and after the upholding event to changes before and after the vote event within the same state.⁸ This specification centers the estimation around the upholding event, which I show is plausibly exogenous, while controlling for endogenous changes around the vote event in governments' willingness to start IDB. To support the identifying assumption, I graphically confirm that regression results are driven by breaks from stable pre-trends around the upholding event, and I show that only historical state-characteristics explain variation in the vote-upholding lag.

I find that IDB led to a significant increase in employee-share, as well as an expansion of the tax base and tax collection. At the same time, I find no significant impacts on a wide set of proxies for confounding determinants of the base and collection. These results directly suggest that plausibly exogenous increases in employee-share lead to growth in tax collection and the base. I further find that the growth in employee-share was concentrated in deciles of the income distribution below the pre-program threshold location, and that state governments lowered the threshold in the post-program period. These additional results are consistent with the specific mechanism in the theory, where the threshold decreases by tracking growth in employee-share further down the income distribution. Finally, a heterogeneity analysis supports the interpretation that the determining difference between employees and self-employed is enforceability.

In a final section of the paper, I model the threshold reform decision, in order to illustrate some of the channels through which enforceability differences in work-types impact the costs and benefits of base expansions. The main prediction that an increase in employee-share causes an expansion of the base is robust to different formulations of the government's objective function, and can operate through administrative, equity, and efficiency cost channels. The model suggests a central elasticity which summarizes the impacts of employee-share on the size of the base, which I calculate using the IDB estimates. Using these calculations, I find that the theory can account for 29.9% of the variation in tax systems between low-income and high-income countries.

⁷In 40% of IDB states, this intermediary period between vote and upholding was over 9 years long.

⁸The estimation is helped by the significant cross-state variation both in the year of the upholding and in the length of the lag between the vote and upholding.

These results have implications for taxation and development. First, they suggest that governments in low-income countries face significant enforceability constraints and may be practically unable, rather than politically unwilling (Keen, 2012), to significantly broaden the income tax base. In turn, the existence of a constrained narrow base violates the central assumption in the normative theory for income tax-based redistribution (Atkinson and Stiglitz, 1976). Instead, it may be desirable to achieve redistribution through indirect taxes (Bachas, Gadenne and Jensen, 2019).⁹ Second, this paper introduces the tax base as a new policy instrument in taxation and development. I find that after controlling for size of base, there is no remaining association between income tax collection and development. This implies that pre-existing theories of income tax growth, which abstract from the base, should be reconciled with this new policy mechanism.

Section 2 discusses related literature. Section 3 describes the micro database and provides stylized facts. Section 4 provides well-identified evidence. Section 5 provides a model and quantifies the theory. Section 6 discusses implications and concludes.

2 Related literature

These results contribute to theories of government growth. My paper is related to theories that suggest exogenous changes in enforceability impact government size (Hinrichs, 1966; Gordon and Li, 2009; LaPorta and Shleifer, 2014). In particular, Kleven et al. (2016) formulate a reduced-form prediction based on aggregate changes in information trails due to firm-size growth. I formulate a mechanism prediction (the tax base) based on dis-aggregate distributional changes in information trails, making my theory complementary to theirs. In addition, I provide the first descriptive micro-evidence and quasi-experimental evidence on the importance of employment-based information trails for long-run income tax collection.¹⁰ In Section 3.5, I discuss how other competing theories of income tax growth, including hypotheses based on public goods (Scheve and Stasavage, 2010), inequality (Peltzman, 1980) and institutions (Acemoglu et al., 2015), are inconsistent with the full set of stylized facts both across and within countries over time. I also discuss aspects of long-run taxation that my theory is silent on, including the role of capital income and information trails on firm sales.

The paper is related to studies of information trails and tax enforcement. The interpretation throughout this paper is that the existence of verifiable paper trails on workers'

⁹At the same time, my results emphasize the important role of governments in supporting growth of employee-employment. In the long-run, this will relax the enforceability constraints and allow governments to tax and redistribute primarily through a broad income tax.

¹⁰Kleven et al. (2016) also provides descriptive, but aggregate, evidence through a positive cross-country correlation between overall employment-share in large firms and total tax collection.

activities in firm records deters employee evasion of income taxes. Kleven et al. (2016) micro-found the enforcement success of information trails and third-party reporting of wage income. Gordon and Li (2009) theoretically investigate the implications of information reporting by financial institutions. Micro-evidence shows that compliance across taxpayers is correlated with extent of information trail coverage, including Kleven et al. (2011), Best (2014), Best et al. (2015), Pomeranz (2015), and Frias, Kumler, and Verhoegen (2015).¹¹ I contribute by carefully documenting across space and time how information trail coverage varies along an income distribution, and relate these patterns to long-run tax collection. My IDB estimates show tax collection impacts of changes in information trails on individual income. These results are complementary to Naritomi (2019), who identifies the collection impacts of changes in information trails on firms' sales.

This paper contributes to long-run studies of tax systems in U.S. states. Serrato and Zidar (2018) study trends in state corporate tax systems and their impacts on state tax collection. Dincecco and Troiano (2015) study revenue impacts of the introduction of the state individual income tax, while Troiano (2017) focuses on the revenue and audit impacts of cooperation agreements between state administrations and the federal tax administration. This paper contributes by studying determinants of state income tax base expansions and their revenue impacts. Place-based economic development programs represent a significant share of current U.S. state and local government expenditure (Kline and Moretti, 2014). The IDB was a historically important place-based program, representing 11% of state and local debt issuance at its peak in the 1960s (Rollinson, 1976). This paper provides the first evidence of the program's impacts on employment, income, and tax systems.

Finally, this paper contributes to the literature on long-run changes in employment structure. I document significant within-country income-heterogeneity in work-types and show how these distributional income-profiles vary with development. These stylized facts are complementary to previous studies, that focus on aggregate differences in work-types (Banerjee and Duflo, 2007; Gollin, 2008). In all surveys, I estimate the income-employee gradient in a robust regression framework; the magnitude and variance of these gradients are empirical moments that can help test competing theories about labor markets and development (Donovan, Lu, and Schoellman, 2019). More broadly, this paper's methodology is related to studies of macro-economic changes using newly constructed micro-evidence (Bicks, Fuchs-Schundeln and Lagakos, 2018; Donovan et al., 2019; Bachas et al., 2019).

¹¹In contexts with weak enforcement capacity, however, the compliance impacts of third-party reporting may be constrained (Carillo, Pomeranz, and Singhal, 2015). In those settings, governments can increase enforcement by requiring third-party agents to withhold taxes (Brockmeyer and Hernandez, 2018).

3 Descriptive evidence on employment structure and tax structure along the development path

I first describe the survey data underlying the cross-country and US historical analysis. Next, I outline the definition of variables and methodology. I then provide novel stylized facts on employment and tax structure along the development path.

3.1 Data

The aim of this section is to construct a distributional profile of employment structure in the individual gross income distribution, and relate distributional employment changes to changes in the size of the income tax base across countries and over time. I select surveys according to three criteria, in order to achieve as high comparability across countries and over time as possible. First, I focus on surveys which contain information on all employment-types that make up the economically active population, following the definition of employment from U.N. System of National Accounts (SNA).¹² This definition has been used in recent cross-country studies of hours worked (Bicks et al., 2018) and unemployment (Feng et al., 2019). Second, I focus on surveys which allow me to calculate a broad measure of individual gross income for the entire active population.¹³ This broad measure includes sources that are potentially included in the determination of individual income taxes: employment income, self-employment income, capital income, and miscellaneous income.¹⁴ I exclude two important sources of income: social transfers, and non-monetary income. Social transfers are excluded because they are not subject to income tax. In contrast, certain non-monetary payments in a work-setting can be subjected to income tax, but I exclude this income source simply because I cannot consistently measure it across all surveys.¹⁵ In 7 countries, I cannot calculate gross individual income with

¹²This criteria eliminates firm surveys, which are often restricted to measuring activities which generate a wage and can omit self-employed and family workers.

¹³This criteria eliminates most administrative tax data-sets, where income is often only observed for the subset of the active population whose income lies above the exemption threshold.

¹⁴Not all countries tax all these sources of income on the individual income tax. I nevertheless maintain this broad measure to ensure comparability of the income distribution across all countries. There are important concerns regarding the precision of self-reported capital income. As an alternative, I can construct the income distribution based on the narrower earned income concept, which excludes capital income and miscellaneous income. Results based on this alternative income definition are similar.

¹⁵The only exception is for respondents that are economically active, but report no monetary income - per example, family workers in agriculture that fully consume their own output. For this limited sample of respondents, in a limited number of countries, I impute an income based on answers to questions about the market-value of the own-consumed output. For more details on the importance of non-monetary income for the main results, please see Appendix A.1.

precision. In those cases, I proxy for gross income using expenditure.¹⁶ Third, I limit the sample to nationally representative data-sets.

These criteria lead me to focus on nationally representative household surveys which cover all types of employment and provide continuous measures of total income. I was able to collect such data in 100 countries with a population of at least one million. The different parts of the worldwide income distribution are almost equally represented, with 20 % from low-income countries; 28% from lower-middle income countries; 21% from upper-middle income countries; and, 31% from high-income countries.¹⁷ For 44 countries, I was able to draw on harmonized data-sets, comprised of Luxembourg Income Study (LIS, 36 countries), Economic Research Forum Micro Data Catalog (ERF; 6 countries), and International Public-Use Micro-data Project (IPUMS; 2 countries). For the remaining 56 countries, which are exclusively in the low-income and middle-income groups, I sourced the data directly from the country's national statistical agency. The database contains 93 living conditions surveys, 5 labor force surveys, and 2 censuses. When multiple years were available, I choose the year closest to 2014. Exact years, data-sources, and sample sizes for all households are given in Appendix A.1.

I construct the within country dimension of the database by combining new and previously used micro sources from the US. I collect data between 1950 and 2010 from Census micro-data extracted via IPUMS USA. Before 1950, the Census did not record work type and continuous measures of income at the individual level. I use the 1935-36 Study of Consumer Purchases, which was considered the precursor to the Census methodology of data on income at the individual level. 300,000 households were interviewed based on sampling units chosen to represent the "demographic, regional, and economic characteristics of the United States" (ICPSR, 2009). Importantly, both the work type and income categories in the 1935 survey are consistent with the later Census-based definitions. To construct a historical pre-1900 profile of employment structure, I use data resulting from a collective effort between Lindert and Williamson (2016) and IPUMS USA.¹⁸ Lindert and Williamson use local tax assessments and occupational directories for 'registered occupations' and local censuses for 'unregistered occupations' in 1870. Labor force counts using the 1 percent US Census sample were provided specifically for the data-project by IPUMS USA. I use the Lindert and Williamson computations of gross earned income, which in-

¹⁶This is mainly due to the inability to measure self employment income in the agricultural sector. See Appendix A.1 for further details. Note that in these countries, I do not attempt to convert expenditure into gross income for the purpose of locating the threshold in the income distribution.

¹⁷I calculate the income classification of a country in the year for which I have survey data using the historical World Bank Classification data.

¹⁸All the representative surveys carried out between 1900 and 1930 are not suitable for my exercise, as they focused on sampling employed wage-earners.

clude wage income, farm income and non-farm business income. However, unlike the surveys from 1935 onward which contain harmonized employee and self-employed variables, I directly build the employment categories in the 1870 using the detailed work titles from the enumerator instructions. Details on the US historical data are in Appendix A.1.

3.2 Definitions and methodology

I code employment as self-employed versus employee on the basis of whether the work-type generates derivative information trails with third-parties that can be used for income tax enforcement. This interpretation does not hinge on whether governments systematically require third-party institutions to report such information, or to act as withholding agents. Indeed, studies have found important compliance effects of information trails even when third-party reporting is not systematically in place (Kleven, 2014). Nonetheless, third-party reporting is wide-spread, and the use of withholding on income taxes, if anything, is more common at lower levels of development (Brockmeyer and Hernandez, 2018). In developing countries where third-party coverage and withholding of employee earnings is limited in scope, information trails include contractual arrangements such as written labor contracts. This information-trails definition of employees is conceptually consistent with the contractual definition of employees in Banerjee and Newman (1993). Based on this definition, I classify as self-employed: own account workers, casual wage-laborers, contributing family workers, employers, and unincorporated household enterprises.¹⁹ This classification is very similar to ILO's classification of informal labor (2002), with the exception of domestic workers in private households, which I classify as self-employed and ILO as employee. Note this classification differs from the underground economy classification (Laporta and Shleifer, 2014), where informality is based on an 'ex post' tax evasion status.

In all surveys, I code employment type based on the primary job in which the respondent spent most hours working during a reference period. I discuss bias arising from this choice of measurement in Section 3.4. I can clearly distinguish employees from employers and from both family and non-family workers in household enterprises. As such, my classification is not confounded by a distinction between firm of size 1 versus larger firms. In countries where seasonal wage-work is arguably most important, I can reasonably distinguish such workers from contract-based employees. Since casual wage-work is most prevalent at lower levels of development, any mis-classification as employee-jobs would lead me to under-state the true transition out of self-employment along the

¹⁹The U.S. tax administration has attempted to construct information trails on self-employed. But even in this high enforcement capacity setting, these efforts have little impact on compliance (Collins et al., 2017)

development path.²⁰

The income tax base is defined as the share of the economically active population whose gross income lies above the income tax exemption threshold. I define the threshold as the minimum level of gross income above which individuals become liable to pay some strictly positive income tax.²¹ I consider a non-married filer without any dependents. Without any further assumptions on filing behavior, the threshold for this filer equates to the standard deduction (or allowance). I focus on this standard deduction, rather than itemized deductions, because it is automatically available to all taxpayers as a fixed amount. This allows me to calculate the threshold in a simple way without having to make any behavioral assumptions about filing and credit take-up (Benzarti, 2018). In addition, while the existence of itemized deductions may vary, the standard deduction is universally available across countries. These choices allow me to create a harmonized measure of the size of the base that is highly comparable across countries and over time. I discuss potential bias arising from standard versus itemized deductions in Section 3.4. To further increase international comparability, I code values of the threshold in all cross-country surveys from a unique source, namely the country tax reports of the IBFD Tax Research Platform.²² I code the value of the threshold in the year of the survey. For the US series, I use the historical IRS series on the personal exemptions.

I measure the distributional employment profiles by calculating the share of employees in the economically active workforce in deciles of a country's gross income distribution, using individual survey weights. I locate the threshold in every gross individual income distribution, and calculate the base as sum of percentiles above the threshold. My employment classification is consistent in the aggregate with the U.N. SNA, which means that all activities identified within the SNA production boundary are captured in the employment estimates in every income decile in every country. As such, the size of the base can be interpreted as the share of the economically active population that is subject to income tax. I use expenditure-side real GDP at chained PPPs in 2011 US\$ from the same year as the country survey data.

²⁰I can always distinguish between working for someone else for pay versus for in-kind payment, and I exclude the latter from the employee category. As such, casual wage-laborers that receive in-kind payment are systematically classified as self-employed. This leaves casual workers that are not paid in-kind as the group that I potentially mis-classify as employee, whenever the survey answers do not provide sufficient precision about the nature of the employee-work. I discuss measurement-induced bias in Section 3.4.

²¹A very limited number of countries, many of which are located in Eastern Europe, have recently moved to the broadest possible income base, where the threshold is 0. The income tax system in all the countries in my sample feature a strictly positive threshold in the year of the survey data. A few countries feature a 0-th bracket, where the filer still has to file an income tax return but the tax rate is 0; in this case, I define as the threshold the value of income above the 1st bracket where there is a strictly positive rate.

²²IBFD is the primary data-source for the country-specific tax summaries produced by several international consultancies (including Deloitte, KPMG, and PwC).

3.3 Results

In this sub-section, I present new stylized facts on changes to employment structure and tax structure along the development path. To summarize findings in the cross-country section, I partition the sample of 100 surveys into ten development groups of 10 countries. In each group and decile of the gross income distribution, I construct the unweighted average of employee-shares in all countries in that group.²³ I complement the cross-country evidence with time-series evidence from the US (1870-2010), with a focus on understanding how the US looked like historically when it was less-developed. The time-series spans almost as large an income range as the cross-country data: in 1870, the US had a similar level of per capita income to India in the survey year.²⁴

Stylized fact #1: Within country employee share increases over the income distribution, and at all levels of income as a country develops

Figure 2 displays the employee distributional profiles in the 10 cross-country development groups. The first stylized fact documents that within each country, the employee-share increases with income, and that across development, the upward-sloping employee-profile moves leftward in the gross income distribution. In other words, in each successive development group, the increase in employee share is concentrated in deciles gradually further down the income distribution. In the poorest countries in the sample, such as Democratic Republic of the Congo, Niger, and Malawi, the employee-share is effectively nil apart from a steep increase in the top percentiles of the income distribution. During the transition from low to lower-middle income countries, such as India, Pakistan, Zambia and Bolivia, the growth in employee share is concentrated in the top three deciles. As the transition to upper-middle income countries takes place, to countries including China, Brazil, South Africa, and Turkey, the growth in employee-share occurs mainly in the four middle deciles (from the 4th to the 7th decile). Finally, when going from upper-middle to high-income countries including the US, UK, Germany, and Japan, the remaining growth in employee share is concentrated in the bottom three deciles.

In Figure 3, I plot the US profiles between 1870 and 2000. In 1870, on the eve of its second industrial revolution, the employee-share in the US was concentrated in the top decile. Over time, the employee-profile gradually moves leftward in the income distribution. For every US historical profile, I construct a 'synthetic' cross-country profile, based on the average of profiles of countries with per capita income within 10% of the US

²³As an alternative, I can construct weighted profiles, using the country's population share in total population of the development group as the country-specific weight. Those results are similar.

²⁴Shortly after the year of the survey, 2004, India graduated to the lower-middle income group. As such, the US historical data spans development levels between the lower-middle and high-income groups.

income in the historical year.²⁵ I insert these synthetic profiles in Figure 3. Interestingly, the profiles look similar both in trends and in levels, when comparing a historical US profile to a cross-country profile at the same level of per capita income. This stylized fact suggests that enforceable employee-income gradually rises further down the income distribution over the development path.

Stylized fact #2: Over development, the tax exemption threshold moves down the income distribution in close co-movement with increases in employee share

The second stylized fact documents a strong co-movement over development between increases in employee share occurring gradually further down the income distribution, and decreases in the threshold. This is illustrated using four countries from the cross-section and four points in time within the US in Figure 1. At lowest levels of development, the threshold is systematically located in the top percentiles of the income distribution, generating a very narrow base between 1% and 5% of the workforce. Over development, we observe a close co-movement between decreases in the location of the threshold and growth in employee share locally to its left. Given the gradual shifts in employee-share profile (fact #1), this co-movement generates a gradual increase in the size of the base over development, as documented in Figure 4. The base is very narrow in low-income countries, taxing only 1-10% of the economically active population. The base expands in middle-income countries to 30-50%, and reaches 90-95% in high income countries. Again, the patterns are similar when comparing the expansion of the base across countries to the historical expansion within the US over time (Figure 4). This stylized fact suggests that governments gradually expand the base over development, by decreasing the threshold which tracks growth in enforceable employee income.

Stylized fact #3: Location of tax exemption threshold systematically maximizes the employee-share on the income tax base

When comparing the poorest to the richest countries in the cross-section, panel B of Figure 4 implies that the base expands by over 4500% over the course of development. The third stylized fact shows that despite this large variation in size over development, the composition of base remains constantly maximized in employee-share throughout. To show this, I locate in all surveys the percentile where the location of the threshold would maximize the employee-share on the base, and calculate the employee-share above this percentile. I then calculate the employee-share in the actual base, as the employee-share above the actual threshold. Finally, I divide the actual employee-share by the maximized share, and plot the ratio against GDP per capita in Panel C of Figure 4. This ratio would be equal to 1 if a country locates its threshold in the exact percentile

²⁵The US in 1870 had similar per capita income to India and Pakistan in their survey-years. The US in 1935 was similar to Indonesia and Peru, and the US in 1950 was similar to Mexico and Uruguay.

which maximizes the employee-share on the base. I find that on average, the actual employee composition of the base achieves 95% of the maximized employee-composition. What is more, this ratio is constant across all levels of development, both across countries and within the US over time.²⁶ This stylized fact shows that the base is systematically a tax on employee income throughout development, consistent with a constant policy-rule that seeks to maximize the enforceability of the base.

Stylized fact #4: Decreases in the exemption threshold are strongly associated with increases in tax collection

Since I am introducing a novel tax policy instrument, I benchmark its association with tax collection against the top marginal tax rate, which is the main policy instrument studied in other theories of income tax growth (Besley and Persson, 2011; Gordon and Li, 2009; Scheve and Stasavage, 2010). The fourth stylized fact shows that differences in size of the base are strongly associated with differences in income tax collection (relative to GDP). In Panel A of Figure 5, I find that size of the base and level of income tax collection increase at similar pace over development: both variables slowly increase between low-income and lower-middle income countries, but witness strong growth during the transition from upper-middle to high-income countries. In Panel B, I find a strongly significant and positive direct relationship between the base and tax collection, while controlling for the tax rate: the expansion from a narrow base (less than 10%) to a broad base (more than 90%) is associated with an increase in income tax collection from 1-2% of GDP to 15-18% of GDP. Combined with fact #3, this fact shows that the highly enforceable base is a key policy tool which drives growth in tax collection. This contrasts with the main previously studied policy tool, the marginal tax rate, which is neither correlated with development (Panel A), nor with income tax collection (Panel B). Finally, as an auxiliary finding, I show in Panel C that once I control for the size of the base, there is no remaining association between income tax collection and development.²⁷ This suggests that the base may be a novel important channel through which fundamental factors, in addition to employment-structure, impact long-run tax collection.

²⁶The extent to which this ratio can deviate from 100% depends on the steepness of employee-income profile. If the employee-share was uniform across the income distribution, any arbitrary location of the threshold would generate a ratio of 100%. As such, this test is informative of policy-rules in settings except for high-income countries, where the employee-share is high and uniform across income (Figure 2).

²⁷There remains variance in the residual collection, at a given level of development. This is particularly true in high-income countries, where additional factors beyond the size of the base contribute to differences in tax collection.

3.4 Robustness

In this sub-section, I provide robustness checks for the stylized facts. In Appendix A.2., I show that the novel facts also hold within Mexico over 50 years (1960-2010). This time-series covers the lower levels of per capita income in the cross-country data. In the Mexican historical data, I uncover the same stylized facts: the employee-share profile is upward-sloping and gradually moves leftward in the income distribution; the threshold gradually moves down the distribution by tracking growth in employee-share and expands the base; the employee-composition on the base is constantly maximized.

In a second robustness check (Appendix A.3), I consider the possibility that the threshold is set for purely redistributive or anti-poverty purposes.²⁸ I show that there is no explicit association between the tax threshold and redistributive thresholds: in less than 5% of countries in the sample is the threshold defined in the tax code as either equal to, or a multiplicative of, the national poverty line or the minimum wage. In addition, I show that there does not either appear to be any implicit relationship between these redistributive thresholds and the tax threshold, as countries with similar levels of income and size of base display enormous variation in the ratio of the tax threshold to the poverty line or the minimum wage.²⁹

In a third robustness check (Appendix A.4), I find limited evidence that the threshold avoids hard-to-tax sectors, including agriculture (Musgrave, 1981), rather than targets enforceable employees. First, I find that only 12% of governments in both low and middle income groups, and 5% of high-income governments, explicitly exempt agricultural income from income taxes.³⁰ Second, I show that only over a limited range of the development path is the agricultural employment profile downward-sloping and leftward-shifting in a manner consistent with movements in the threshold. Over this limited range, agriculture is virtually entirely made up of self-employment. I also find that none of the easy-to-tax sectors - mainly manufacturing and government - have the required upward-sloping and leftward-shifting distributional profiles to be confounding determinants.

In a fourth robustness check (Appendix A.5), I consider potential biases from the survey methodology and measurement of variables. First, systematic under-reporting of self-employed income, due to measurement issues (de Mel, McKenzie and Woodruff,

²⁸I collect the poverty lines that are set by the national government, rather than the international poverty lines defined by the World Bank.

²⁹Of course, the threshold redistributes through the tax system directly, as it is defined based on an individual's income. In Section 4, I provide an estimate of the impact of employee-share on the threshold, in a setting where income was constant.

³⁰This includes India and Pakistan. I code this in the IBD country reports. I only count instances where agricultural income is entirely exempt. There are cases where self-employed operating in agriculture may deduct specific costs to determine their tax liability, but I do not include such instances (Appendix A.4).

2009) or evasion (Kleven et al., 2011) would push the true self-employed profile to the right of the measured profile. However, studies suggest evasion-motives are much less present among self-employed in survey data than in administrative data (IRS, 2008). Similarly, I find no evidence of bunching in the micro-data, which is consistently found in administrative data and related to evasion.³¹ Increases in enforcement capacity and survey precision over development would lead to *rightward*-shifts of the self-employed profile. As such, systematic misreporting either does not introduce bias in levels, or it leads me to understate the true leftward shift in self-employed over development. Second, individuals may hold multiple jobs at the time of the survey or a different job outside the survey reference period. This is unlikely to generate significant bias in the representativeness of my estimates, as these different jobs are most likely to fall within the same category of my employment classification.³² Finally, I abstract from modeling itemized tax deductions, but frequent take-up of deductions would introduce a wedge between the size of my measured base and the size of the true base. This wedge would lead me to overstate the true size of the already narrow base in low-income countries. In the Appendix, I argue that the most plausible form of systematic bias would lead me to understate the importance of the base as a determinant of tax collection.

Finally, I use a regression setting to investigate the robustness of the employee-income slope (Appendix A.6). In all countries, I estimate the likelihood at the respondent-level of being employee as a function of income, and find that the gradient is robust to the inclusion of controls for sector (4 industry dummies), geography (urban dummy), and education (three levels of educational completion dummies).

3.5 Discussion

These stylized facts are strongly consistent with the proposed theory of long-run income tax growth. They suggest that the rise of a modern broad-based tax system occurs through the expansion of the base which tracks growth in enforceable income along the income distribution. Fact #1 constitutes the first micro-evidence along the full development path consistent with the importance of information trails and enforceability for long-run tax collection (LaPorta and Shleifer, 2014; Kleven et al., 2016).³³ Fact #2 provides an extension to these theories, by introducing the threshold as a new policy mech-

³¹This is consistent with Hsieh and Olken (2014), who also use non-tax administrative data and find little evidence for bunching of firms around regulatory cut-offs in India, Indonesia, and Mexico.

³²In developing countries, most own-account workers with multiple jobs are likely to simultaneously work in other self-employed jobs, including contributing to the family farm, or working in the household enterprise. Similarly, most individuals that engage in seasonal wage-labor are likely to hold self-employed jobs outside of demand season.

³³In the next section, I provide the first plausibly well-identified evidence consistent with these theories.

anism through which growth in enforceable income covered by information trails leads to growth in tax collection. Through this mechanism, the base grows by over 4500% over development.³⁴ Despite this large growth in size, fact #3 shows that the base composition systematically maximizes its employee-share. This is consistent with a high employee-share being a necessary condition for an enforceable base. It suggests that high enforceability of the base is a constant policy rule throughout development.³⁵

The proposed theory is consistent with variation in both tax collection and income tax policies across countries that span the full range of development levels today, and within developed and developing countries over the long-run.³⁶ I argue that this contrasts to the main competing theories of income tax growth. One set of alternative theories argue that demand for public goods (most noticeably through wars, as in Scheve and Stasavage, 2010) or inequality (Peltzman, 1980) lead to higher collection through increasingly progressive income tax systems.³⁷ Another set of theories argues that expansion of democracy should lead to higher collection (a review of these theories is provided in Acemoglu et al., 2015)³⁸ These theories have in part been derived in the context of developed countries' historical trends up to the 1950s.³⁹ Since 1960, Mexico has seen a continued expansion of its base and collection, while the U.S. base has stagnated. Since 1980, China's base has dramatically expanded, while India's base has remained constant in size (Piketty and Qian, 2009). These more recent periods of growth and stagnation of the income tax system are directly consistent with corresponding periods of increased growth or stagnation in employee-share. In contrast, these periods did not see the changes to expenditure or inequality required by demand theories (Alvaredo et al., 2018), nor the change in democratization required by institutional theories.⁴⁰ In addition, it is not a priori clear how

³⁴In the cross-country setting, the least developed countries have a base which is on average 2% in size, while it is 92% in the most developed countries.

³⁵Fact #3 also provides systematic documentation to the policy observation, made across a variety of settings and over time, that a country's personal income tax is essentially a tax on employees (Bird and Oldman, 1964; Musgrave, 1981; Bird and Zolt, 2003).

³⁶As I show in the following section, the proposed theory is also consistent with observed variation within tax systems of individual US states over time.

³⁷These effects can be permanent, both through the "Ratchet effect" and through investments in fiscal capacity (Besley and Persson, 2011).

³⁸The evidence on the impact of democracy on income taxes is mixed, however. Aidt and Jensen (2009) find that suffrage expansion reduces individual income and corporate income taxes. Scheve and Stasavage (2010) find no association between democracy and income tax progressivity in a long-run analysis of OECD countries. Acemoglu et al. (2015), on the other hand, find a positive impact of lagged democracy on total tax collection in a large sample of countries over time. However, these effects could be driven by changes in other taxes apart from personal income taxes.

³⁹The long-run data in the US shows that my proposed theory is also consistent with the trends in tax systems over this historical period.

⁴⁰Per example, China's level of democratization has stayed constant since 1950, according to the widely used Polity measure. These arguments are based on descriptive patterns across a limited number of coun-

these alternative theories would predict a tax base composition that is constantly high in employee-share over development (fact #3). Finally, these competing theories show changes in top marginal tax rates as the main mechanism result for income tax growth. However, I find that the top marginal rate is effectively uncorrelated with income tax collection across countries at different levels of development (Figure 5).

There are important aspects of long-run taxation that the theory is silent on. By focusing on enforceability of earned income, I abstract from capital income. When the income tax was implemented in the U.S. in 1913 in the 99th percentile of the income distribution, capital constituted an important source of taxable income (Piketty and Saez, 2003). In currently low-income countries with narrow bases, it is similarly plausible that the evolution of capital income, and the societal demand to tax it, also impact base expansions.⁴¹ By focusing on individual earners' enforceability, I abstract from the important issue of information trails on firm's sales (Pomeranz, 2015; Naritomi, 2019). A similar research design employing descriptive and quasi-experimental evidence could be employed to study whether, and how, changes in information trails on firms can explain the gradual increase in reliance on VAT observed across development (Kopczuk and Slemrod, 2006).

4 Direct estimate of impact of employment structure on the tax base and tax collection

While the stylized facts are consistent with the proposed theory, they do not allow me to rule out confounding determinants of tax systems that often vary jointly with employment-structure over development. In this section, I therefore provide direct evidence of employment structure's impact on the income tax system, using the setting of US states over time. I first provide background on the data and source of variation. I then discuss the identification strategy. I present graphical evidence and regression-based results, before providing mechanism results and robustness checks.

tries. In the following section, I show a direct impact of employment-structure on collection, in a setting which where there was no change in demand for public goods, inequality, and political institutions.

⁴¹Capital income remains more concentrated among top incomes than earned labor income. As such, when the base is broad and covers middle and low-income individuals, the threshold would be ineffective as a tool to increase the burden of taxation on capital income. In such a setting, the increase in marginal rates at high levels of income would be a more appropriate policy response. A joint study of the evolution of earned income and capital income could shed light on the current debate about income versus wealth taxation in developing countries (Londono-Velez, 2019; Piketty, Yang and Zucman, 2019).

4.1 Background and program details

Background and data

US states represent a potentially compelling setting to study development of tax systems. Each state chooses tax policies, including an exemption threshold and rates for personal income taxes. Each state is responsible for the associated collection and enforcement costs.⁴² The cost of collection in the average state tax administration was 4.3% of taxes in 1960; this is comparable to the average cost of collection of national tax administrations in low-income countries today, which 3.8%.⁴³

Similar to the cross-development setting, the average state has seen a significant growth in tax collection since 1950 which is driven by an increase in personal income taxes (Panel A, Figure 6). Over the same period, the average state underwent a significant transition from self-employment to employee-jobs. Consistent with stylized fact #1, within the average state the employee-share increases with income, and over time the upward-sloping employee-profile moves leftward in the state's income distribution (Panel B, Figure 6). Over time, the state income tax threshold gradually decreases in the income distribution, appearing to track growth in employee-share locally to its left, consistent with stylized fact #2. Finally, Panel B shows that the employee-share above the threshold remains constant and high, consistent with an expansion of a highly enforceable base in stylized fact #3. The replication of the stylized facts in the U.S. states increases the number of settings where the theory is consistent with variation in tax collection and policies. It also reinforces any external validity of the direct evidence from this section for the cross-country setting in the previous section.

I combine data from several sources to construct variables that are as comparable as possible to the cross-country section. I calculate employment structure in the active workforce using decennial Census data between 1940 and 2010. The 'class of worker' category in the Census data allows me to classify respondents as either self-employed or employee in a manner that is consistent over time within US states and with the cross-country classification. The advantage of the Census data is that it provides a measure of employee-share of the workforce. The disadvantage is that it is not yearly. I interpolate the numerator and denominator between Census years using a natural cubic spline. I find similar results when using the continuously available measure of the employee-share of income. I rely on the US historical state tax calculator 1900-2007, constructed by

⁴²State tax authorities also make investments in administrative and enforcement capacity similar to national tax authorities. These include filing processes, withholding, and auditing systems.

⁴³The cross-country measure is from Jensen and Lagakos (2019). The cost includes capital outlays and the payroll of the tax administration. More details are found in Appendix Section B.1.

Bakija (2009), to construct measures of the state-year income tax structure. In particular, I calculate the tax threshold for an individual single earner who claims the standard deduction. This definition of the threshold matches the definition used in the cross-country setting. To construct tax collection levels, I use state government finance publications between 1929 and 2010 published by the US Census. I express tax collection as a fraction of total state personal income, from the historical BEA series.⁴⁴

Program details

To establish a direction of causality from employment structure to the income tax system, I exploit variation in the effective implementation date of the Industrial Development Bonds (IDB) program. Through the IDB program, sub-state government units (most often counties and boroughs) issued bonds to finance the acquisition or construction of facilities and equipment for lease to private firms (Cobb, 1993). According to an early Federal definition, the IDB program was intended to attract industrial plants to rural communities characterized by under-employment among self-employed (Area Redevelopment Administration Commission [ARA], 1962). The IDB would act as a positive level-shifter in employee-share. The dominant justification for government intervention was that inadequate local credit for industrial firms constituted the barrier to expansion in rural communities (ARA, 1962).

Importantly, IDB issuances were revenue bonds, which are secured exclusively by the future revenues of the project. This is in contrast to general obligation bonds which are secured by the credit of the issuer. This distinction implies that IDB did not directly lead to increased revenue need to support the credit quality of the state government.

By enabling private industry expansion, IDB was considered to be directly incompatible with the 'public purpose' provision. According to this provision, government debt may only be issued for public purposes. Implementation therefore required the state House to vote in a legal statute which exempted IDB from the public purpose provision. In the cases where the public purpose provision was directly written in the state Constitution, the vote-in led to a constitutional amendment. But there was no legal historical precedent to such development program. The voted statute or amendment, and by extension any bond issuance, would therefore remain legally uncertain until the highest state court would litigate to uphold its constitutionality through a specific court case (Abbey, 1965; Cobb, 1993; Pinsky, 1972; Rollinson, 1976).⁴⁵ The court case would most

⁴⁴The construction of all the variables used in this section is discussed in detail in Appendix B.1.

⁴⁵The fact that IDB was issued as revenue bonds, rather than general obligation, was often the basis of the argument for not violating the public purpose provision. In the case of *Wayland v. Snapp*, the Arkansas Supreme Court "(...) chose to uphold the issuance of the revenue bonds by invoking the doctrine that rev-

often be triggered by any of the administrative steps that the local government was required to fulfill, such as the set-up of local administrative units devoted to handling the IDB issuance process.⁴⁶

I code the year of vote-in and the year of upholding in the 28 states that implemented the IDB program.⁴⁷ As a precautionary measure, I code these events from both legal reviews and government publications, but there is almost no disagreement across sources.⁴⁸ The time-lag between vote-in upholding has mean of 6.67 years and standard deviation of 6.77, an exceeds 9 years in over 40% of IDB states.⁴⁹

4.2 Identification strategy

The legal uncertainty surrounding the IDB program suggests that the vote-in event would be a fuzzy treatment. Instead, I center the program impact around the court upholding event which, as I show below, effectively begins the program. To account for the possibility that the willingness to implement the program may be endogenous to changes in the political and economic environment, I include a dummy variable for the intermediary period between the vote-in event and the upholding event. This control captures any state-time varying observable and unobservable shocks to the political and economic environments which occur around the time of the endogenous policy-decision to vote. This includes willingness to start the program in response to shocks to employment structure or changes in 'need' or 'taste' for tax revenue. The estimation of impacts is therefore based on the upholding event, which I show is plausibly exogenous, rather than the vote event, which is most likely endogenous.

This discussion motivates the estimation of the following regression

$$y_{st} = \beta + \alpha \mathbf{1}(\text{Vote-in})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \lambda \mathbf{X}_{st} + \mu_s + \gamma_t + \varepsilon_{st} \quad (1)$$

where s denotes state, t denotes time, $\mathbf{1}(\text{Vote-in})_{st}$ is a dummy taking a value of 1 when the state has voted in the IDB program but it has not been upheld in the state court

enue bonds do not violate a credit clause because they are retired through lease revenues of the project, not out of tax funds" (Pinsky, 1972).

⁴⁶I provide additional information on the institutional details of the IDB program and its compatibility with the public purpose provision in Appendix B.2.

⁴⁷Starting in 1971, a series of regulatory reforms led to the merge of IDB with other sub-state funding programs. Most importantly, these reforms expanded the scope of projects approved under municipal bond funding, shifting the nature of funding from rural industrialization towards public-goods projects in infrastructure and environmental conservation.

⁴⁸In those few cases where there exists a deviation between sources, I pick the earliest case as the year of the upholding. Details on the leading court cases and sources are in Appendix B.2.

⁴⁹The distribution of the time-lag across all IDB states in in Appendix B.2.

system, and $\mathbf{1}(\text{Upheld})_{st}$ is a dummy taking on a value of 1 when the state supreme court has upheld the legality of the constitutional amendment. Exogenous control variables in \mathbf{X}_{st} include dummies for election years, and the first-stage variables used in Besley, Persson, and Sturm (2010) to instrument for political competition and tax and expenditure policies.⁵⁰ All standard errors are clustered at the state level to allow for correlation over time within a state. The estimation in (1) assesses program impacts by comparing changes in outcomes before and after the upholding event, relative to changes before and after the vote-in event within the same state. Variation to estimate θ comes both from the within-state lag between vote and upholding, and from the cross-state variation in timing of the upholding event and in length of lag.

To interpret the coefficient on $\mathbf{1}(\text{Upheld})_{st}$ as the causal impact of the IDB program requires the identifying assumption that, conditional on the vote-in and the included controls, the timing of the court upholding decision is uncorrelated with the error term in (1). I provide two pieces of supporting evidence. First, below I provide graphical evidence that the regression effects are driven by clear deviations from a stable pre-trend following the upholding event. Groups with differing length of lag between vote-in and upholding are on parallel trends prior to the upholding event, suggesting that groups with longer lag constitute a valid counterfactual for groups with shorter lag.

The second supporting evidence is that the only significant predictors of the time lag are state-specific, time-invariant historical characteristics. Table 1 reports the results from non-parametric Cox proportional hazards models. These models use state time-varying and time-invariant regressors to predict the probability of the upholding event occurring, conditional on the vote having occurred. A civil law dummy significantly predicts a higher conditional probability of upholding. This result is consistent with studies across US states (Berkowitz and Clay, 2005) and across countries (La Porta, Lopez-de-Silanes and Shleifer, 2008).⁵¹ In addition, states that have experienced historical defaults on public-private funding schemes take longer to uphold. This is consistent with a legal literature which argues the 'public purpose' provision was more entrenched in states that have experience with the adverse outcomes of IDB-type funding schemes (Sbragia,

⁵⁰The first stage instruments measure the share of the state population which are subject to either a literacy test or a poll tax (or both), and which were abolished after the 1965 Voting Rights Act. In Appendix B.3, I show that the main estimates are robust to including additional, potentially endogenous control variables: log per capita income, the existence of right to work laws, the existence of a state corporate income tax, and the firm-size coverage of state unemployment insurance.

⁵¹One interpretation of this result, drawing on Berkowitz & Clay (2005), is that civil law produces a Constitution with more statutory components, rather than framework provisions, and that the existence of statutory laws created more frequent demand for constitutional change among affected groups as the political and economic climates change over time. This explanation is consistent with the difference in IDB litigation procedures observed across states: civil law origins states were more likely to vote in statutes, as opposed to Acts, which in turn were likely to be revised more quickly.

1996).⁵² Conditional on these two characteristics, I find that none of 14 political and economic potential determinants are significant predictors of the time-lag.

I make two sample selections to estimate (1). First, I center the sample in every state around the year of the upholding event, in a time-window which ranges from 5 years prior to the vote-in event to 5 years following the upholding event. Consequently, θ will only pick up short-run impacts of the upholding decision. I do this primarily to increase the likelihood of capturing the isolated, direct employee-impact of IDB. This choice decreases the likelihood of capturing confounding medium and long-run impacts, including endogenous policy-responses (Besley and Case, 1994) and indirect economic effects (such as spill-overs and agglomeration effects; Kline and Moretti, 2013). This short-run specification also reduces concerns over inadequate controls for long-run trends in employment structure.⁵³ The main drawback is that my estimates are silent on the long-term impacts of IDB. Second, I limit the sample to IDB states, since states that never implement IDB may be fundamentally different in ways that are hard to control for. The inclusion of non-IDB states could improve estimation of controls in (1), and I show results are robust to their inclusion in Appendix B.3.

4.3 Graphical results

Before turning to the regression results, I provide graphical event-study evidence. I study outcomes in two groups which constitute 60% of IDB states: those with a time-lag from vote-in to upholding between 5 and 10 years, and more than 10 years. I show the yearly average of outcomes in each group over five years prior to the vote-in, the full intermediary period, and ten years after the upholding event. This event-study allows me to graphically illustrate the estimation strategy in (1), and to flexibly visualize the patterns of outcomes relative to upholding and vote events directly in the raw data.

The results are displayed in Figure 7. Panel A studies the impacts on cumulative issuance of IDB debt.⁵⁴ The panel displays a sharp and large increase in issuance immediately following the upholding event. In contrast, there is no issuance following the vote-in, which suggests the upholding decision was a strictly binding constraint on the effective start of the program. I also do not find any issuance in the years immediately preceding the upholding, suggesting there was no meaningful anticipation of when the court decision would occur.

⁵²I show in Appendix B.3. that results are not driven by changes over time in unobservable confounders that are correlated with these two historical characteristics.

⁵³In Appendix B.3., I show that results are robust to different strategies to control for trends.

⁵⁴I code the full set of IDB issuances in every state and year, using the historical series of Moody's Municipal and Government Manual.

In Panel B, I consider the impacts on employee-share of earned income. I use this variable as a proxy for employment structure because it is available on a continuous basis, unlike the employee-share of employment variable.⁵⁵ I find a clear positive deviation from trend in employee-share that coincides with the immediate post-upholding years. This deviation occurs relative to a stable pre-trend over 15 years, which includes the vote event. Note also that the two groups are on parallel-trends throughout the pre-upholding period, suggesting states with longer time-lag constitute a potentially valid counterfactual for states with shorter time-lag.

In Panel C, I study whether the upholding event is associated with reforms to the state income tax threshold. With a two-year lag to the upholding event, I find a large reduction in the nominal value of the threshold. Again, the sharp decrease contrasts to a long and stable pre-period, where both groups are on parallel-trends. Two comments are in order. First, Panels B and C provide compelling graphical evidence of positive impacts of IDB on employee-share and the base, which are directly observable in the raw data. Both variables are constant throughout the long pre-upholding period, and see clear deviations from trend that coincide with the immediate post-upholding years. In addition, the base expansion occurs with a short lag to the increase in employee-share. Second, Panel C directly shows a real policy-response to the upholding event: state governments engaged in active reforms to lower the threshold and increase the base.⁵⁶

4.4 Regression results

I confirm the graphical results in a regression format that estimates (1). Results are displayed in Table 2. In column (1), I find that there is a 1.7 percentage point increase in the employee-share of employment in the 5-year period following the upholding event. Column (3) shows that the upholding event led to a significant reduction in the ratio of the threshold to per capita income. This result implies that the threshold decreased in the income distribution, and the base increased. Despite the increase in size of the base, the upholding event did not change the employee-share on the base (Column 5).

To support the interpretation that these results are evidence of a direct impact of employee-share on the income tax, Table 3 shows absence of impact on confounding de-

⁵⁵This variable excludes all transfers received from Federal and state government. The denominator contains all sources of earned income: employee farming, employee non-farming, self employed farming, self-employed non-farming. The denominator excludes all sources of non-earned income, such as dividends and interest payments. I find similar regressions results when using either employee-share of earned income or employee-share of employment (Appendix B.3).

⁵⁶An alternative 'reform' is to leave the nominal value of the threshold unchanged. In the context of income growth, this 'passive reform' will increase the size of the base by lowering the threshold in the real income distribution.

terminants of tax systems. In Panel A, I find no impact of upholding on non-income tax revenues (Columns 1-3), or other income tax instruments (Column 4). Similarly, I find no significant impact of upholding on investments in enforcement capacity (Column 5) and administrative capacity (Column 6). These results suggest that the upholding decision did not coincide with a change in the state's general 'taste' or need for revenue.

In Panel B, Col (1) shows the upholding event led to a statistically and economically insignificant change in per capita income.⁵⁷ Several factors likely contribute to this null result. At the micro level, an individual may switch from self-employment to employee-job without increasing gross income if the transition brings access to insurance. This prediction is derived in the Lewis dual sector model (1954), whose setting of a small, industrialized center surrounded by a large self-employment sector fits well the official description of IDB counties (ARA, 1962).⁵⁸ Consistent with this prediction, I find that the upholding event led to an increase in employer social insurance supplements to salaries.⁵⁹ At the macro level, previous studies of local development programs in the US found income effects only in the long-run over 40 years (Kline and Moretti, 2013). My estimates are in the short-run over 5 years. In addition, the previous programs studied were broad in scope and achieved long-run impacts through agglomeration and spillover effects. In Appendix B.4, I find no such effects in a county-level evaluation of the IDB program; this is consistent with the Federal description of the IDB as being locally targeted and narrow in scope (ACIS, 1965).

The final columns of Panel B show no association between changes in the upholding status and income inequality (Col 2); generosity of social assistance (Col 3); labor regulations (Col 4); political competition (Col 5); and, Democratic vote-share (Col 6).⁶⁰

4.5 Mechanism

The results so far show a direct impact of increases in employee-share on the base. In this sub-section, I show that this effect is driven by the rise in employee-share below the threshold. I estimate (1) separately for the employee-share in every decile of the state

⁵⁷I measure income using the BEA's net earnings by place of residence. In Appendix B.3. I show that the null result is robust to using alternative measures of income from the Census and the IRS.

⁵⁸For additional qualitative description on the IDB counties consistent with this interpretation, see Appendix sections B.2. and B.4.

⁵⁹These results are not reported, but gladly available upon request. During the historical period of IDB, access to government social insurance schemes such as unemployment insurance was limited to employees working in firms above a certain size-threshold. Note, however, that this increase in demand for social insurance is self-funded via social security taxes.

⁶⁰The construction of these variables is explained in detail in Appendix B.1.

income distribution.⁶¹ Figure 8 reveals that the upholding event led to a statistically significant increase in the employee-share concentrated in the middle of the income distribution.⁶² Given the upward-sloping employee-share in the average state's income distribution (Figure 6), this distributional impact implies a leftward shift of the employee-profile. Figure 8 further shows that the increase in employee-share occurred in deciles locally to the left of the pre-IDB threshold location, and that the threshold decreased in the distribution. These distributional results are closely consistent with stylized facts #1-#2. They provide plausibly well-identified evidence on the specific mechanism in the proposed theory, whereby exogenous increase in employee-share cause governments to lower the threshold to capture this growth in enforceable income.

At the center of the theory is the assumption that self-employed and employees differ in information coverage and enforceability. I now provide a heterogeneity exercise to support this interpretation. The intuition, which I formalize in the following section, is that there is an administrative cost associated with taxing a new income segment.⁶³ This cost is the expenditure needed to collect information trails to assess income of all new taxpayers, and it is increasing in the share of self-employed on the segment.⁶⁴ The prediction is that base expansions will be less responsive to changes in employment-structure, when the cost of expansion is less strongly increasing in self-employed share. To test the prediction, I leverage the staggered signature in the 1950s-1970s of Exchange of Information Acts (EoI) between state tax administrations and the Federal IRS (Penniman, 1980). The EoI provided states' tax administrations with a flow of information trails from the Federal IRS.⁶⁵ These information trails were relevant for the state's own tax assessments, since Federal and state income taxes are based on similar sources, and the IRS had a stronger capacity to collect these trails in the historical period. Under an EoI, state administrations could therefore in part substitute their own assessment needs with the Federal information trails, which makes the cost of base expansion less strongly increasing in self-employed share. Under this interpretation, the existence of an EoI should make changes in the base less responsive to changes in employment-structure.

Formally, I estimate

⁶¹I construct the employee-share in every decile in every state, using the decennial Census records. I then interpolate between years, using a natural cubic spline.

⁶²This is consistent with qualitative descriptions (Cobb, 1993), which highlights that the plants opened under IDB were predominantly in low-skill industries.

⁶³I am grateful to Roger Gordon for the initial suggestion which led to this heterogeneity analysis.

⁶⁴This is because it is more costly to collect information trails on self-employed than on employees (OECD, 2006). Models of optimal commodity tax bases (Yitzhaki, 1979; Wilson, 1989) emphasize assessment expenditures as important social costs of any tax system. Recent cross-country evidence shows that such expenditures can be a quantitatively important share of taxes collected by governments (Jensen and Lagakos, 2019).

⁶⁵This is consistent with the positive revenue impacts of EoI signatures found in Troiano (2017).

$$y_{st} = \beta + \pi \mathbf{1}(\text{EoI})_{st} + \alpha \mathbf{1}(\text{Vote-in})_{st} + \phi \mathbf{1}(\text{Vote-in})_{st} \times \mathbf{1}(\text{EoI})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \beta \mathbf{1}(\text{Upheld})_{st} \times \mathbf{1}(\text{EoI})_{st} + \lambda \mathbf{X}_{st} + \mu_s + \gamma_t + \varepsilon_{st} \quad (2)$$

where $\mathbf{1}(\text{EoI})_{st}$ takes a value of 1 if state s has signed an EoI with the IRS in year t . In column 4 of Table 2, I find that the positive responsiveness of base expansions to the court upholding is much larger and statistically significant in non-EoI states. This is despite the fact that the upholding event led to statistically similar increases in employee-share in both EoI and non-EoI states (Column 2, Table 2).⁶⁶ This heterogeneity result is consistent with the interpretation that base expansions depend on the difference in information trails between employees and self-employed.

In the final column of Table 2, I find a large and positive impact of the upholding event on growth in income tax collection (relative to GDP). Given I found no impact on confounding determinants of collection in the previous sub-section, I interpret this result as direct evidence that rise in income covered by information trails through increases in employee-share cause growth in tax collection. Interestingly, the positive impact on tax collection is concentrated in non-EoI states, where the positive impact on the base was also concentrated. This could suggest that, in this particular setting, the main impact of employment-structure on tax collection is through the base.

Robustness I explore the robustness of the main results in Appendix B.3. I find similar results in a regression which removes $\mathbf{1}(\text{Vote-in})_{st}$ and uses only the cross-sectional difference in timing of upholding as a source of variation. I find that results remain unchanged when allowing civil law and historical default states, the only determinants of the vote-uphold time-lag, to be on independent and fully non-parametric time paths. Results are robust to the inclusion of a linear trend interacted with the cross-section of 1940 per capita income. This result alleviates concerns that my main specification does not adequately capture differential convergence patterns in employment structure across states. Finally, results are robust to removing those IDB states that simultaneously participated in other place-based development programs (Kline and Moretti, 2013).

Taken together, these results provide plausibly well-identified evidence both for the direct impact of employment-structure on tax collection, and for the tax base mechanism. Potential threats to this interpretation remain, but the confounding state-year variation

⁶⁶The signing of EoI is staggered over the same period as the implementation of IDB, and states are balanced across EoI signature status by the time of upholding. This suggests that difference in results in EoI vs non-EoI states is not due to differences in sample size across the two groups. In addition, I find that neither the vote nor the upholding event changed the likelihood of signing an EoI (results not reported).

would have to: cause a clear deviation from trend in employment and tax outcomes that coincides with the upholding event; not be captured by any of the proxies in Table 3; and, lead to heterogeneous impacts across EoI vs non-EoI states.

5 Model

In this section, I provide a characterization of the local reform decision to increase the size of the base by lowering the threshold by a small amount. This local reform-environment is consistent with most observed reforms. My characterization of income tax base changes is complementary to recent income tax models that also use a local-reform environment but focus on rate changes (Saez and Stantcheva, 2016). At the same time, I draw on the indirect tax design literature, where Yitzhaki (1979) and Wilson (1989) derive optimal tax formulae for the number of commodities to be included in the commodity tax base. In particular, I focus on the same social costs of base expansions - efficiency costs and administrative costs - as in those models. At the end, I use the model to interpret the mechanism results in the US states and to quantitatively relate the US estimates to the cross-country variation in size of base.

5.1 Setup and empirical prediction

I consider a fixed distribution of income z across workers with pdf $h(z)$ and cdf $H(z)$. I assume exogenous employment shares of self-employed and employees at each income level, denoted respectively φ_z and $1 - \varphi_z$. I assume that φ_z is decreasing with z , similar to stylized fact #1.

If the agent reports income $z \geq K$, then she is liable to pay $\tau(z - K)$. Otherwise, she is not liable for income tax. K is the threshold and τ is the marginal tax rate.⁶⁷ I assume linear utility to abstract from income effects. I assume agents have access to an evasion technology which allows them to pay c in order to mis-report their true income and instead report income at K to evade taxes. I assume that for employees, the existence of information trails which cover all of their income makes it infinitely costly to evade: $c^E(z) = \infty \forall z$. For self-employed, I assume information trails only partially cover their income, making it costly but not impossible to evade: $c^{SE}(z) > 0 \forall z$. I assume $c^{SE}(z)$ is increasing and convex in z .⁶⁸ In this setting, there will exist a marginal buncher at

⁶⁷I assume that the government cannot explicitly differentiate between employees and self-employed in the tax system. I make this restriction to be consistent with the pre-dominant policy observed in practice in my sample of countries (according to the IBFD reports). Allowing the government to set different thresholds and marginal rates depending on employment-type would be an interesting extension to the current model.

⁶⁸These cost-functions are consistent with the micro-evidence from random audits in Denmark (Kleven et

income \bar{z} who is indifferent between bunching and full compliance : $\bar{z} - c(\bar{z}, \bar{z} - K) = K + (1 - \tau)(\bar{z} - K)$. All self-employed with income $z : K \leq z \leq \bar{z}$ will under-report and bunch at the threshold. An increase in τ unambiguously leads to more evaders: $\frac{\partial \bar{z}}{\partial \tau} > 0$. An increase in the threshold will lead to less evaders if the marginal gain from compliance is larger than the marginal gain from under-reporting after the threshold decrease. I will assume this condition hold, that is

$$\frac{\partial \bar{z}}{\partial K} < 0 \text{ if } \tau > c_K(\bar{z}, \bar{z} - K) \quad (3)$$

I model an administrative cost of taxing an additional income segment z which increases in the self-employed share: $a(z) = a(\varphi_z)$ with $\frac{\partial a}{\partial \varphi_z} > 0$. I assume this cost-function is constant in z , but convex in φ_z . This cost represents the administrative expenditure which is required to assess taxes on the segment z by building information trails. Yitzhaki (1979) emphasizes assessment and collection expenditures as important social costs of a tax system. Consistent with this view, recent cross-country data shows these expenditures can represent a significant fraction of total taxes collected, especially in developing countries (Jensen and Lagakos, 2019). Modeling the cost to be increasing in the self-employed share is motivated by the policy literature (summarized in OECD (2006) and IMF (2015)). This literature documents that it is much more costly to collect information trails from a large number of self-employed than to aggregate information trails on employees from a limited number of employers.⁶⁹ Stylized fact #3 is consistent with this view, suggesting governments seek to minimize the share of self-employed taxpayers for whom it is significantly more costly to assess their income.

The net revenue base reflects evading self-employed and administrative costs

$$R = \int_{z \geq \bar{z}} \tau(z - K) \varphi_z dH(z) + \int_{z \geq K} \tau(z - K) (1 - \varphi_z) dH(z) - \int_{z \geq K} a(\varphi_z) dH(z) \quad (4)$$

Consider a reform which locally decreases the threshold: $dK < 0$. This reform will have three effects on revenue: a mechanical gain, a behavioral cost, and an administrative cost. The mechanical gain, dM , reflects the increase in revenue collected on the infra-

al., 2011), which found no evasion of employee income, and partial evasion of self-employed income. For my main prediction to hold, I do not require that employees incur infinite cost of evasion; I only need that the cost-functions are such that self-employed evade more than employees at the same level of income.

⁶⁹I assume that the government complies with the expenditure specified by the cost function a to tax a new income segment. In other words, I abstract from the possibility that the administration anticipates evasion by self-employed and abandons the costly collection of information trails on their activities. This simple model also abstracts from the possible interactions that may exist between investments to change the cost function a and reforms to the threshold.

marginal agents, assuming no behavioral responses

$$dM = -dK\tau \left[\int_{z \geq \bar{z}} \tau(z-K) \varphi_z dH(z) + \int_{z \geq K} \tau(z-K) (1 - \varphi_z) dH(z) \right] \quad (5)$$

$$\geq 0 \text{ if } dK < 0$$

The behavioral loss, dB , reflects the efficiency cost and consequent loss in revenue due to behavioral responses of the marginal agents

$$dB = -\frac{\partial \bar{z}}{\partial K} dK \tau(\bar{z} - K) \varphi_K \quad (6)$$

$$\leq 0 \text{ if } dK < 0$$

where I have used the local approximation that $\varphi_K \approx \varphi_{\bar{z}}$. This approximation is plausible if the last buncher is not located too far above the threshold. Finally, the local threshold decrease will lead to an additional administrative cost $dC = dK \cdot a(\varphi_K) < 0$ if $dK < 0$. Relative to a mechanical gain in revenue, the base expansion therefore incurs two social costs - an efficiency cost, and an administrative cost. Balancing these social costs against the revenue gain is achieved when $dB + dC + dM = 0$. This yields the characterization for the location of the threshold

$$\frac{K^{Social}}{\bar{z}} = \frac{1}{\left[1 + \left[\frac{\text{Mech gain} - a(\varphi_K)}{\text{Beh loss}} \right] \cdot [\varepsilon_{\bar{z}, K} \varphi_K]^{-1} \right]} \quad (7)$$

where $\text{Mech gain} = \int_{z \geq \bar{z}} \tau(z-K) \varphi_z dH(z) + \int_{z \geq K} \tau(z-K) (1 - \varphi_z) dH(z)$, $\text{Beh loss} = h(\bar{z}) \bar{z}$, and where $\varepsilon_{\bar{z}, K}$ denotes the elasticity of the marginal buncher with respect to the threshold. An increase in the employee-share locally around the threshold causes a decrease in both of these social costs, leading to a decrease in the threshold. The model delivers the main empirical prediction

Empirical prediction: An increase in employee share locally around the threshold leads to an expansion of the income tax base through lowering of the threshold

$$\frac{\partial K^{Social}}{\partial(1 - \varphi_K)} < 0 \quad (8)$$

Discussion

This model emphasizes two social costs - administrative and efficiency - through which increases in employee-share locally to the left of the threshold lead to an expansion of the base. The main prediction is consistent with the stylized cross-development facts (Section 3) and the direct results from US states (Section 4). The model-environment has abstracted from re-distributional considerations. Incorporating standard vertical equity preferences, through welfare weights that depend on income, will affect the optimal level of K but not (directly) its dependence on employment-structure.⁷⁰ Another possibility would be to incorporate horizontal equity preferences for a 'fair tax base'. Fairness could relate to the idea that the base should not discriminate in terms of compliance between groups, for a same level of income (Brautigam, Fjeldstad and Moore, 2008). Such preferences would be consistent with the stylized facts, but given the lack of evidence on them, I leave their incorporation into the model for future work.⁷¹

5.2 Relating theory to US states direct evidence and cross-development stylized patterns

In this sub-section, I relate the model to the estimates obtained in the U.S. states (Section 4). There, I argued that under an EoI the state administration could rely on information trails built by the IRS for its own assessment purposes. This allows state administrations to substitute some of their own need to construct information trails with the pre-existing trails built by the the IRS. One possible way to model the existence of an EoI is therefore as a reduction in the convexity of the administrative cost function a . In such a set-up, the responsiveness of the base to an increase in employee-share would be larger in non-EoI states.⁷² This is consistent with the heterogeneity result in Table 2.

⁷⁰This is because my main prediction considers a change in the employment-structure, at a given level of income. Of course, the positive empirical association between income and employee-status (stylized fact #1) means that vertical concerns will implicitly lead to redistribution across employment-types.

⁷¹Consider the employee-group and self-employed group on the base. One potential fairness criteria for compliance would be that the employee's share in effective contribution to revenue should be equal to its share in statutory contribution. This could be modelled by a mis-representation index, where society faces horizontal inequity loss from any deviation from a situation of perfect relative representation across groups

$$\text{Horizontal inequity} = \left(1 - \left| \frac{\text{Employee-share on statutory income tax base}}{\text{Employee-share on compliant income tax base}} \right| \right)$$

So long as self-employed evade more than employees, the inequity cost associated with a lowering of the threshold will always be smaller when the employee-share at the threshold is larger. This horizontal equity aversion predicts that increases in employee-share cause expansions of the base (facts #1-#2), and can rationalize the maximized employee-share on the base (stylized fact #3).

⁷²Stronger convexity in the non-EoI setting implies $\frac{\partial a^{\text{non-EoI}}}{\partial \varphi_z} > \frac{\partial a^{\text{EoI}}}{\partial \varphi_z} > 0$. In this case, the derivative in (8) is larger in absolute value in the non-EoI setting than in the EoI setting.

The theory suggests that an appropriate elasticity to summarize the impact of changes in employment structure on the base is $\varepsilon_{[1-H(K)], [1-\varphi_z|z < K]}$. This elasticity indicates the percent increase in the size of the base, $1 - H(K)$, following a 1 percent increase in the employee-share beneath the threshold, $1 - \varphi_z|z < K$. I derive a value of 3.89 for this elasticity, using the results from the US states.⁷³ I then investigate what share of the variation in size of the base across countries (Panel B, Figure 4) can be accounted for by combining the observed cross-country variation in employee-share (below the threshold) and the US states-based value of the elasticity. I find that the elasticity-based predicted change in base can account for 29.9% of the difference in size of the base between a typical low-income country, India, and the US.⁷⁴ The estimated elasticity can account for 68.3% of the variation between a typical lower-middle income country, Peru, and the US.⁷⁵

6 Conclusion

In this paper, I provide evidence that the increase in information trails through the long-run transition from self-employment to employee-employment is a determinant of income tax collection along the full development path. The theory is that increases in information-covered income through growth in employee-share occur gradually further down a country's income distribution over development. In turn, the government increases collection by gradually expanding the base to capture growth in enforceable employee-income. Using a new micro data-base covering 100 countries today and several countries over the long-run, I provide new stylized facts consistent with the theory. I directly test the theory in the context of US states, where I show positive impacts of plausibly exogenous increases in employee-employment on the base and collection.

My findings challenge the policy view, dating back to Kaldor's Indian tax mission in 1956, that a broad-based income tax can and should be at the heart of tax systems in currently developing countries. This paper has documented the existence of significant

⁷³Using the evidence from Figure 8, the upholding event led to a 3.6 percent change in the employee-share beneath K . Using the regression results from column 3 in Table 2, the upholding event led to a 14 percent increase in the size of the base. This implies an elasticity value of $14/3.6 = 3.89$. An alternative would be to compute the elasticity by using values of the key parameters in (7) from the literature. Unfortunately, we lack robust evidence on several of these, including the extensive margin elasticity $\varepsilon_{\bar{z}, K}$.

⁷⁴In the year where I have survey data in India, it was classified as a low-income country. It has since become a lower-middle income country.

⁷⁵Using Figure 1, between India and the US, there is a 347% increase in the employee-share beneath the 98th percentile (the location of the Indian threshold). Using the elasticity estimate, the implied increase in the size of the base is 1349%, while the actual increase in size of the base between India and the US is 4500% (moving from the 2nd to the 92nd percentile of the gross income distribution). Using Figure 1, between Peru and the US, there is a 117 percent increase in the employee-share beneath the (Peruvian) threshold, which implies a predicted increase in the size of the base by 455%. The actual variation in size of the base between Peru and the US is 666%.

enforceability constraints in these countries, which limit the extent to which governments can meaningfully broaden the base. The normative foundation for this policy view is the Atkinson-Stiglitz theorem (1976), which is based on the assumption that governments can implement a flexible income tax system along the full income distribution. My results show this assumption is ill-suited to the reality in developing countries, where the base is constrained and often covers only a fraction of the income distribution. Once we recognize the practical constraints on income taxes, it may be desirable to achieve redistribution through indirect taxes (Bachas et al., 2019).

This paper introduces the tax base as a new policy instrument in public finance and development. In contrast to previously studied tax instruments (most importantly, tax rates), I found large cross-development variation in the size of the base which is closely associated with levels of collection. In fact, I found that after controlling for size of base, there is no remaining association between income tax collection and development (Figure 5, Panel C). This suggests that pre-existing theories of income tax growth, which abstract from the base, should be reconciled with this new policy mechanism.

The micro-database allowed me to document new facts on long-run changes in employment structure along an income distribution. A robust finding has been the close match when comparing historical employee-income profiles of developed countries to profiles of current countries at similar levels of development. Models of structural employment change largely abstract from heterogeneity in work-types along an income distribution, but future work could be devoted to make these models compatible with the documented long-run changes in employee-income profiles.

References

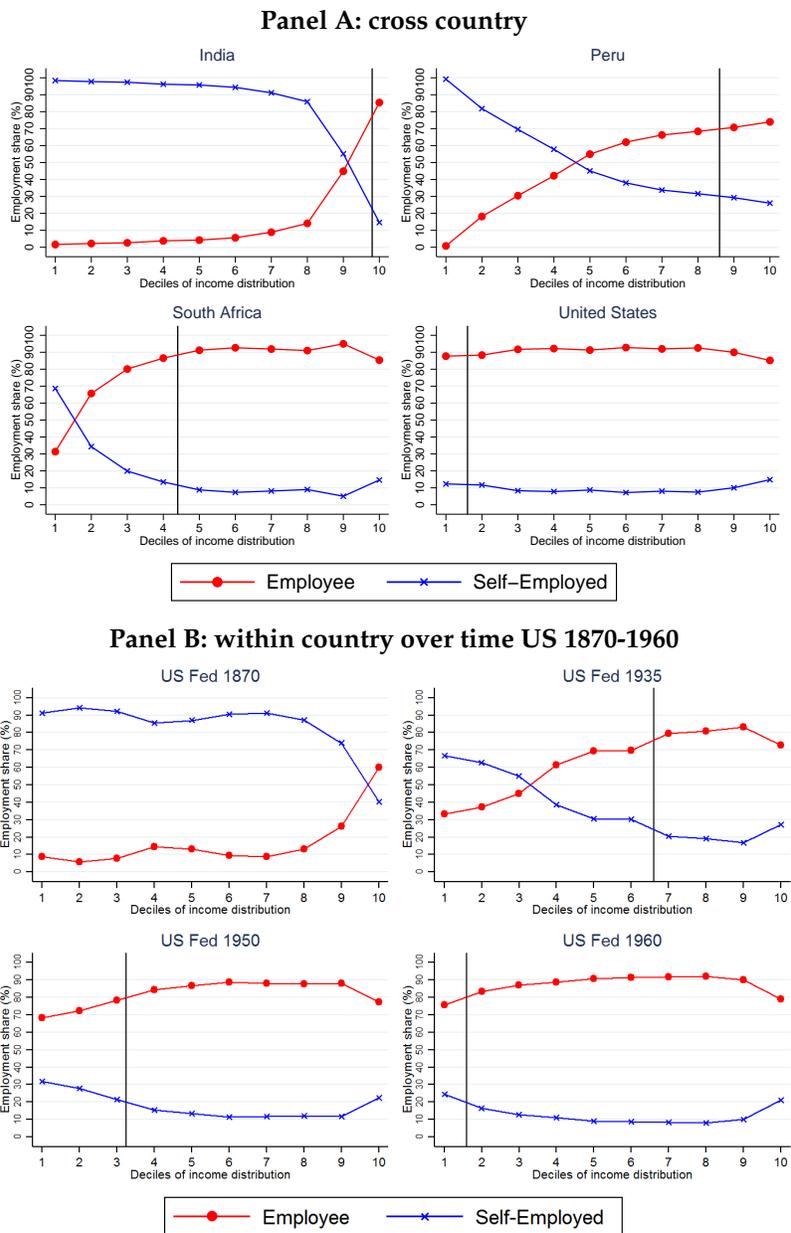
- [1] Abbey, A. [1965]. "Municipal Industrial Development Bonds", *Vanderbilt Law Review*, 19.
- [2] Acemoglu, D. S. Naidu, P. Restrepo, J. Robinson [2015]. "Democracy, Redistribution, and Inequality," in *Handbook of Income Distribution*, Volume 2B, North-Holland.
- [3] Aidt, T., P. Jensen [2009]. "Tax Structure, Size of Government, and the Extension of the Voting Franchise in Western Europe, 1860-1938," *International Tax and Public Finance*, 16(3).
- [4] Alvaredo, F., L. Chancel, T. Piketty, E. Saez, G. Zucman [2018]. *World Inequality Report*. Paris: World Inequality Lab.
- [5] Area Redevelopment Administration [1962]. *Industrial Development Bond Financing: A Commission Report*. Washington D.C: National Archives.
- [6] Atkinson, A., J. Stiglitz [1976]. "The Design of Tax Structure: Direct versus Indirect Taxation", *Journal of Public Economics*, 6.
- [7] Avila, J., J. Londono-Velez [2019]. "Can Wealth Taxation Work in Developing Countries? Quasi-Experimental Evidence from Colombia," *Working Paper*.
- [8] Bachas, P., L. Gadenne, A. Jensen [2019]. "Informal Consumption and Indirect Tax Design: Evidence from Consumption Diaries around the World", *Working Paper*.
- [9] Bakija, J. [2009]. "Documentation for a Comprehensive Historical U.S. Federal and State Income Tax Calculator Program", *Working Paper*.
- [10] Banerjee, A., E. Duflo [2007]. "The Economic Lives of the Poor", *Journal of Economic Perspectives*, 21(1).
- [11] Banerjee, A., A. Newman [1993]. "Occupational Choice and the Process of Development", *Journal of Political Economy*, 101(2).
- [12] Benzarti, Y. [2018]. "How Taxing is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs", *NBER Working Paper Series*, No. 23903.
- [13] Best, M. [2014]. "The Role of Firms in Workers' Earnings Response to Taxes: Evidence from Pakistan", *Working Paper*.
- [14] Best, M., A. Brockmeyer, H. Kleven, J. Spinnewijn, M. Waseem [2015]. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan", *Journal of Political Economy*, 123(6).
- [15] Besley, T., A. Case [1994]. "Unnatural Experiments? Estimating the Incidence of Endogenous Policies", *NBER Working Paper Series*, No. 4856.

- [16] Besley, T., T. Persson [2011]. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press.
- [17] Besley, T., T. Persson, D. Sturm [2010]. "Political Competition, Policy and Growth: Theory and Evidence from the US", *Review of Economic Studies*, 77(4).
- [18] Berkowitz, D., K. Clay [2005]. "American Civil Law Origins: Implications for State Constitutions", *American Law and Economic Review*, 7(1).
- [19] Bicks, A., N. Fuchs-Schündeln, D. Lagakos [2018]. "How do Average Hours Worked Vary with Development? Cross-Country Evidence and Implications", *American Economic Review*, 108(1).
- [20] Bird, R., O. Oldman [1964]. *Readings on Taxation in Developing Countries*. Johns Hopkins University Press.
- [21] Brautigam, D., O. Fjeldstad, M. Moore [2008]. *Taxation and State-Building in Developing Countries: Capacity and Consent*. Cambridge University Press
- [22] Brockmeyer, A., M. Hernandez [2018]. "Taxation, Information, and Withholding: Evidence from Costa Rica", *Working Paper*.
- [23] Carillo, P., D. Pomeranz, M. Singhal [2017]. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement", *American Economic Journal: Applied Economics*, 9(2).
- [24] Cobb, J. [1993]. *The Selling of the South: the Southern Crusade for Industrial Development 1936-1990*. University of Illinois Press.
- [25] Collins, B., J. Hoopes, D. Reck, M. Sebastiani, J. Slemrod [2017]. "Does Credit-Card Information Reporting Improve Small-Business Tax Compliance?," *Journal of Public Economics*, 149.
- [26] de Mel, S., D. McKenzie, C. Woodruff [2008]. "Measuring Micro-enterprise Profits: Musk We Ask How the Sausage is Made?," *Journal of Development Economics*, 88.
- [27] Dincecco, M., U. Troiano [2015]. "Broadening the State: Policy Responses to the Introduction of the Income Tax", *NBER Working Paper No. 21373*.
- [28] Donovan, K., J. Lu, T. Schoellman [2019]. "Labor Market Flows and Development," *Working Paper*.
- [29] Frias, J., T. Kumler, E. Verhoegen [2015]. "Enlisting Workers in Improving Payroll-Tax Compliance: Evidence from Mexico," *NBER Working Paper No. 19385*.
- [30] Gollin, D. [2008]. "Nobody's Business But My Own: Self-employment and Small Enterprise in Economic Development", *Journal of Monetary Economics*, 55(2).
- [31] Gordon, R., W. Li [2009]. "Tax Structure in Developing Countries: Many Puzzles and a Possible Explanation", *Journal of Public Economics*, 93(7-8).

- [32] Hinrichs, H. [1966]. *A General Theory of Tax Structure Change During Economic Development*. The Law School of Harvard University.
- [33] Hsieh, C-T., B. Olken [2014]. "The Missing 'Missing Middle'," *Journal of Economic Perspectives*, 28(3).
- [34] ICPSR based on US Department of Labor [2009]. *Study of Consumer Purchases in the United States, 1935-1936*. ICPSR study 8908. Michigan: ICPSR.
- [35] Internal Revenue Service [2008]. "Differences in Income Derived from Survey and Tax Data," *Statistics of Income Paper Series*.
- [36] International Labour Organization [2002]. "Report VI: Decent Work and the Informal Economy", in *International Labour Conference 90th session*. Geneva: ILO.
- [37] International Monetary Fund [2015]. "Current Challenges in Revenue Mobilization - Improving Tax Compliance," *IMF Policy Paper*.
- [38] Jensen, A., D. Lagakos [2019]. "Government Productivity in Tax Collection", *Working Paper*, Harvard Kennedy School.
- [39] Keen, M. [2012]. "Taxation and Development - Again," *IMF Working Paper Series*, 12/220.
- [40] Kopczuk, W., J. Slemrod [2006]. "Putting Firms into Optimal Tax Theory," *American Economic Review*, 96(2).
- [41] Kleven, H. [2014]. "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 28(4).
- [42] Kleven, H., M. Knudsen, C. Kreiner, S. Pedersen, E. Saez [2011]. "Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark", *Econometrica*, 79(3).
- [43] Kleven, H., C. Kreiner, E. Saez [2016]. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries", *Economica*, 83.
- [44] Kline, P., E. Moretti [2013]. "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority", *Quarterly Journal of Economics*, 129(1).
- [45] Kline, P. E. Moretti [2014]. "People, Places, and Public Policy: Welfare Economics of Local Economic Development Programs", *Annual Review of Economics*, 6.
- [46] La Porta, R., A. Shleifer [2014]. "Informality and Development", *Journal of Economic Perspectives*, 28(3).
- [47] La Porta., R., F. Lopez-de-Silanes, A. Shleifer [2008]. "The Economic Consequences of Legal Origins", *Journal of Economic Literature*, 46(2).

- [48] Lindert, P., J. Williamson [2016]. *Unequal Gains: A New History of American Growth and Inequality*. Princeton University Press.
- [49] Musgrave, R. [1981]. "Income Taxation of the Hard-to-Tax Groups", in R. Musgrave (editor), *Fiscal Reform in Bolivia*, Harvard University Press.
- [50] Naritomi, J. [2019]. "Consumers as Tax Auditors", *American Economic Review*, 10(9).
- [51] OECD [2006]. *Tax Administration in OECD and Selected Non-OECD Countries: Comparative Information Series*. Paris: OECD.
- [52] Peltzman, S. [1980]. "Growth of Government," *Journal of Law and Economics*, 23(2).
- [53] Penniman, C. [1980]. *State Income Taxation*. Johns Hopkins University Press.
- [54] Pinsky, D. [1972]. "State Constitutional Limitations on Public Industrial Financing: an Historical and Economic Approach", *U. of Pennsylvania Law Review*, 111(3).
- [55] Piketty, T., E. Saez [2013]. "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118(1).
- [56] Piketty, T., L. Yang, G. Zucman [2019]. "Capital Accumulation, Private Property, and Rising Inequality in China, 1978-2015," *American Economic Review*, 109(7).
- [57] Pomeranz, D. [2015]. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax", *American Economic Review*, 105(8).
- [58] Rollinson, M. [1976]. *Small Issue Industrial Development Bonds*. Capital Publishing Corporation.
- [59] Saez, E., S. Stantcheva. [2016]. "Generalized Social Welfare Weights for Optimal Tax Theory", *American Economic Review*, 106(1).
- [60] Sbragia, A. [1996]. *Debt Wish: Entrepreneurial Cities, U.S. Federalism, and Economic Development*. University of Pittsburgh Press.
- [61] Scheve, K., D. Stasavage [2010]. "The Conscription of Wealth: Mass Warfare and the Demand for Progressive Taxation," *International Organization*, 64.
- [62] Serrato, J.C., O. Zidar [2018]. "Structure of State Corporate Taxation and its Impacts on State Tax Revenues and Economic Activity", *Journal of Public Economics*, 167.
- [63] Troiano, U. [2017]. "Intergovernmental Cooperation and Tax Enforcement", *NBER Working Paper Series*, No. 24153.
- [64] Wilson, J. [1989]. "On the Optimal Tax Base for Commodity Taxation", *American Economic Review*, 79(5).
- [65] Yitzhaki, S. [1979]. "A Note on Optimal Taxation and Administrative Costs", *American Economic Review*, 475-480.

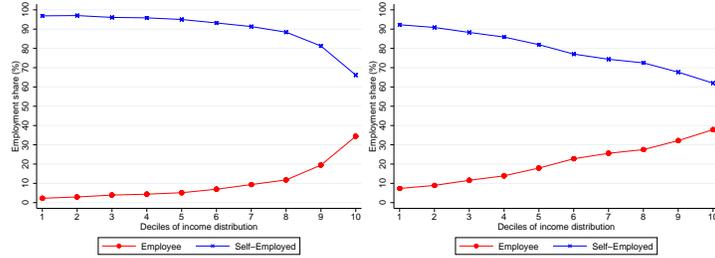
FIGURE 1: EMPLOYEE SHARE OVER INCOME DISTRIBUTION AND DECREASE IN INCOME TAX EXEMPTION THRESHOLD



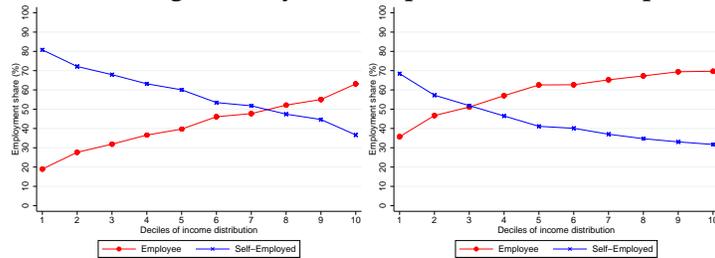
Notes: These graphs plot the employment-shares of employees and self-employed over deciles of the income-distribution, for different countries (Panel A) and within the US over time (Panel B). The share of each work-type is defined as the share of total employment in the decile of the income-distribution. Employees are defined as individuals working in jobs whose activity generate information trails that can be leveraged for income tax enforcement (Section 3.2). In each graph, the black solid denotes the location of the personal income tax (PIT) exemption threshold, which is the value of gross income above which a tax-filer becomes liable to pay income taxes (Section 3.2). Each graph is constructed from nationally representative micro surveys. Source: Section 3.2 and Appendix Section A.1.

FIGURE 2: EMPLOYEE SHARE: REPRESENTATIVE DEVELOPMENT PROFILES

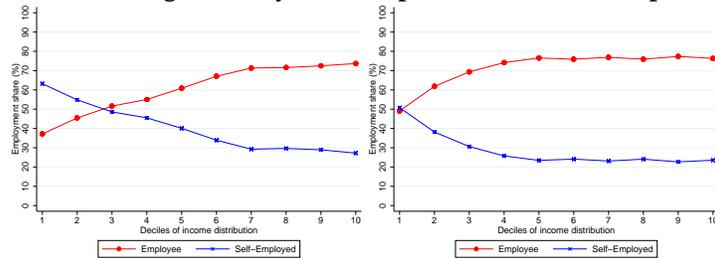
Profile for average country at \$1065 pc [LHS] and \$2226 pc [RHS]



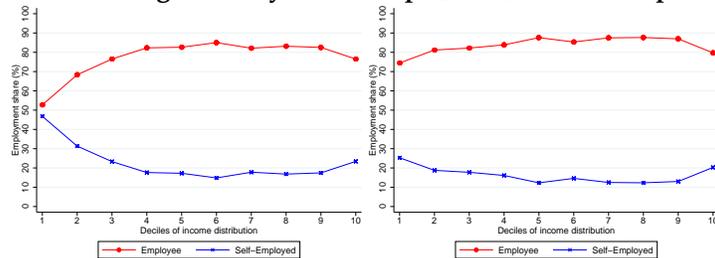
Profile for average country at \$3239 pc [LHS] and \$5796 pc [RHS]



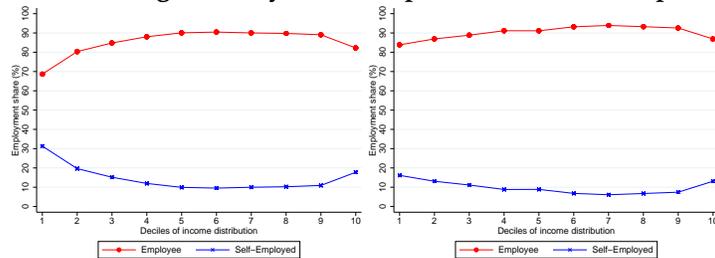
Profile for average country at \$8826 pc [LHS] and \$11257 pc [RHS]



Profile for average country at \$17141 pc [LHS] and \$27960 pc [RHS]

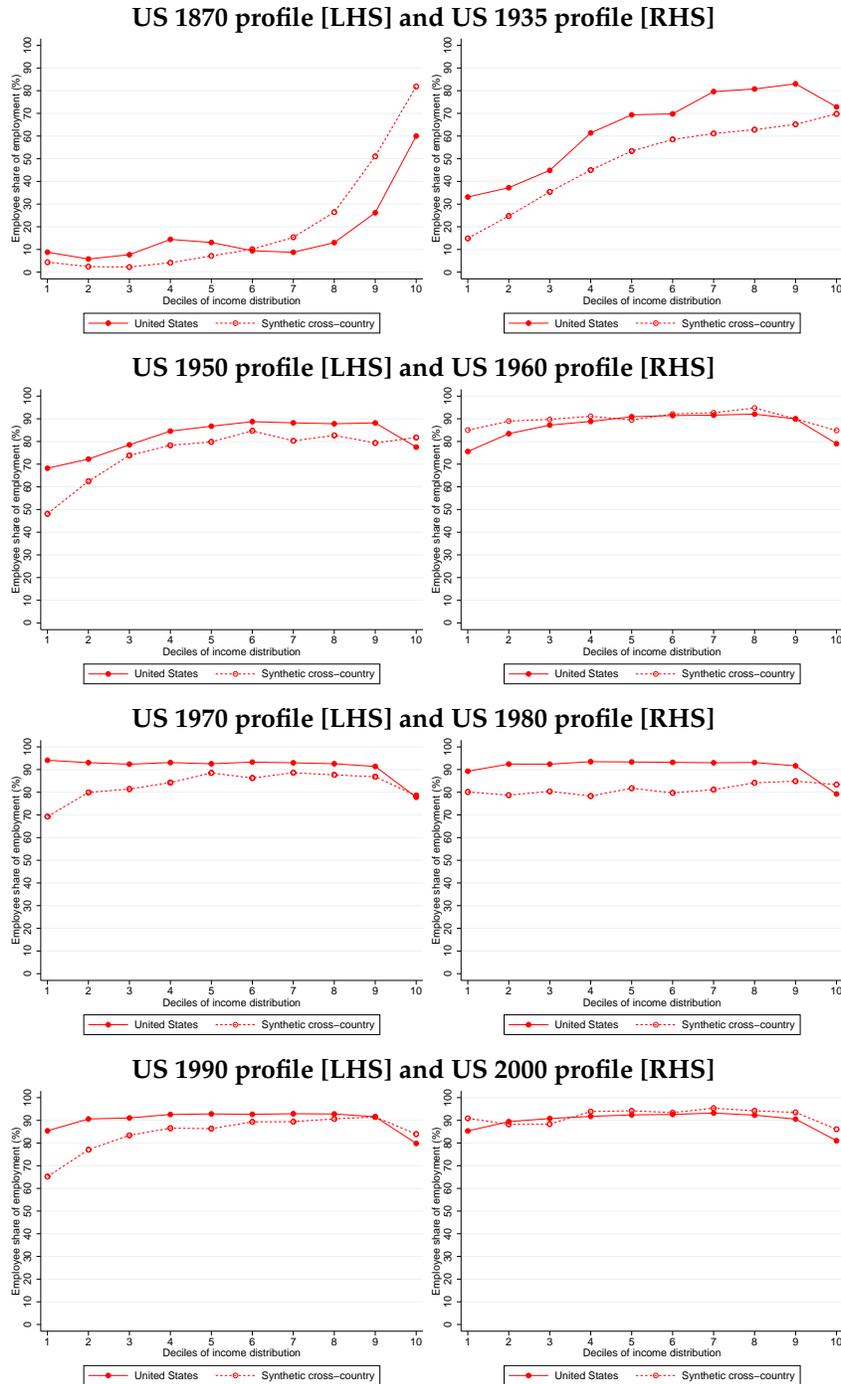


Profile for average country at \$38224 pc [LHS] and \$53878 pc [RHS]



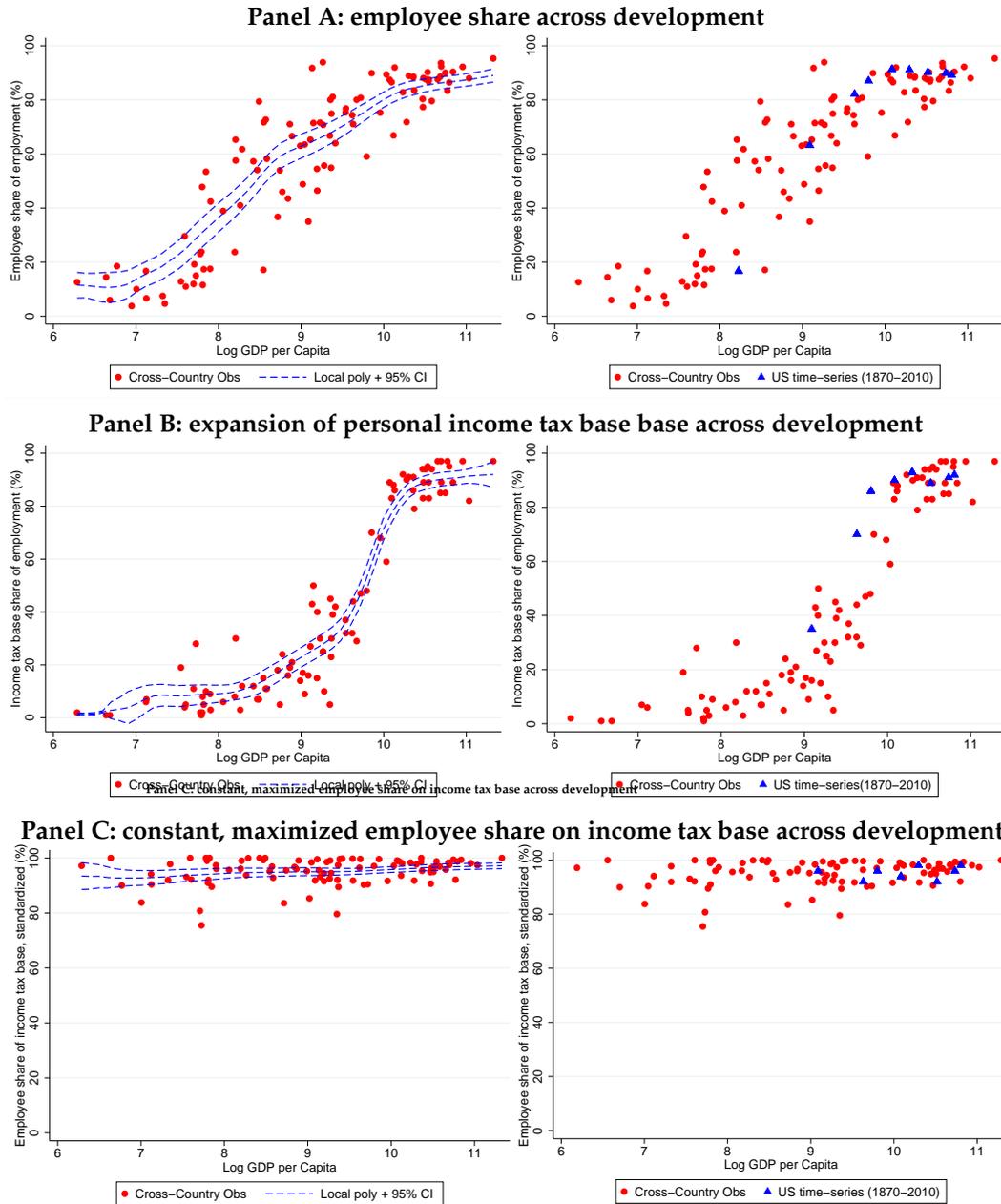
Notes: These figures plot the employment-shares of employees and self-employed over deciles of the income-distribution, for representative countries at different levels of per capita income. The share of each work-type is defined as the share of total employment in the decile of the income-distribution. Employees are defined as individuals working in jobs whose activity generates information trails that can be leveraged for income tax enforcement (Section 3.2). To construct this graph, I partition the cross-country sample into ten groups of ten countries, based on their level of per capita income. Within each group, I calculate the unweighted average of the employee-share and the self-employed share profiles across the 10 countries. I plot this average profile for every group, and indicate the average per capita income of the group. I use expenditure-side real GDP at chained PPPs in 2011 US\$ from the same year as the country-survey year. Source: Section 3.2 and Appendix Section A.1.

FIGURE 3: EMPLOYEE SHARE: WITHIN US AND PAIRED SYNTHETIC CROSS-COUNTRY PROFILES



Notes: These panels plot employee-shares over the income distribution in the US between 1870 and 2000, in comparison to employee-profiles based on current countries at similar levels of per capita income. In every graph, the solid red line indicates the employee-share of economically active employment across deciles of individual gross income in the US, in the indicated year. In every graph, the hollow and dashed line indicates the employee-share of economically active employment across deciles of individual gross income in a synthetic country based on the cross-country sample. This synthetic country is constructed as the average of the countries in the contemporaneous cross-country sample which have real per capita income that lies within 10% of the per capita income of the US in the historical year. As an example, the synthetic country for the US in 1870 includes contemporaneous India and Pakistan, while the synthetic country for the US in 1935 includes Indonesia and Peru. I use the expenditure-side real GDP at chained PPPs in 2011 US\$ from Penn World Tables to construct the synthetic country profiles. Source: Section 3.2-3.3 and Appendix Section A.1.

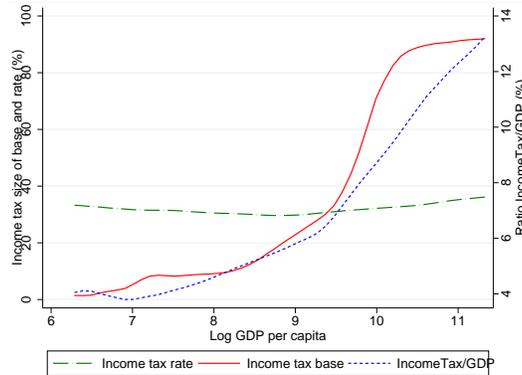
FIGURE 4: EMPLOYEE SHARE, SIZE AND EMPLOYMENT-COMPOSITION OF INCOME TAX BASE



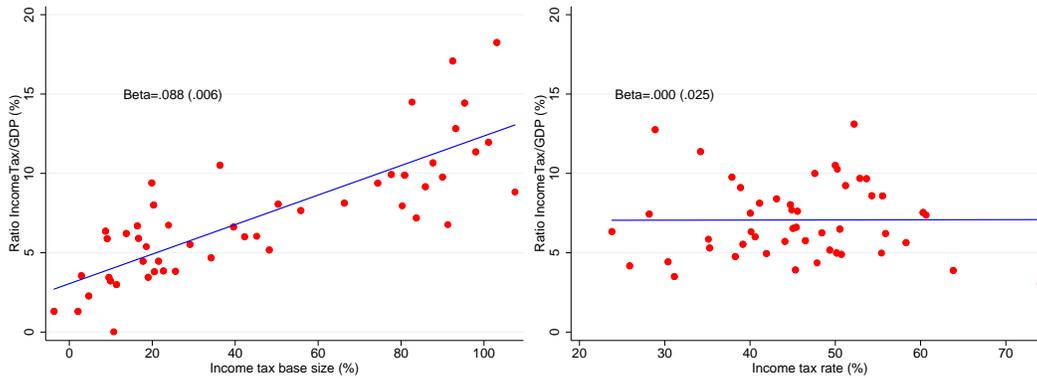
Notes: These figures document changes in employment and tax structure across development, using the 100 countries in the cross-country micro-database. In each panel, the left-hand scatter-plot is based on the full cross-country sample, while the right-hand scatter-plot overlays the US historical series onto the cross-country plot. Income per capita in the cross-country sample and in the US historical sample is measured with expenditure-side real GDP at chained PPPs in 2011 US\$ from Penn World Tables. In the left-hand plot, the dashed lines indicate a local polynomial fit with 95% confidence bands. Panel A plots the employee-share of employment against per capita income. Employee-jobs are defined as those which activity generate information trails that can be used for income tax enforcement. Employment is the total number of individuals in the economically active workforce, following the definition of the U.N. System of National Accounts. Panel B plots the size of the income tax base against per capita income. The size is defined as the sum of percentiles in the gross individual income distribution that lie above the income tax exemption threshold, and represents the share of the economically active workforce that is liable to pay income taxes. See Section 3.2 for more details on the measurement of employee-jobs and the exemption threshold. Panel C plots the employee-share composition of the tax base. This is the employee-share above the exemption threshold. It is expressed relative to the highest possible employee-share composition that can be obtained in the country. Source: Section 3.2 and Appendix Section A.1.

FIGURE 5: INCOME TAX BASE, RATE AND COLLECTION ACROSS DEVELOPMENT

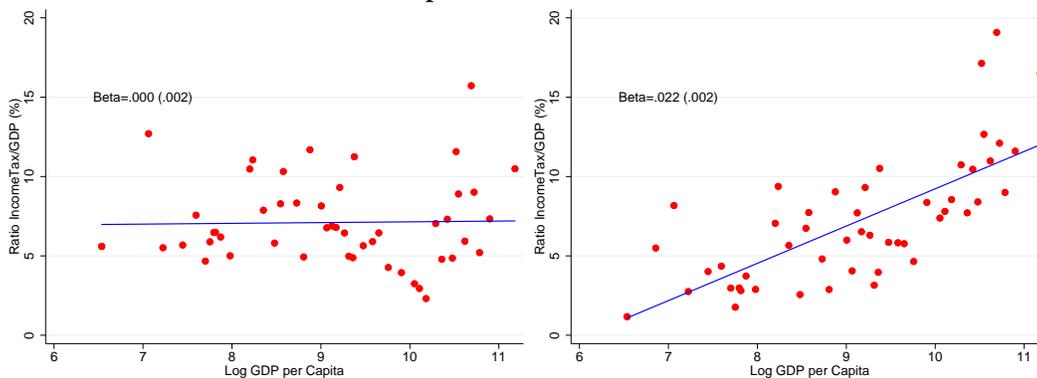
Panel A: PIT base, rate and collection across development



Panel B: conditional association between PIT collection and base [LHS], rate [RHS]



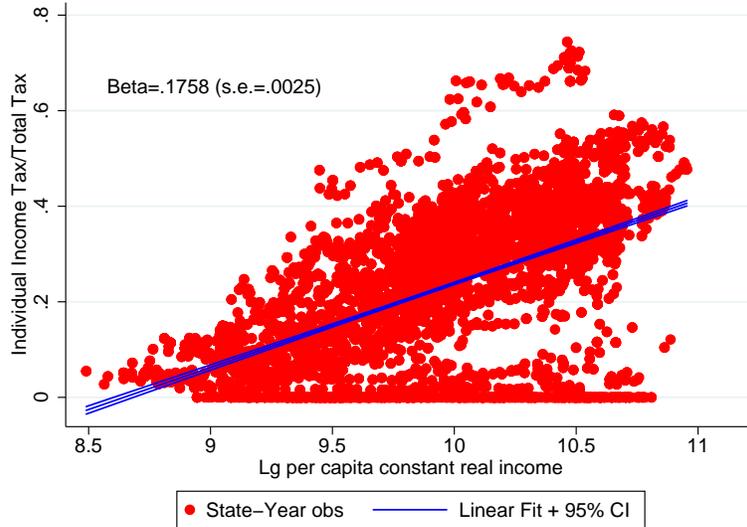
Panel C: collection across development, conditional on PIT base [LHS], rate [RHS]



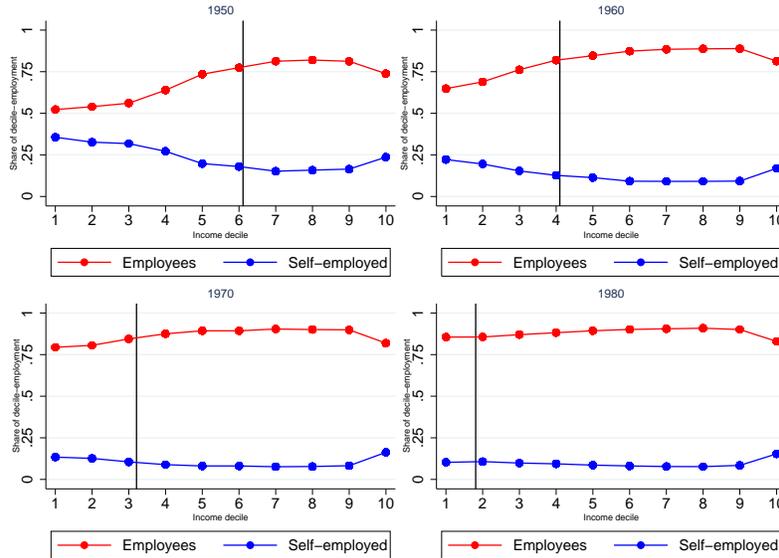
Notes: All graphs use the 100 countries contained in the cross-country micro data-base. Panel A plots the evolution of PIT collection, base, and rate across development. Each line corresponds to a local polynomial fit between log GDP per capita: income tax to GDP (red solid line); size of income tax base (blue short dash line); and, top marginal income tax rate (MTR, green long dash line). The size of the income tax base in a country is defined as the sum of percentiles in the gross individual income distribution that lie above the income tax exemption threshold, and represents the share of the economically active workforce that is liable to pay income taxes. Panel B plots the conditional association between income tax to GDP and income tax base (MTR), controlling for MTR (income tax base). Per example, to construct the right-hand side figure, I first regress both [income tax/GDP] on income tax base on MTR, and calculate residuals. I then group observations into fifty equal-sized (2 percentile-point) bins based on the tax base residuals, and scatter the means of [income tax/GDP] and tax base residuals within each bin, adding back the sample mean of each variable to ease interpretation. Panel C plots the association between residual income tax and development, after regressing income tax on PIT base (left-hand graph) or MTR (right-hand graph). In Panels B and C, the solid line shows the best linear fit, estimated on the underlying cross-country data. Source: Section 3.2. and Appendix Section A.1.

FIGURE 6: RISE OF INCOME TAX, TAX BASE EXPANSION AND EMPLOYEE SHARE: US STATES

Panel A: State income tax share of total taxes: all states 1939-2010

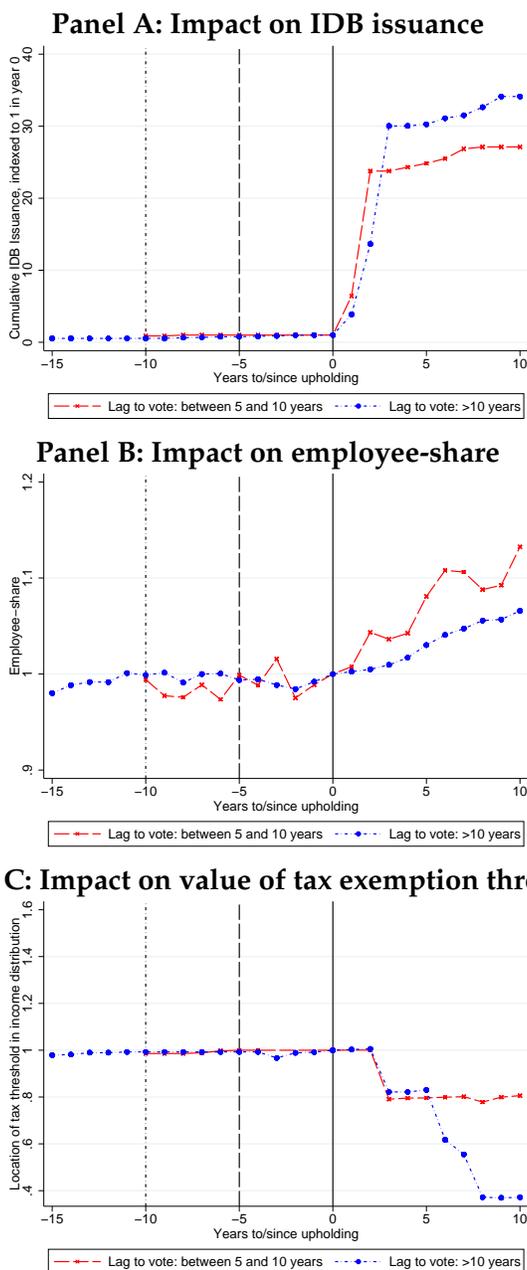


Panel B: employee share and state exemption threshold: average state 1950-1980



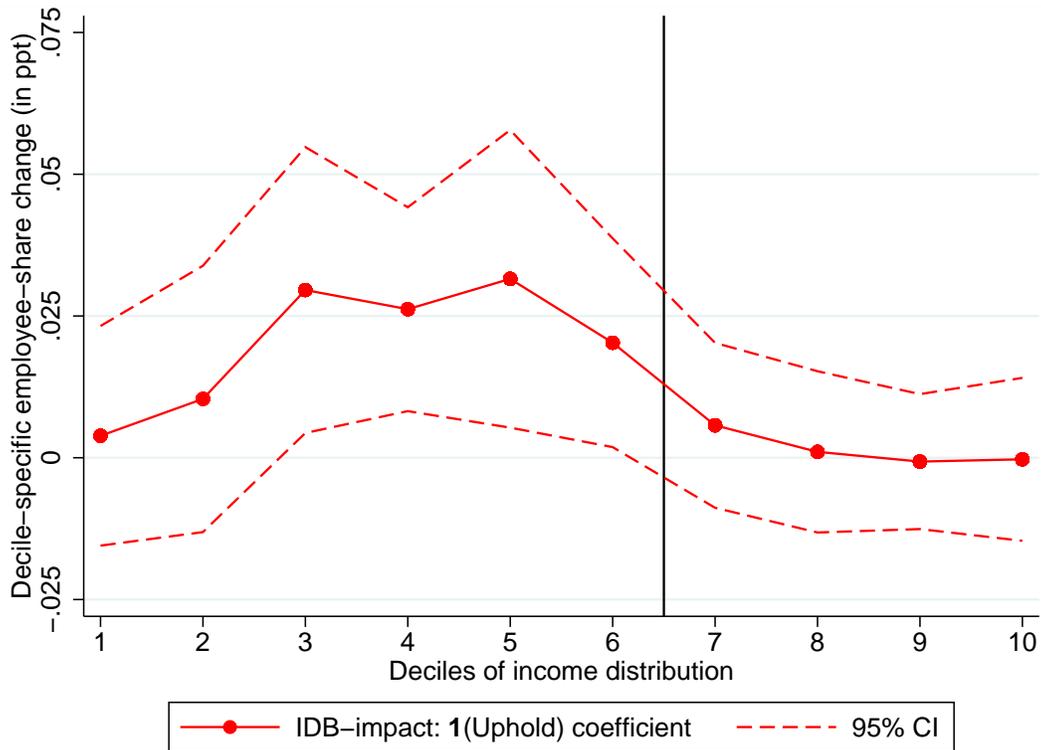
Notes: Panel A plots the state income tax share of total state taxes, using all state-year between 1939 and 2010, against log per capita income. State income tax is the sum of corporate income and personal income tax revenues. The tax revenue data is from the State Government Finance series, published by the Census. The real per capita income is constructed as the per capita income in a state-year from the historical BEA series, deflated by the historical CPI. The solid lines denote the linear fit with a 95% confidence interval from the regression on the full underlying state-year observations. Panel B plots the employment-shares of employees and self-employed over deciles of the gross income income distribution, for the average state in the US between 1950 and 1980. An employee-job is defined as a job which activity generates an information trail that is relevant for income tax enforcement. Employees are expressed as a share of the total economically active workforce. In every graph, the black solid line represents the location of the State income tax exemption threshold in the gross income distribution. This threshold represents the value of gross income above which single filer becomes liable to pay State income taxes, and is constructed using the state tax-calculator from Bakija (2009). For more details on the construction of the exemption threshold and employee-job, please see Section 4.1. Every graph in Panel B represents the average US continental State in the indicated year. Source: Section 4.1 and Appendix B.1.

FIGURE 7: IMPACT OF UPHOLDING EVENT ON OUTCOMES



Notes: These graphs provide graphical evidence on the impact of the Industrial Development Bonds (IDB) program. Every graph shows two series: the long-dash, square line indicates the group of States for which the lag between the vote and upholding event lasted between 5 and 10 years; the short-dash, circle line represents that group of States which the lag between vote and upholding lasted more than 10 years. The vote event is the year that the State legislature decided to vote on an IDB statute or provision, while the upholding event is the year that the voted IDB statute or provision was upheld as constitutional by the State supreme court. For more details on these events and the causes of the lag between them, please see Section 4.1. In every series, the variable of interest is indexed to 1 in the year of the upholding event, which is denoted with a solid vertical line. The distance between the vertical and the long-dashed (short-dashed) line demarks the intermediary time-period after the vote event but before the upholding event in the States for which the lag from vote to upholding was between 5 and 10 years (longer than 10 years). In Panel A, the outcome is the cumulative number of IDB issuances +1 (to avoid dividing by 0 in years with no cumulative issuances). In Panel B, the outcome is the employee-share of gross income. In Panel C, the outcome is the nominal value of the exemption threshold. Source: Section 4.1-4.2 and Appendix Section B.2.

FIGURE 8: DISTRIBUTIONAL IMPACTS OF UPHOLDING EVENT ON EMPLOYEE SHARE



Notes: This graph displays the impacts of the IDB upholding event on the distribution of employee-share across deciles of the State income distribution. The graph plots coefficients $\hat{\theta}_j$ on the 1(Uphold) indicator in a regression on employee-share in decile $j = 1, \dots, 10$, using specification 1. Each circle denotes the decile- j point estimate $\hat{\theta}_j$. The dashed lines denote the 95% confidence interval of the point-estimate, where standard errors are clustered at the state level. The black solid line denotes the average location of income tax exemption threshold in the IDB States, calculated in the year prior to the vote event. Source: Section 4.4 and Appendix Section B.1.

TABLE 1: DETERMINANTS OF CHANGE TO IDB LITIGATION-STATUS

LHS=1(Upheld)								
Panel A: Politics								
1(Civil law)	5.435 (3.201)***	12.781 (5.212)***	12.117 (4.592)***	9.789 (4.815)***	12.558 (5.489)***	10.291 (3.884)***	12.956 (5.031)***	12.320 (4.921)***
1(Hist rail default)		.337 (.105)***	.567 (.334)	.378 (.120)***	.340 (.116)***	.671 (.449)	.336 (.108)***	.428 (.215)*
Covariate	-	-	SupremeC appointed?	Election dummies	Voting restrictions	Population total	Neighbor uphold?	RightWork laws?
Signif 10 pct?	-	-	N	N	N	N	N	N
Panel B: Economics								
1(Civil law)	11.020 (6.981)***	10.266 (3.905)***	11.029 (4.182)***	12.803 (5.134)***	13.102 (6.031)***	10.547 (4.370)***	12.776 (5.221)***	11.461 (5.988)***
1(Hist rails)	.358 (.092)***	.677 (.455)	.571 (.350)	.332 (.148)**	.294 (.078)***	.461 (.208)*	.338 (.121)***	.358 (.115)***
Covariate	Redev counties	LabForce growth	Manuf growth	Employee share	Income per cap	PIT base	PIT take	TotT take
Signif 10 pct?	N	N	N	N	N	N	N	N

Notes: This table reports the results of estimating non-parametric Cox proportional hazard models, where hazard rates are reported. The unit of observation is a state-year. A State enters the sample in the year where the State legislature votes in a statute or provision authorizing the IDB program. The State drops from the sample once the highest instance of the State court system upholds the constitutionality of the voted statute or provision in a leading court case. Each column in each panel reports the coefficients on a dummy for civil law origins, a dummy for historical rail defaults, as well as a description of the included covariate and whether the included covariate is significant at the 10 percent level. The test for significance relates to the null that the coefficient is equal to one. Panel A includes political covariates that could determine a change in litigation status: whether State supreme court members are appointed or elected; whether it is a State election year; whether voter restrictions are in place; the total population size; whether a geographically adjacent State has upheld the IDB statute; and, whether right-to-work laws are in place. Panel B includes economic potential determinants: the share of counties in a State that are classified as “redevelopment areas” and eligible for local development financing according to Federal criteria; the yearly labor force growth; the yearly manufacturing workforce growth; the employee-share of employment; the income per capita; the size of the State income tax base; the ratio of State income taxes to GDP; the ratio of State total taxes to GDP. *, **, *** denote significance at the 10 percent, 5 percent, 1 percent level. Standard errors robust to clustering at the state level. Sources: Section 4.1-4.2 and Appendix Section B.2.

TABLE 2: EFFECTS OF IDB PROGRAM ON EMPLOYMENT AND INCOME TAX OUTCOMES

	(E-share)		(K/y)		(E-share above K)		Log(PIT/GDP)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Vote)	.003 (.005)	.002 (.005)	-.357 (.232)	-.343 (.258)	-.004 (.010)	-.003 (.010)	.049 (.057)
1(Vote) x 1(EoI)		.009 (.011)		.401 (.392)		.000 (.017)	-.065 (.081)
1(Uphold)	.017 (.005)***	.012 (.007)**	-.639 (.278)**	-.804 (.335)**	-.007 (.012)	.000 (.013)	.162 (.079)**
1(Uphold)x1(EoI)		.016 (.011)		.759 (.319)**		-.008 (.019)	-.256 (.165)
F-test: 1(Uphold) + 1(Uphold)x1(EoI) (p-value)		7.72 (.003)		0.02 (.895)		0.16 (.696)	0.60 (.446)
Mean outcome variable	0.771	0.771	7.084	7.084	0.798	0.789	.972
State FE	x	x	x	x	x	x	x
Year FE	x	x	x	x	x	x	x
State-year controls	x	x	x	x	x	x	x
States	28	28	28	28	28	28	28
State-year Obs	466	466	466	466	466	466	466

Notes: This table reports results from estimating the following regression

$$y_{st} = \beta + \pi \mathbf{1}(\text{EoI})_{st} + \alpha \mathbf{1}(\text{Vote-in})_{st} + \phi \mathbf{1}(\text{Vote-in})_{st} \times \mathbf{1}(\text{EoI})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \beta \mathbf{1}(\text{Upheld})_{st} \times \mathbf{1}(\text{EoI})_{st} + \lambda \mathbf{X}_{st} + \mu_s + \gamma_t + \varepsilon_{st}$$

where s denotes state and t denotes time. $\mathbf{1}(\text{Vote-in})_{st}$ indicates whether a vote has occurred in the state House to allow issuance of IDB but the IDB has not yet been upheld, $\mathbf{1}(\text{Upheld})_{st}$ indicates whether the State court system has upheld the constitutionality of the voted IDB statute or provision. The vote-in and upholding events are mutually exclusive. $\mathbf{1}(\text{EoI})_{st}$ is an indicator variable taking a value of 1 when a State has passed an exchange of information agreement with the Federal Internal Revenue Service. In Columns 1-2, the outcome variable is the employee-share of the economically active workforce. In Columns 3-4, the outcome variable is the ratio of the State income tax exemption threshold, K , to State per capita income, y . K represents the value of gross income above which an individual filer becomes liable to pay State income taxes. In Columns 5-6, the outcome variable is the employee-share of employment with income above the exemption threshold. In Column 7, the outcome variable is (log) ratio of State individual income taxes to GDP. In the middle of the table is reported the mean of the outcome variable in the period before the vote event. The state-year controls, \mathbf{X}_{st} , are election year dummies and the first stage political competition and policy instruments from Besley et al. (2010). *, **, *** denote significance at the 10%, 5%, 1% level. Robust standard errors clustered at the state level in parentheses. Source: Section 4.2 and Appendix Section B.1.

TABLE 3: EFFECTS OF IDB PROGRAM ON CONFOUNDING OUTCOMES

Panel A	Log(CorpIncTax/GDP) (1)	Log(SalesTax/GDP) (2)	Log(PropertyTax/GDP) (3)	PIT MTR (4)	1(PIT Withholding) (5)	Tax Administration (6)
1(Vote)	-.012 (.058)	.068 (.052)	-.041 (.027)	-.001 (.001)	.057 (.090)	-.015 (.052)
1(Vote) x 1(EoI)	-.095 (.086)	-.003 (.138)	.019 (.033)	.005 (.002)*	.331 (.209)	.088 (.069)
1(Uphold)	-.055 (.135)	-.047 (.049)	-.031 (.049)	-.001 (.002)	.018 (.154)	-.029 (.080)
1(Uphold)x1(EoI)	-.075 (.109)	.112 (.158)	-.039 (.078)	.004 (.003)	.285 (.229)	.087 (.093)
Mean outcome variable	.991	2.521	.860	.129	.014	2.492
State FE, Year FE, Controls	x	x	x	x	x	x
States	28	28	28	28	28	28
State-year Obs	466	466	466	466	466	466

Panel B	Income per Capita (1)	Top 1 percent income share (2)	Max Unemp Benefits (3)	1(Right to Work Laws) (4)	Political Competition (5)	Democratic Vote Share (6)
1(Vote)	26.465 (23.404)	.234 (.190)	-19.481 (21.058)	.008 (.010)	.003 (.011)	-.024 (.011)*
1(Vote) x 1(EoI)	5.444 (54.439)	-.419 (.213)*	9.114 (22.076)	-.003 (.007)	-.031 (.023)	.011 (.024)
1(Uphold)	-8.318 (35.369)	.051 (.223)	-10.707 (18.734)	.050 (.046)	-.004 (.014)	-.017 (.017)
1(Uphold)x1(EoI)	40.544 (55.482)	-.285 (.286)	5.254 (27.838)	-.040 (.038)	-.022 (.022)	-.013 (.023)
Mean outcome variable	7641.297	13.433	307.366	.234	-.112	.539
State FE, Year FE, Controls	x	x	x	x	x	x
States	28	28	28	28	28	28
State-year Obs	466	466	466	466	466	466

Notes: This table reports results from estimating the same regression as in Table 2. In Panel A, the outcome variables are ratios of different tax revenue sources to GDP; the marginal state income tax rate; an indicator for existence of state income tax withholding; the number of State tax departments. In Panel B, the outcome variables are: income per capita; the top 1% share of gross income; maximum State weekly unemployment benefits; an indicator for right-to-work laws; the measure of political competition from Besley et al. (2010); and, the Democratic vote-share across all State-specific elections. *, **, *** denote significance at the 10%, 5%, 1% level. Robust standard errors clustered at the state level in parentheses. Source: 4.2 and Appendix Section B.1.

Online Appendix for 'Employment Structure and the Rise of the Modern Tax System'

Anders Jensen *

October 2019

*Harvard Kennedy School and NBER. Contact: anders_jensen@hks.harvard.edu.

A Cross-development material

A.1 Data sources and construction of variables

In this sub-section, I provide additional details on the novel micro data-base used in Section 3. I first outline the underlying data and construction of variables in the cross-section of countries. I then outline the data and variables construction used in the historical US time-series.

A.1.1 Cross-country: sources and methodology

The cross-country database contains micro-data collected from 100 countries around the world, to document changes in employment structure transformation in as many incremental stages over development as possible. I chose to focus on countries with at least 1 million citizens. The selection of a survey in a particular country had to satisfy three criteria. First, it must be nationally representative. Second, it must survey respondents in all forms of work arrangement as opposed to, per example, only salaried workers. Third, it must contain continuous information on all sources of income, instead of, say, only wage earnings.

Given these criteria, the preferred type is a living conditions survey. This type of survey will often dominate a labor force survey, for three reasons. First, the living conditions survey usually contains information on a broader range of income sources which, especially in the context of less-developed countries, can be quite important in order to construct the lower deciles of the country's income distribution. Second, it is not always clear what the underlying sample design is for the labor force survey, and it could potentially omit individuals which in the context of this study should be included in the survey, such as casual wage day laborers and household family workers; on the other hand, the scope of a living conditions survey is usually to assess the conditions of a nationally representative sample of individuals, which should include all the alternative work type patterns. Third, the sample size of a living condition survey is typically larger than that for a labor force survey, which does not have to imply better quality of data, but usually is due to sampling design which attempts to survey all geographical areas in the country. Basic health and demographics surveys are discarded, because they do not contain information on work arrangements and income.

The data collection effort resulted in 100 surveys, which are detailed in Table A.1, displaying for each country: the year of the survey; the per capita income group; the survey type; the coverage; the sample size; and, the original source. The income group corresponds to the World Bank classification of the country in the year of the survey. The micro database covers all levels of development: 20% of surveys from low-income countries; 28%

from lower-middle income countries; 21% from upper-middle income countries; and, 31% from high-income countries. 93 out of the 100 data-sets are living condition surveys, 5 are labor force surveys, and the remaining 2 are censuses. In low and lower-middle countries, I obtain almost all surveys directly from the national statistics office, or the relevant government agency. In these countries, the average sample size is substantially larger than the corresponding Living Standards and Measurement Survey (LSMS) from the same country.

The construction of the employee variable is based on questions similar to the 'class of worker' question in the US Census. All cross-country surveys were chosen to ensure the highest possible international comparability. Two features in particular serve that purpose, and are common across all surveys. The first feature is the high level of detail in the categories of the 'worker-class' question. In all surveys, I can therefore distinguish between employees and employers. This removes the possibility that employers of large firms are counted as employees, in which case the comparison of employee versus self-employed would partially be confounded by a firm size comparison. In addition, I can systematically distinguish between employees and both family and non-family workers in household enterprises. I can also systematically distinguish between employees that work for a salary versus employees that work for in-kind payments. Finally, and related to the previous point, I can distinguish between casual daily wage laborers and 'regular' employees in the countries where seasonal work is arguably most prevalent. It is true, however, that I cannot systematically distinguish casual wage laborers, and non-regular wage earners more generally, from contract-based regular employees. Taken together, this discussion implies that, with the exception of daily wage laborers, I can construct employee and self-employment categories in a consistent and internationally comparable manner across all countries. The second advantageous feature of all surveys is that my definition of employee versus self-employed is systematically based on an 'objective' worker-class question. In contrast, certain surveys allow respondents to choose 'informal sector' in response to the worker-type question. As discussed in the main text, my employee classification is closely related to the ILO concept of formality. Nonetheless, the specific definition of formality embedded in surveys is likely to vary across countries in ways that are hard to measure, and relying on such responses would reduce the transparency of comparisons across countries. As such, I discard all surveys where I cannot construct the employee classification based on a detailed and objective 'class of worker' question.

I focus on calculating gross income from all sources in order to be conceptually consistent with the broadest possible income-definition in the tax code. This leads me to calculate four sources of income: wage income, self-employment income, capital income,

and miscellaneous income (such as lottery receipts). Most importantly, I ensure that I can calculate both employee and self-employment income with precision. In this context, the most significant challenge is to calculate self-employment income in agriculture in less-developed countries. Agricultural earned revenue includes the value of crops sold to others. I do not attempt to create a monetary value of in-kind sales, as offering and receipt of in-kind goods and services is not subject to tax. Agricultural capital revenue includes the sale of live-stock, income from rental of equipment, and share-cropping income. From this revenue I attempt to subtract costs, which include expenditure on inputs, wages paid out to workers, and new investments. In a limited number of countries, I do not observe any agricultural revenue for respondents that are self-employed in agriculture. These are most often contributing family workers on farms where the full output is consumed by the family. In this limited number of cases, I construct the income as the market value of the own-consumed output, as estimated by the respondent. In all surveys, I exclude two sources: social transfers, and in-kind goods and services. I exclude social transfers because it falls outside the concept of taxable income. The monetary value of in-kind goods and services are sometimes included in taxable income, often on a presumptive basis. However, apart from the mentioned case above, I exclude this source of income because I cannot measure it consistently across all surveys. Non-monetary income is often more important for less wealthy individuals, and is more prevalent in less developed countries. In the surveys where there exists systematic data on the monetary value of non-monetary income, I can confirm that the inclusion of these sources of income does not change the distributional employee-profile. That is because these sources of non-monetary income are too small in magnitude to overturn the decile-ranking of individual income.

In 7 countries, I cannot calculate gross individual income with precision. These countries are: Democratic Republic of Congo, Liberia, Ethiopia, Malawi, Mali, Burkina Faso, Cambodia. In the case of DRC, Liberia, Ethiopia, and Malawi, I do not comprehensively observe either agriculture sales or costs, so I cannot calculate agricultural self-employment income. In Mali, Burkina Faso, and Cambodia, I do not comprehensively observe costs of non-agriculture own-account workers, so I cannot calculate non-agricultural self-employment income. In these 7 cases, which are among the poorest in the micro-database, I instead calculate total individual expenditure, and use it as a proxy for total income. There exists a set of low-income countries in which I have both good income and expenditure data. In results not reported, I can confirm that the employee-share profiles are very similar when using either income or expenditure to calculate the x-axis distribution. As mapping expenditure into income is difficult, I do not attempt to locate the income tax exemption threshold in these 7 countries.

While I define the employee-status based on the respondent's primary job activity, I at-

tempt to calculate income from all activities reported during the reference period. The main issue that arises in this context is the allocation of income which is reported at the household, rather than individual, level. For sources of earned income that are not at the individual level, I assign equal portions of this income to each economically active member of the household that reports having undertaken this activity during the reference period. Per example, the value of sold crops will be distributed equally among all household members that report having contributed to the family farm, either as a first or secondary activity. For sources of non-earned income reported at the household level, I assign an equal portion to each economically active member, such as in the case of property rental income.

Whenever a country's tax code is based on annual amounts and the reference period in the country's survey module is not, I construct the annual income distribution. I multiply the regular amount by the number of periods in the year – e.g. if wage income was reported monthly, I multiply it by the number of months that the wage income is reported to have been received during the past year. In the case where no periodicity exists, I assume that the flow was occurring during the whole year with the same pattern as during the reference period.

In every country survey, I limit the sample to the economically active population, following the definition of employment from the U.N. System of National Accounts. This definition is also used in Bicks, Fuchs-Schundeln, & Lagakos (2018), and in Feng, Lagakos, & Rauch (2018), which study respectively how hours worked and unemployment vary with development. I code employment-type based on the primary job in the reference period. The primary job is often explicitly defined as the job in which the respondent spent most hours during the reference period. The reference period in the Luxembourg Income Study (LIS) is annual, while it is predominantly monthly in the remaining surveys. The extent to which the periodicity and the focus on the primary job introduce biases in the representativeness of my employment-categories is discussed in Section A.6.

In addition to income and employment-categories, the micro-database also contains variables on education, sector, and geographical location. The geographical location measures whether a respondent lives in an urban area or not. I do not attempt to harmonize this variable, and use the urban definition in the surveys, which may therefore vary from country to country. I use variables to indicate three levels of education completion: not completed primary; completed primary but not high school; completed high school. I chose to not distinguish further levels of education, in order to maximize the number of surveys where I could create consistent measures. Finally, I code the sector of the primary job. The aim was to create a set of sectoral categories which are consistent with the ISIC classification. I create four sectoral categories: agriculture; manufacturing ; services; and,

public administration. I define these four categories in relation to the divisions of the ISIC 4.4 classification, where: agriculture contains Section A; manufacturing and construction contains Sections B to F; services contains Sections G to M, and S to U; and, public administration and education contains Sections N to R. As such, the manufacturing sector also contains mining and construction; the services sector also contains wholesale and retail trade, transportation, IT, finance, and activities of household enterprises; and, public administration also contains education, social work, and entertainment. Most of the industry codes in the surveys do not contain a pre-existing ISIC classification. To the best extent possible, I therefore first map the survey-categories to ISIC divisions, and then to my 4 sectoral categories. I do not include the sectoral variable if the survey has data only on a subset of the categories - per example, if a survey records that a job is not in the agricultural sector, but does not specify which non-agricultural sector it belongs to. These variables are used in the regression analysis in Section A.5.

A.1.2 Historical US time-series: sources and methodology

The historical federal profiles in the US between 1950 and 2010 were constructed using the decennial Census samples, extracted from the IPUMS USA database. I exclude all respondents that are not active in the labor force during the reference period. I calculate the individual income distribution, based on the measure of gross income at the individual level. To construct the income distribution, I use the measure of total, pre-tax, personal income. Farm and non-farm business income, as well as wage income, are consistently recorded in every Census sample. I use the detailed 'class of worker' question, which allows me to assign unpaid family workers to the self-employed category. Consequently, the self-employed category includes employers, own account workers, self employed that are not incorporated, and self-employed that are incorporated. Given the resemblance with the categories contained in the cross-country surveys, there is strong comparability between these US historical profiles and the cross-country profiles constructed in Figure 3. I apply individual weights to estimate the employee-share of every decile of the income distribution in every decade.

Before 1950, the decennial Census does not report total personal income at the individual level. The 1940 1 percent sample does contain wage and salary income, but no business income nor farm income, which are required to construct a personal gross income distribution. Instead, I use the 1935-36 Study of Consumer Purchases. The scope of the study was to "ascertain for the first time in a single national survey the earning and spending habits of inhabitants of large and small cities, villages, and farms" (ICPSR Study 8908, 2009). The survey was the result of a joint effort by the Bureau of Labor Statistics and the Bureau of Home Economics of the Department of Agriculture, and is meant to have

been the sampling-methodology predecessor for the income-component in Census. The survey contains both a labor force component, where respondents gave information on income and housing, and for a subset of the total sample, a living conditions component where respondents gave additional information on expenditure. The primary sampling units were chosen to represent “the demographic, regional, and economic characteristics of the United States” (ICPSR, 2009). From these areas, a randomly selected group of approximately 700,000 families were screened in a first wave. From this first wave, 300,000 families were chosen to supply basic income and housing info, and a subset of 61,000 families were selected to provide additional expenditure information. It is important to understand the selection criteria into the different waves. The ICPSR accompanying documentation explains that in order to be selected out of the first wave, the requirements were: “families include at least two members, with husband and wife married for at least one year, and with no more than the equivalent of ten boarders for the survey year (...) farm families had to live in a setting that met the Census definition of a farm; the family itself must operate the farm (or in the southeast, be a sharecropper) and have conducted farming activities for at least one year” (ICPSR Codebook, 2009). Families were admitted to the first wave “without restriction in terms of occupation, income, employment status, or whether they were drawing or had drawn relief during the year.” Selection into the second-wave where the survey included expenditure components, was based on the following criteria: “non-farm families must have had at least one wage earner in a clerical, professional, or business occupation. A minimum income for the survey year of \$500 was required in the largest cities and \$250 in the smaller cities and rural areas (...) Families that had received relief were excluded from this third wave.” These criteria produce a highly selected sample for the second-wave respondents, and hence I base the analysis on the sample of first-wave respondents.

The ICPSR data-sample that I use for the 1935 Federal profile is based on a random sub-sample of approximately 5,000 families who only completed the first-wave ‘labor force’ component of the survey. The ICPSR sub-sample was created in the following way: “a sampling fraction of 1 schedule for entry for every 83 schedules counted was chosen” from the urban sample, creating 3200 schedules from the larger urban areas and 1800 schedules from the more rural areas”; the ICPSR sample consists of schedules “spread across both the rural and urban portions of the original investigation.” The employee classification is based on ‘status of employment’ question, which is identical to the (non-detailed) ‘class of worker’ question used in all US Censuses from 1950 onward. I code as an employee any individual respondent who reports being a “salaried worker/wage earner.” I code as self-employed any respondent who reports being “self-employed”, and any respondent who does not specify a type of work but declares to be working, is

above age 20 and who has substantial work-related income. I exclude all respondents that are employed on work-relief projects in their primary job. As such, the sample closely resembles the economically active workforce definition used in the cross-country sample. Total gross income only exists at the household level. Rather than try to assign income at the individual level within the household, I focus on the work-type of the head of household. I then rank individuals based on the reported total income, and estimate the employee-share in each income decile.

The 1935-36 survey marked a clear shift in focus of the surveys conducted by the Bureau of Labor Statistics. Indeed, the surveys carried out prior to the 1930s focused on measuring family income and expenditure patterns of the U.S. employed workers and their families. Consequently, the available surveys, including the “Cost of living in the United States, 1917-1919” (ICPRS 7711, 1986) and the “Cost of living of industrial workers in the United States and Europe, 1888-1890” (Haines, 2006) contain data from families of wage earners or salaried workers in industrial locales scattered throughout the U.S. In order to construct a historical profile before the 1930s, I use data from Lindert & Williamson (2016), which studies incomes in the U.S. between 1650 and 1870.

Unlike previous work which approaches the measurement of income during this historical period from the production-side or the expenditure-side, Lindert & Williamson build estimates of income based on personal income records, assembling nominal earnings from free labor and property income. The approach to estimating income in Lindert & Williamson derives from combining information about income and labor force participation counts across occupation-space-time. This amounts to building ‘social tables’ across occupations within a given space-time frame, and the approach is conceptually similar to social accounting matrices that were used in development economics in the 1970s and 1980s. The authors provide a significant effort to capture all occupation categories in a given space-time. They draw on data from local tax assessments and occupational directories for ‘registered’ occupations, and local censuses for ‘unregistered occupations’. These same data sources usually provide counts of the total number of individuals across the different occupations. The authors combine previous work with new estimates from local sources to derive personal earned income across occupation-space-time. In some instances, the occupation-space-time income reported was not at the annual level, and the authors bring the estimates to such level by making assumptions on the full-time number of hours spent (the assumptions are discussed in Lindert & Williamson, 2016). The authors also collect data on property income by assuming rates of return on wealth estimates that vary across occupation-space-time, and combine this with earned income to derive measures of total income.

I construct a historical 1870 profile based on the data kindly provided by Peter Lindert.

This cross-section builds upon the 1870 1 percent US Census sample delivered to the authors by IPUMS USA, which included sampling weights at the individual-level. The 1 percent sample contains space-occupation counts, which are then merged with the authors' estimate of total income at the same level. I extend their analysis and classify all available occupation categories as either self-employed or employee. I use the detailed description of each occupation category to code employment-type. Per example, all occupations where a reference is made to 'manager' are coded as employee cells. The enumerator instructions for the sample design are particularly useful for my exercise in that they highlight very clearly the need to distinguish between self-employed and employee status: "Do not call a man a 'shoemaker', 'bootmaker', unless he makes the entire boot or shoe in a small shop. If he works in a boot and shoe factory, say so (...) Cooks, waiters, etc., in hotels and restaurants will be reported separately from domestic servants." The occupation category only exists for the head of household. The measure of total income includes own labor earnings in agriculture and non-agriculture, farm and non-farm operating income, and property income. This is a comprehensive measure of gross income before taxes and transfers which is not identical to, but closely resembles, the measure used in the more recent Federal US and cross-country samples. I apply the sampling weights initially provided by IPUMS USA. I estimate the employee-share in every decile of the individual gross income distribution, for the population that is active in the labor force.

In all the profiles, I locate the Federal income tax exemption threshold in the income distribution. Note that there was no Federal income tax in 1870. In all profiles from 1935 onward, I use the historical IRS series which provide details on the nominal value of the standard deduction of a single filer.

Table A.1: Cross-Country Data Sources

Country	Year	Per Capita Income Group	Survey type	Coverage	Sample Size	Original source
Albania	2009	Upper Middle	Labor Force	National	18,997	National Institute of Statistics
Argentina	2009	Upper Middle	Living Conditions	Urban	47,862	National Institute of Statistics and Census
Australia	2014	High	Living Conditions	National	16,801	Luxembourg Income Study
Austria	2013	High	Living Conditions	National	5,102	Luxembourg Income Study (LIS)
Azerbaijan	1995	Low	Living Conditions	National	8,901	Living Standards Measurement Study (LSMS)
Bangladesh	2010	Low	Living Conditions	National	19,664	Bangladesh Bureau of Statistics
Belgium	2000	High	Living Conditions	National	2823	Luxembourg Income Study (LIS)
Belize	1999	Lower Middle	Labor Force	National	15,167	Central Statistical Office
Bolivia	2007	Lower Middle	Living Conditions	National	16,130	National Institute of Statistics
Brazil	2009	Upper Middle	Living Conditions	National	191,810	National Institute of Geographics and Statistics
Bulgaria	2007	Upper Middle	Living Conditions	National	6,941	National Institute of Statistics
Burkina Faso	2014	Low	Living Conditions	National	32,023	National Institute of Statistics and Demographics
Cambodia	2009	Low	Living Conditions	National	31,959	Ministry of Planning
Cameroon	2007	Lower Middle	Living Conditions	National	51,836	National Institute of Statistics
Canada	2013	High	Living Conditions	National	27,344	Luxembourg Income Study (LIS)
Chile	2009	Upper Middle	Living Conditions	National	90,610	Social Observatory, University Alberto Hurado
China	2013	Upper Middle	Living Conditions	National	14,782	Luxembourg Income Study (LIS)
Colombia	2009	Upper Middle	Living Conditions	National	170,220	National Directory of Statistics
Costa Rica	2009	Upper Middle	Living Conditions	National	19,594	National Institute of Statistics and Census
Czech Republic	2013	High	Living Conditions	National	7,653	Luxembourg Income Study (LIS)
Cote d'Ivoire	2008	Lower Middle	Living Conditions	National	59,699	National Institute of Statistics
Dem. Rep. of the Congo	2004	Low	Living Conditions	National	72,685	National Institute of Statistics
Denmark	2013	High	Living Conditions	National	88,696	Luxembourg Income Study (LIS)
Dominican Republic	2009	Upper Middle	Living Conditions	National	30,430	National Statistics Office
Ecuador	2009	Lower Middle	Living Conditions	National	78,865	National Institute of Statistics and Censuses
Egypt	2010	Lower Middle	Living Conditions	National	34,069	Economic Research Forum (ERF)
El Salvador	2014	Lower Middle	Living Conditions	National	20,361	Center for Labor and Social Studies (CEDLAS)
Estonia	2013	High	Living Conditions	National	6,576	Luxembourg Income Study (LIS)
Ethiopia	2010	Low	Living Conditions	National	18,864	Living Standards Measurement Study (LSMS)
Finland	2013	High	Living Conditions	National	11,112	Luxembourg Income Study (LIS)
France	2010	High	Living Conditions	National	14,440	Luxembourg Income Study (LIS)

Table A.1: Cross-Country Data Sources (continued)

Country	Year	Per Capita Income Group	Survey type	Coverage	Sample Size	Original source
Georgia	2010	Lower Middle	Living Conditions	National	4,811	Luxembourg Income Study (LIS)
Germany	2014	High	Living Conditions	National	14,915	Luxembourg Income Study (LIS)
Ghana	2010	Low	Living Conditions	National	62,042	Ghana Statistical Service
Greece	2013	High	Living Conditions	National	6,115	Luxembourg Income Study (LIS)
Guatemala	2014	Lower Middle	Living Conditions	National	22,118	Luxembourg Income Study (LIS)
Honduras	2009	Lower Middle	Living Conditions	National	98,028	National Institute of Statistics
Hungary	2014	High	Living Conditions	National	2,718	Luxembourg Income Study (LIS)
Iceland	2010	High	Living Conditions	National	4,133	Luxembourg Income Study (LIS)
India	2004	Low	Living Conditions	National	59,487	Luxembourg Income Study (LIS)
Indonesia	2011	Lower Middle	Living Conditions	National	111,824	Statistics Indonesia
Iraq	2011	Lower Middle	Living Conditions	National	176,042	Economic Research Forum (ERF)
Ireland	2010	High	Living Conditions	National	3,508	Luxembourg Income Study (LIS)
Israel	2014	High	Living Conditions	National	11,770	Luxembourg Income Study (LIS)
Italy	2014	High	Living Conditions	National	6,258	Luxembourg Income Study (LIS)
Jamaica	2002	Lower Middle	Living Conditions	National	18,943	Living Standards Measurement Study (LSMS)
Japan	2008	High	Living Conditions	National	7,840	Luxembourg Income Study (LIS)
Jordan	2010	Upper Middle	Living Conditions	National	15,472	Economic Research Forum (ERF)
Kenya	2005	Low	Living Conditions	National	62,175	National Bureau of Statistics
Kosovo	2000	Lower Middle	Living Conditions	National	14,167	Living Standards Measurement Survey (LSMS)
Liberia	2014	Low	Living Conditions	National	18,089	Institute for Statistics
Lithuania	2008	Upper Middle	Living Conditions	National	15,837	National Statistics Office
Luxembourg	2013	High	Living Conditions	National	4,373	Luxembourg Income Study (LIS)
Malawi	2011	Low	Living Conditions	National	56,218	National Statistical Office
Mali	2014	Low	Living Conditions	National	37,175	Living Standards Measurement Study
Mexico	2011	Upper Middle	Living Conditions	National	17,682	National Institute of Statistics and Geography
Mongolia	2003	Low	Labor Force	National	49,948	National Statistical Office
Morocco	2009	Lower Middle	Living Conditions	National	10,769	Ministry of Economy and General Affairs
Mozambique	2014	Low	Living Conditions	National	9,128	National Institute of Statistics
Namibia	2009	Upper Middle	Living Conditions	National	44,614	National Planning Commission
Netherlands	2013	High	Living Conditions	National	23,935	Luxembourg Income Study (LIS)
Nicaragua	2014	Lower Middle	Living Conditions	National	9,250	Center for Labor and Social Studies (CEDLAS)
Niger	2011	Low	Living Conditions	National	3,859	Living Standards Measurement Survey (LSMS)
Nigeria	2011	Lower Middle	Living Conditions	National	23,289	National Bureau of Statistics
Norway	2013	High	Living Conditions	National	23,993	Luxembourg Income Study (LIS)

Table A.1: Cross-Country Data Sources (end)

Country	Year	Per Capita Income Group	Survey type	Coverage	Sample Size	Original Source
Pakistan	2001	Lower Middle	Living Conditions	National	75,519	Federal Bureau of Statistics
Palestine	2011	Lower Middle	Living Conditions	National	25,947	Economic Research Forum (ERF)
Panama	2010	Upper Middle	Population and Housing Census	National	314,118	IPUMS-International
Papua New Guinea	1996	Lower Middle	Living Conditions	National	8,660	Living Standards Measurement Survey
Paraguay	2009	Lower Middle	Living Conditions	National	18,419	National Statistics Office
Peru	2009	Upper Middle	Living Conditions	National	95,199	National Institute of Statistics
Poland	2013	High	Living Conditions	National	39,993	Luxembourg Income Study (LIS)
Puerto Rico	2005	High	Population and Housing Census	National	35,416	IPUMS-International
Romania	1997	Lower Middle	Living Conditions	National	35,995	Luxembourg Income Study (LIS)
Russia	2013	High	Living Conditions	National	6,079	Luxembourg Income Study (LIS)
Rwanda	2000	Low	Living Conditions	National	32,679	National Institute of Statistics
Serbia	2007	Upper Middle	Living Conditions	National	17,375	Living Standards Measurement Survey (LSMS)
Sierra Leone	2003	Low	Living Conditions	National	23,022	National Office of Statistics
Slovakia	2009	High	Living Conditions	National	4,704	National Statistical Office
South Africa	2012	Upper Middle	Living Conditions	National	7,105	Luxembourg Income Study (LIS)
South Korea	2006	High	Living Conditions	National	13,178	Luxembourg Income Study (LIS)
Spain	2013	High	Living Conditions	National	10,728	Luxembourg Income Study (LIS)
Sri Lanka	2008	Lower Middle	Labor Force	National	66,381	Department of Census and Statistics
Sudan	2009	Lower Middle	Living Conditions	National	48,845	Economic Research Forum (ERF)
Sweden	2005	High	Living Conditions	National	11,607	Luxembourg Income Study (LIS)
Switzerland	2013	High	Living Conditions	National	7,961	Luxembourg Income Study (LIS)
Taiwan	2013	High	Living Conditions	National	23,474	Luxembourg Income Study (LIS)
Tajikistan	2007	Low	Living Conditions	National	1,503	State Statistical Agency
Timor Leste	2007	Lower Middle	Living Conditions	National	9,094	National Statistics Directorate
Tunisia	2009	Upper Middle	Living Conditions	National	50,371	Economic Research Forum (ERF)
Turkey	2011	Upper Middle	Labor Force	National	37,121	National Statistical Institute
Tanzania	2010	Low	Living Conditions	National	20,559	National Bureau of Statistics
Uganda	2011	Low	Living Conditions	National	13,618	National Bureau of Statistics
Ukraine	2010	Lower Middle	Living Conditions	National	10,428	State Statistics Service
United Kingdom	2013	High	Living Conditions	National	20,002	Luxembourg Income Study (LIS)
United States	2013	High	Living Conditions	National	63,859	Luxembourg Income Study (LIS)
Uruguay	2009	Upper Middle	Living Conditions	National	132,559	National Institute of Statistics
Venezuela	2006	Upper Middle	Living Conditions	National	166,506	National Institute of Statistics
Zambia	2014	Lower Middle	Living Conditions	National	11,921	Central Statistical Office

Notes: for details on this table, please see Section A.1.

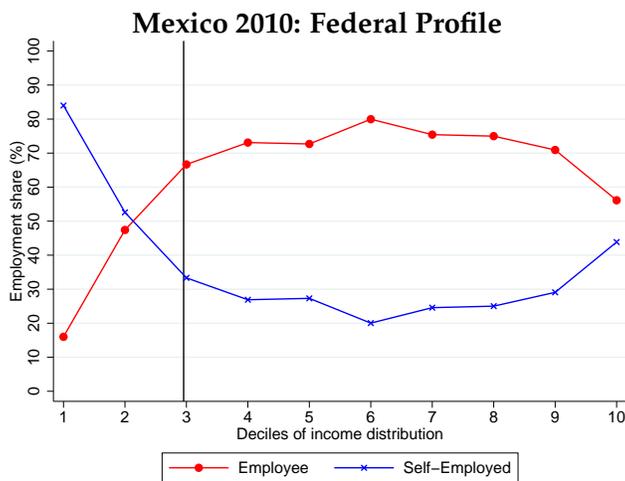
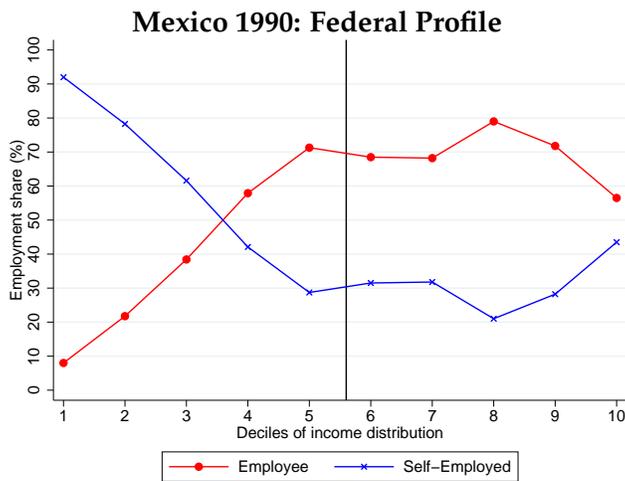
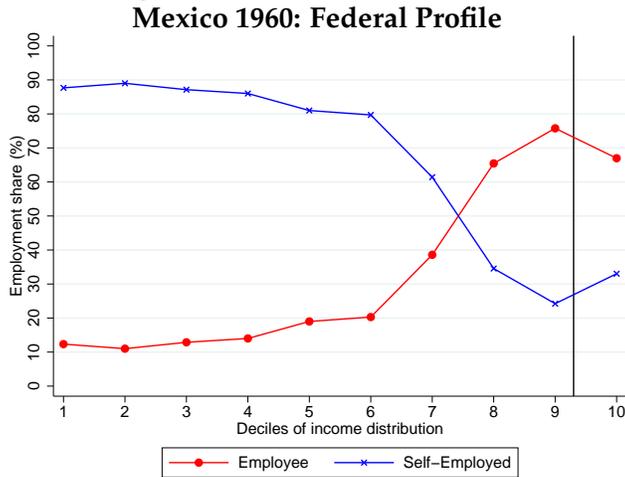
A.2 Additional historical profile: Mexico 1960-2010

As a robustness check to the stylized facts, I show that they also hold over the long-run in a currently developing country, Mexico. I focus on Mexico because it has variables of income and employee-jobs that are consistently defined over a long period of time, namely 1960-2010. The data is extracted from IPUMS International. The disadvantage is that only earned income is measured consistently over this period - as opposed to total income, which further includes capital income and 'other' income. I use answers to the 'class of worker' question. The only inconsistency over time in this question is that the 2010 sample groups household assistants together with salaried workers, whereas in previous samples, these categories are separated. As such, I am over-estimating the true employee-share in the 2010 profile. Importantly, day laborers are separated from salaried workers, and I can assign the former to the self-employment category in all years. There also exists a category for unpaid family workers, which I assign to the self-employment category. I construct the sample of respondents that are economically active, and use survey weights to construct individual earned income distributions in 1960, 1990, and 2010. For the years 1990 and 2010, I code the value of the exemption threshold from OECD's Personal Taxes database. For 1960, I use the historical archives of the Mexican Tax Authority.¹

The results from this exercise are displayed in Figure A.1. I uncover the same stylized facts that were found both in the cross-country sample and in the historical US series: the employee-share profile is upward-sloping and gradually moves leftward in the gross income distribution; the exemption threshold gradually moves down the distribution and expands the size of the income tax base; and, the employee-composition on the tax base is constantly maximized.

¹Available at: <http://www.dof.gob.mx/index.php>.

Figure A.1: Mexico 1960-2010



Notes: The circle-line (small cross-line) indicate the employee-share of the economically active workforce in a decile of the Mexican personal gross income distribution. An employee-job is defined as a job whose activity generates an information trail that can be leveraged for income tax enforcement purposes. For more details on this variable, please see Section 3.2. In every profile, the vertical solid line denotes the location of the Federal individual income tax exemption threshold. This threshold is the nominal value of gross (pre-tax) income above which a single filer becomes liable to pay income tax. Each historical profile is built from the Census micro-data from IPUMS International. The values of the exemption thresholds are from the OECD's Personal Taxes database, and the official archives of the Mexican revenue service. Source: Section 3.4 and Appendix Section A.2.

A.3 Redistributive targeting

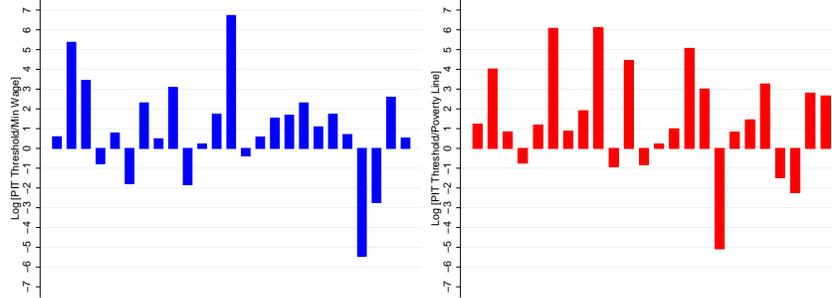
In this robustness check, I provide evidence to suggest that the exemption threshold is not set to target social assistance or anti-poverty in the income distribution. Governments define thresholds of income that are used as inputs in formulas to provide social assistance and anti-poverty relief. I use the national poverty line and the minimum wage values as proxies for the 'social redistribution' threshold. I first show that only very rarely is the income tax threshold explicitly defined to be either equal to, or a multiplicative of, this social redistribution threshold. In 5% of countries in the cross-sectional sample, the tax code defines the exemption threshold to be a multiple of this redistribution threshold. These countries are: Mozambique, Bolivia, Paraguay, Turkey, and Slovakia. As an example, in Mozambique the exemption threshold is equal to 36 times the minimum wage, while in Paraguay it is equal to 120 times the minimum wage. I use the country-specific IBFD tax summaries to document this pattern. There exists a much more frequent explicit relation between redistributive thresholds and social security contributions. Indeed, several countries use (a multiplicative of) the minimum wage to define an exemption threshold for employee contributions.

Even if there exists no explicit relation defined in the tax code, governments may nonetheless implicitly maintain an association between the tax threshold and the social assistance threshold. To investigate this, I collect data on the value of the national poverty line and the minimum wage in all countries in the cross-sectional sample. I try to collect the data in as close a year as possible to the survey and tax exemption threshold year. I use harmonized data from ILO on the statutory nominal gross monthly minimum wage. Data is missing in 8 countries: Austria, Denmark, Finland, France, Kosovo, Sweden, Switzerland, Palestine. There does not exist a similar harmonized database on the value of the national poverty line for my sample. The World Bank collects cross-country data on the share of the population that falls below both international and national poverty lines, but such data does not directly disclose the value of the national lines used. I was able to collect relevant data in 88 of the 100 countries in my sample. The missing countries are: Albania, Austria, Hungary, Kosovo, Panama, Papua New Guinea, Romania, Serbia, Slovakia, Timor Leste, Ukraine, and Venezuela. Importantly, I collect the poverty line that is set by the national government, rather than the value of the international poverty line in local currency. Some governments do incorporate international criteria to determine poverty lines. Per example, some low-income countries base their poverty calculations on the minimum nutritional intake concept used by the World Bank to define international poverty; and, some European countries adopt the EU-wide definition of poverty as 60% of median income. The important point is that the poverty lines I collect are based

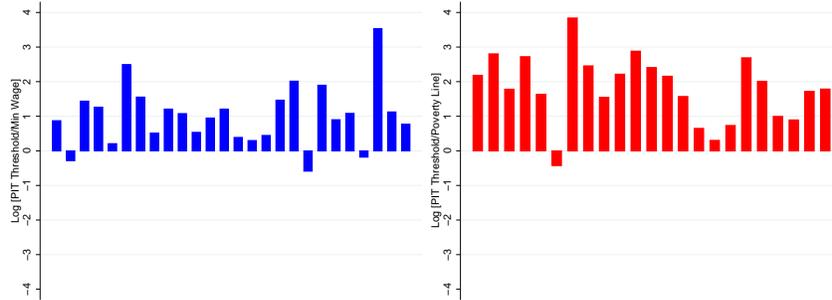
on an active decision made by the government, similarly to the definition of the tax exemption threshold. In some countries, the government defines several poverty lines, per example on a regional basis or on an urban-rural basis. I always pick the poverty line in each country with the highest value. Since poverty lines in developing countries are most often below the tax exemption threshold, this decreases the likelihood to observe that the two thresholds are far away from each other in value.

The results are displayed in Figure A.2. The three panels separate countries into development groups: low and lower middle income; higher middle income; and, high income. I construct the ratio of the income tax exemption threshold to the minimum wage, and of the exemption threshold to the poverty line. In the left-hand graphs, the bars represent country-specific ratios using the minimum wage, while the right-hand graphs display the ratio using the poverty line. Finally, within each graph, I sort the countries by GDP per capita. I take the log of the ratio, as this allows me to display all country-ratios on the same graph. Therefore, a bar-value below 0 means that the exemption threshold is located below the minimum wage/poverty line in the specific country. There is no obvious, confounding trend which emerges from Figure A.2. Within all development groups, countries with similar per capita income, and hence similar size of tax base (Figure 4), display very large variation in the relative value of the tax threshold to the redistribution threshold (note the log-scale of the y-axis). This holds even for countries at similar levels of development within the same region: the ratio for the minimum wage (poverty) is 0.48 log points (1.90 log points) in Burkina Faso, while it is 3.07 log points (6.09 log points) in Uganda; it is 2.09 log points (2.79 log points) in Bolivia, and 0.51 log points (2.65 log points) in Honduras. The highest-income countries often locate both the poverty and the minimum wage thresholds above the tax exemption threshold. But apart from this feature, there is not any systematic relationship between the relative location of tax and redistribution thresholds, and per capita income. Taken together, these findings suggest that the tax exemption threshold is not set to target social assistance in the income distribution.

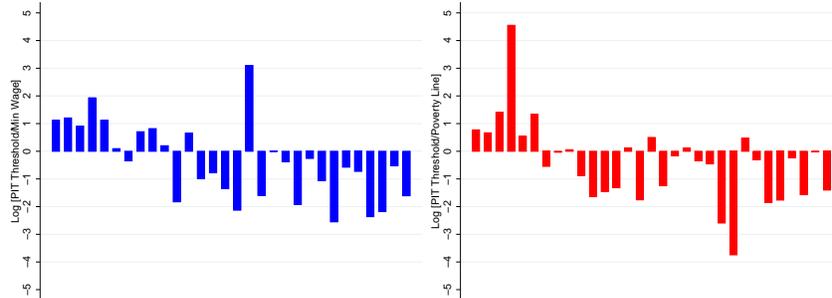
Figure A.2: Redistributive Targeting
Panel A: Low and Lower-Middle Income Countries



Panel A: Upper-Middle Income Countries



Panel C: High-Income Countries



Notes: In every graph, a bar represents a country-observation from the cross-country micro-database. The three panels demarcate countries according to their per capita income group: low and lower-middle; upper-middle; high income. Within each graph, countries are ranked in ascending order of per capita income. Within each group, the left-hand graph shows the log of the ratio of the income tax exemption threshold to the minimum wage; the right-hand graph shows the log of the ratio of the income tax exemption threshold to the poverty line. All thresholds are expressed in annual and local currency. Source: Section 3.4 and Appendix Section A.3.

A.4 Sectoral distributional profiles

In this robustness check, I consider whether the location of the exemption threshold is targeting sectoral structure, rather than employment structure. I study whether the threshold appears to be set such as to avoid a 'hard to tax' sector, agriculture, or whether the threshold is set to capture the 'easy to tax' sectors of manufacturing and public administration (Musgrave, 1981). In order to investigate this confounding hypothesis, I first consider whether the tax exemption explicitly targets any sector. In particular, I use the IBFD country-reports in all countries in the cross-section, and report whenever income from agriculture is fully exempt from individual income taxation. I do not take into account instances where tax codes allow self-employed to deduct costs specifically related to agricultural work - per example, from the purchase of a tractor for farming. This is because my measure of the exemption threshold in all countries is the standard deduction, which is granted regardless of taxpayer behavior, and not the itemized deduction, which requires the taxpayer to itemize deductions. I chose the former measure because it can be constructed in a comparable way across space and time, as discussed in Section 3.2. I find that agricultural income is fully exempt only in 11% of low-income countries; 12% of middle-income countries; and, 5% of high-income countries. This list includes Mali, Morocco, and Sierra Leone in Africa; and, India and Pakistan in Asia.

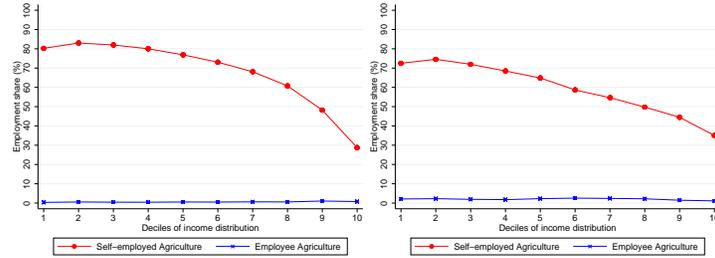
As an alternative approach, I consider whether changes in sectoral distributional profiles over development could account for the movement in the exemption threshold. I create four sectoral categories in all the surveys in the cross-section: agriculture; manufacturing and construction; trade and services; and, public administration. I define these four categories in relation to the divisions of the ISIC 4.4 classification. The construction of the sector variable is described in detail in Section A.1.1.

Using these harmonized sector variables, I first study the distributional profiles of agricultural employment. I construct these profiles in the same way as the employment profiles in the main text (Section 3.2). The results are displayed in Figure A.3. At lowest levels of development, agriculture is prevalent everywhere except for the top of the income distribution. And, in the transition from low-income to middle-income group, the downward-sloping agriculture-profile gradually shifts leftward in the distribution. This pattern is similar to stylized fact #1, such that stylized fact #2 would be consistent with a setting where the exemption threshold targets the non-agricultural sector which increases gradually further down the income distribution. However, Figure A.3 also reveals that in these same income groups, virtually all agricultural work is concentrated among self-employed with no information trails. On the other hand, Figure A.3 reveals that in the transition from middle-income to the high-income group, the agricultural profile has be-

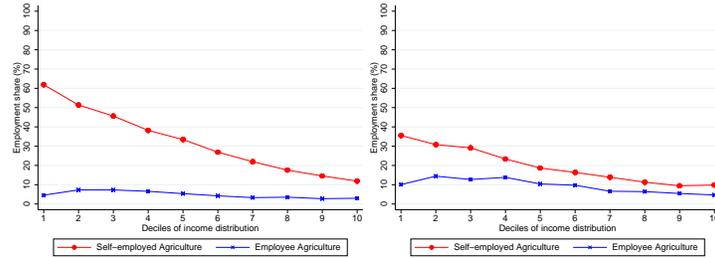
come very small in magnitude and almost entirely flat in the distribution. During this same transition, there continues to be an important transition between self-employment and employee-jobs outside of the agricultural sector, which is associated with further decreases in the location of the threshold. These facts suggest that movement out of agriculture could account for the expansion of the tax base, but only in a limited range of the development path, where it is fully confounded by movements out of self-employment. In contrast, movements out of self-employment can account for the expansion of the base over the full development path, including over a range of development where it cannot be confounded by movements out of agriculture.

I now consider whether the movement of the exemption threshold is consistent with targeting of 'easy to tax' sectors. I focus on manufacturing and public administration. Since work in these sectors is strongly correlated with having an employee-job, I study the sectoral profiles conditional on employee-job. Results are displayed in Figure A.4. The distributional profile of easily taxable sectors would have to be upward-sloping in the income distribution, and move leftward as the country develops, in order to be a confounding factor. This is not borne out in the observed profiles. The public administration profile is upward-sloping at some development levels, but the magnitude of the slope is quantitatively small, and there is no consistent left-ward shift over development. The public sector share at the top of the income distribution is most likely driven by central administration workers, while the share towards the lower end of the distribution is probably made up in part by field-workers in health and education. While located at very different parts of the income distribution, these jobs share the common feature of being easy to tax - in the sense that the government, as the direct employer, perfectly observes the salaries. The manufacturing distributional profile is largely flat in the income distribution. The level-shift upward and then downward of the manufacturing profile is consistent with the inverse-U shaped aggregate importance of manufacturing over development that other work has documented. Taken together, these facts do not suggest that the stylized facts #1-#4 are confounded by sectoral transitions over the development path.

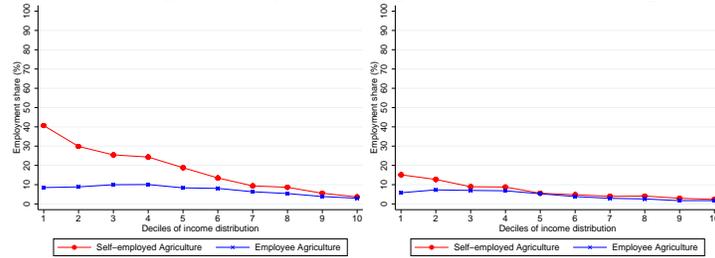
Figure A.3: Distributional Profiles of 'Hard-to-Tax' Sectors
Profile for average country at \$1065 pc [LHS] and \$2226 pc [RHS]



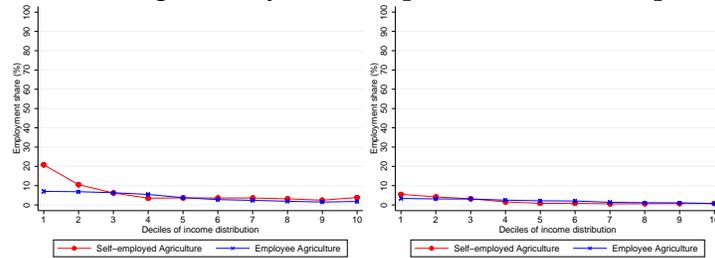
Profile for average country at \$3239 pc [LHS] and \$5796 pc [RHS]



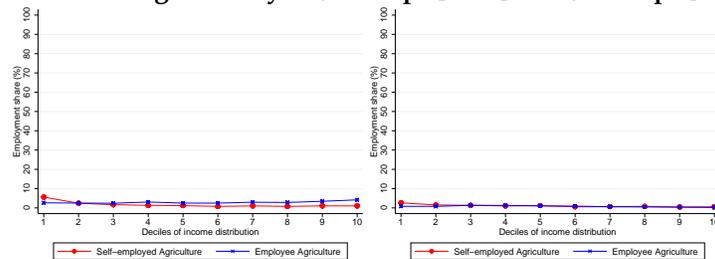
Profile for average country at \$8826 pc [LHS] and \$11257 pc [RHS]



Profile for average country at \$17141 pc [LHS] and \$27960 pc [RHS]

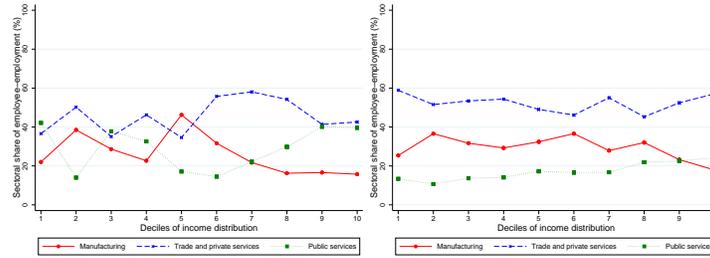


Profile for average country at \$38224 pc [LHS] and \$53878 pc [RHS]

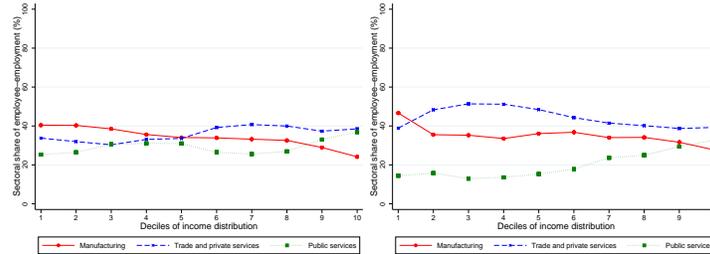


Notes: These figures plot the employment shares of self-employed agricultural workers and of employee agricultural workers, over deciles of the income distribution, for representative countries at different levels of per capita income. The share of each group is defined as the share of the total economically active workforce in the decile of the income distribution. To construct this graph, I partition the cross-country sample into ten groups of equal size, based on their level of per capita income. Note that I am limited to the group of countries where there exists sectoral data. Within each group, I calculate the unweighted average employment-share of agricultural self-employed and agricultural employee. I plot this average profile for every group, and indicate the average per capita income of the group. I use expenditure-side real GDP at chained PPPs in 2011 US\$ from the same year as the country-survey year. Source: Section 3.4 and Appendix Section A.4.

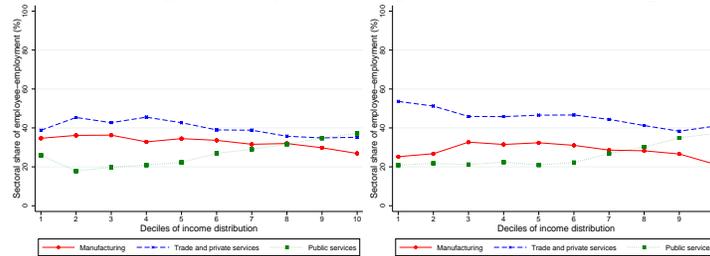
Figure A.4: Distributional Profiles of 'Easy-to-Tax' Sectors
Profile for average country at \$1065 pc [LHS] and \$2226 pc [RHS]



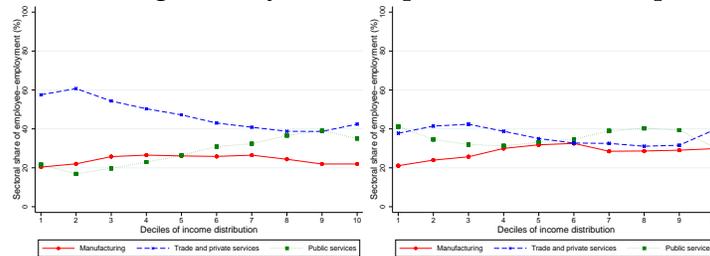
Profile for average country at \$3239 pc [LHS] and \$5796 pc [RHS]



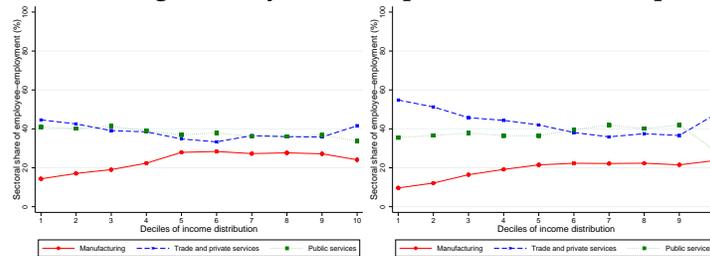
Profile for average country at \$8826 pc [LHS] and \$11257 pc [RHS]



Profile for average country at \$17141 pc [LHS] and \$27960 pc [RHS]



Profile for average country at \$38224 pc [LHS] and \$53878 pc [RHS]



Notes: These figures plot the sectoral shares of employees over deciles of the income distribution, for representative countries at different levels of per capita income. Sectors are defined according to the ISIC classification (Section A.4). The share of each sector is defined as the share of the total employee workforce in the decile of the income distribution. To construct this graph, I partition the cross-country sample into ten groups of equal size, based on their level of per capita income. Note that I am limited to the group of countries where there exists sectoral data. Within each group, I calculate the unweighted average sectoral shares by income decile. I plot this average profile for every group, and indicate the average per capita income of the group. I use expenditure-side real GDP at chained PPPs in 2011 US\$ from the same year as the country-survey year. Source: Section 3.4 and Appendix Section A.4.

A.5 Potential biases resulting from methodology

In this subsection, I discuss the potential biases that can arise from the survey methodology and the measurement and construction of variables. I code employment type based on the primary job in which the respondent spent the most hours during the reference period. Many individuals have many jobs at the same time (Banerjee & Duflo, 2007). But this will affect the representativeness of my estimates only to the extent that these jobs fall in different categories in my classification. An individual who contributes on the family farm while being an own-account worker within the same reference period would be classified as 'self-employed' in both jobs. In surveys where the reference period is not yearly, there may be bias in the measure of employment structure if the employment type in the reference period is not representative of the entire year. This is potentially important in developing countries, where there is strong seasonality in job type. This introduces bias to the extent that the jobs at different periods of the year fall in different employment structure categories, which I argue is unlikely in a developing country context. Indeed, individuals that are casual wage laborers during the harvest season are unlikely to be regular full-time employees in the non-harvest season. Rather, they are likely to be own-account workers or contributing family workers. In this case, the individual would be classified as self-employed during all periods of the year, despite the different jobs held at different periods of the year.

A second source of bias comes from the fact that I cannot systematically separate casual wage work from contract-based wage-work. I can always distinguish between working for someone for pay versus for in-kind payment, and I exclude the latter from the employee category. As such, casual wage-laborers that receive in-kind payment are systematically classified as self-employed. This leaves the group of casual workers that are not paid in-kind as the group that I potentially mis-classify as employee, whenever the survey answers do not provide sufficient precision about the nature of the employee-work. Since the transition over development involves a movement out of casual wage labor into contract-based wage labor, this mis-classification will lead me to under-state the true growth in employee-share along the development path.

Another potential source of bias arises from the possibility that self-employed misreport their true amount of income. This is unlikely to introduce a major bias, for three reasons. First, self-employed face much smaller, if any, incentives to mis-report income in survey data than in administrative tax data. Consistent with this idea, studies from the U.S. that match survey and administrative sources find that self-employed sources of income are systematically larger in survey data, while there are no significant discrepancies for employee income-sources (IRS, 2008). Second, the model in Section 5 does predict under-

reporting of income among self-employed locally around the exemption threshold. But while the standard bunching model predicts a steepened employee-share locally around the exemption threshold, it also predicts a decrease in employee-share further to the left of the threshold. This is not borne out in the data: instead, I observe a gradual increase in the employee-share over the full distribution. There is therefore no strong evidence of evasion-driven bunching behavior in the survey data. Under-reporting of income by self-employed could also be due to by non-evasion motives. Woodruff et al. (2009) show that recall error, which is more present when the reference period is not annual, lead self-employed to under-estimate their income. Both in the case of evasion and recall error, the true self-employed distributional profile would lie to the right of the observed one. If development is associated with increases in the ability to detect under-reporting among self-employed, this would generate gradual leftward shifts of the employee-share profile. Similarly, if development is associated with a decrease in recall error, either due to changes in survey methodology or to an increase in accounting tools and book-keeping, this would similarly imply a leftward shift of the employee-profile at increasing levels of development. Both evasion detection capacity and measurement precision, which plausibly grow with development, therefore lead me to under-estimate the true progressive rightward shifts in the employee-profile due to transitions from self-employment to employee-employment.

Finally, bias could be introduced from the construction of the income tax base. I construct the tax base as the share of the individual income distribution that lies above the single-filer standard deduction (or allowance). As explained in the main text (Section 3.2), this choice is made to construct the tax base in the most transparent way without making any behavioral assumptions and in a way that can meaningfully be compared across countries. Notwithstanding, there exists features of tax systems which allow taxpayers to further reduce their tax liability, including deductible expenses. If a significant number of filers makes use of such additional features, this introduces a wedge between the size of base measured in this paper, and the size of the 'effective' base. The extent of existence of these features varies significantly across countries. Per example, there is a growing policy debate in the US on the large number of taxpayers that do not pay any Federal income tax. There exists no consistent evidence across countries at different levels of development on the extent to which the effective tax base is reduced through credits and deductions. Even if taxpayers in all countries in my sample made use of these deductions, it is likely that the size-wedge between my measured base and the effective base is larger in more developed countries. This is simply because the potential wedge in less-developed countries is bounded above by the small size of my measured base. In this case, I am overstating the variation in size of base across levels of development (Panel

B, Figure 4). Perhaps more importantly, a size-wedge that increases with development means that I am understating the strength of the association between size of tax base and income tax collection (Panel B, Figure 5). This point is also supported by the observation that the variance in residual tax collection, controlling for the statutory size of tax base, is larger in more developed countries (Panel C, Figure 5). This discussion suggests that bias introduced by the wedge between my measure of the base and the effective size of tax base only strengthens the main finding of the tax base being a first-order determinant of tax collection across development.

A.6 Robustness of employee-income gradient in regression setting

In this subsection, I investigate the employee-income gradient in a regression setting. This provides a complementary characterization of the relationship between income and employee-status to the income-decile distributional profiles in the main text. In addition, it allows me to begin to investigate which factors contribute to changes in the income-employee gradient over development. I focus on three characteristics: sector, location, and education. These are observable characteristics which vary over development, due to the sectoral movement from agriculture to manufacture to services; the rural-urban migration; and, the rise in higher education.

I use the four sectoral categories described in the Section A.1.1. In addition, I create a dummy variable equal to 1 if a respondent lives in an urban area. I do not attempt to harmonize this variable, and use the urban definition directly in the surveys. Finally, I use education variables to code four dummies, indicating if a respondent has: not completed primary; completed primary but not high school; completed high school. I chose to not distinguish further levels of education, in order to maximize the number of surveys where I could create consistent measures.

To visualize the impact of controlling for a characteristic on the employee-income gradient, I estimate the following regression in every country c

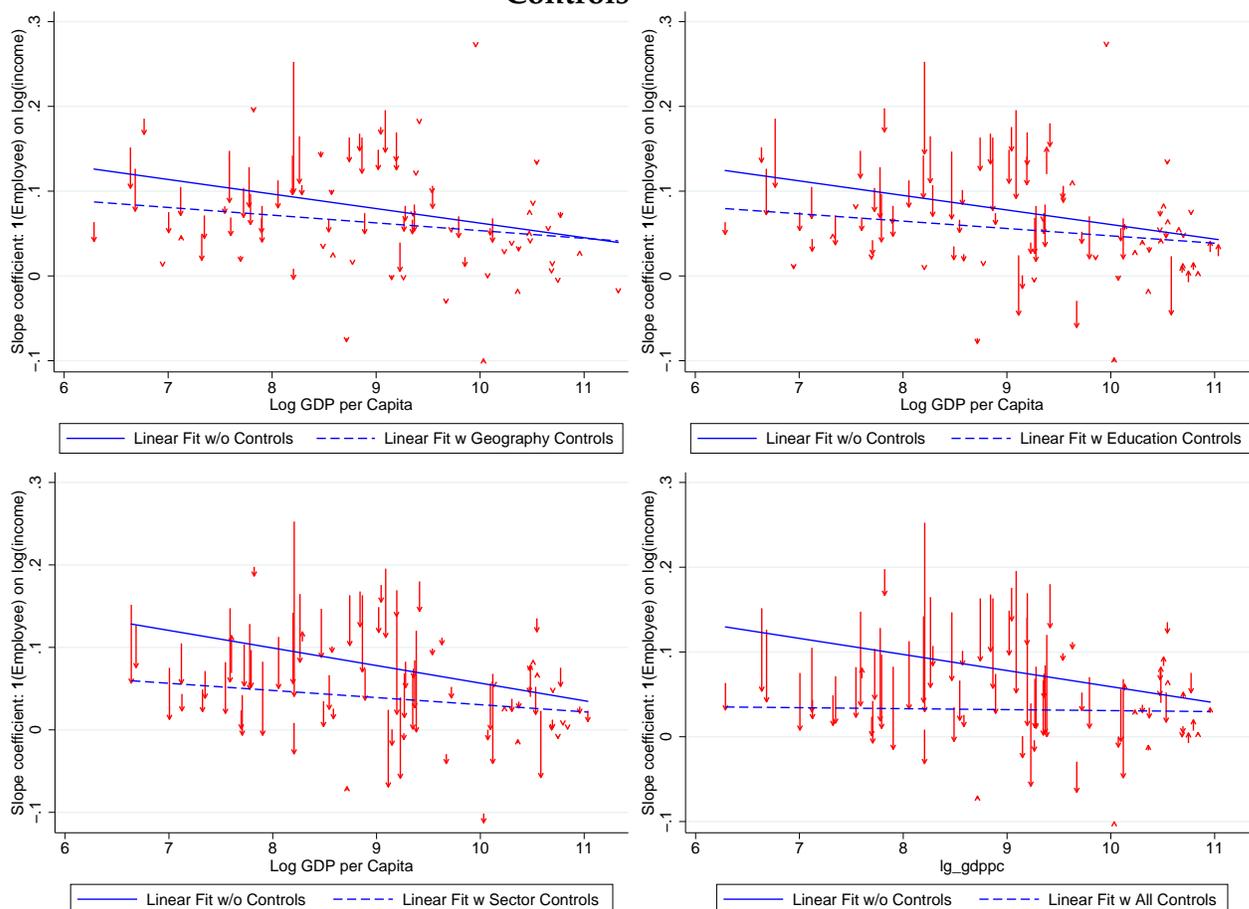
$$\mathbf{1}(\text{Employee})_i = \alpha + \theta \mathbf{X}_i + \beta \log(\text{income})_i + \varepsilon_i$$

where income_i is the individual gross income of individual i used to construct the income distribution (Section 3.2), $\mathbf{1}(\text{Employee})_i$ is a dummy equal to 1 if an individual is an employee (Section 3.2), and \mathbf{X}_i contains the control indicator variables (sector, education, urban). I obtain a country-specific slope-coefficient β^c from estimating this regression separately in every country. In every graph, I plot these coefficients $\beta_{with\ control}^c$ together with coefficients from estimating the regression without controls, $\beta_{no\ control}^c$ against log per capita income. The two coefficients for a particular country are denoted by the beginning ($\beta_{no\ control}^c$) and end ($\beta_{with\ control}^c$) of a vertical arrow. This regression is a linear probability model, which has the advantage that the slope-coefficient is directly interpretable. The disadvantage is that the slope-coefficient is not informative in settings where the relationship between employee and $\log(\text{income})$ is strongly non-linear. This is the case in less (most) developed countries, where the likelihood of being an employee is very small (large) apart from the very top (top and bottom) of the income distribution (Figure 3). Results (not reported) are similar when using the employee-share differential between the top and bottom deciles.

The results are displayed in Figure A.6. In most countries, there remains a strong positive

association between income and employee-status in both urban and rural areas (upper-left panel). The within-education group gradients are somewhat smaller, but continue to be large especially in low and middle-income countries (upper-right panel). The most significant reduction in slopes comes from controlling for industry fixed effects. Nonetheless, the income-gradients continue to be larger in less-developed countries. Finally, the full inclusion of controls leads to a flat relationship between magnitude of income-employee gradient and development. This suggests that, while neither of the factors considered entirely confounds the employee-income gradient, the joint variation in the three factors can contribute to the observed changes in gradients in stylized fact #1.

Figure A.6: Employee-Income Gradients Across Countries Without and With Controls



Notes: Each dot in every scatter-plot represents a country-specific slope coefficient based on estimating the regression in Section A.6. Each of the four graphs show slope-coefficients when including controls for: geography (North-West quadrant); education (NE); sectors (SW); geography, education, and sector (SE). In each graph, the start-point of an arrow represents the country-specific slope-coefficient without the control, and the end-point of an arrow represents the slope-coefficient after including the control. All slope-coefficients are plotted against log GDP per capita, measured using expenditure-side real GDP at chained PPPs in 2011 US\$. In every graph, the solid (dashed) line represents the linear OLS fit of the slope-coefficients without control (with control). For more details on the construction of the different control variables, please see Section A.1.1. These graphs are constructed using the full cross-country survey sample. Source: Section 3.4 and Appendix Section A.6.

B US states material

B.1 Data sources and construction of variables

In this sub-section, I describe the construction of variables used in the US states analysis (Section 4).

B.1.1 Employment and earnings

I construct the aggregate employment-share variables using decennial Census data at the state level between 1930 and 2010. The data is extracted from IPUMS USA. In each decennial data-extract, I exclude from the sample any individual that is not economically active during the reference period and for whom the general class of worker variable is 0 (“N/A”). I also exclude, when possible, any individual who reports total personal income either equal to 9999999 (“N/A”) or strictly negative. In the IPUMS USA data, total personal income corresponds to the respondent’s total pre-tax personal income or losses from all sources for the previous year. I code as self-employed (employee) a respondent who responds ‘self-employed’ (‘works for wages’) in the class of worker category. This classification in IPUMS USA is consistent with the classification used in the cross-development sample, in the sense that I code the employment-type based on the primary job of the respondent in which they spent the most time during the reference day or week. Within each decennial extract, I apply person-weights to estimate, for each state, the representative total number of respondents, the total number of employee respondents, and the total number of self-employed respondents. I then calculate the employee-share as the ratio of total number of employee respondents to the total number of employee and self-employed respondents. I interpolate the numerator and denominator between Census years using a natural cubic spline (Herriot & Reinsch, 1973).

I construct the employment shares by income decile of the income distribution of each state, in 1935 and in every decade between 1950 and 2010. The 1950-2010 data is extracted from the IPUMS USA database. The definitions of type of work and industry are the same as those used to construct the state-year aggregate employment shares. I rank all respondents within a given state according to the reported total personal income. The personal income reported measures each respondent’s total pre-tax personal income. Importantly, throughout the sample period, this measure is largely comparable: it includes in all samples, wage, farm and business components. I then apply person-weights and partition each state’s income distribution into ten deciles (ten bins of equal sample size). Within each decile, I estimate the conditional proportions of employees and self-employed to construct the employee-shares by income decile. In years before 1950, the decennial US Census does not provide reported income and occupation-category at

the level of the individual. I use the 1935-36 Study of Consumer Purchases in the United States, which had the scope to 'ascertain for the first time in a single national survey the earning and spending habits of inhabitants of large and small cities, villages, and farm'. I access this data under the ICPSR data archive reference #08908. I discuss the 1935 data-sample and construction of variables in more detail in Section A.1.2. I construct the deciles of the state-specific income distribution and estimate the employment shares specific to each decile-state. I use these data to construct the profile of employment-share and self-employment share over deciles of each state's income distribution, for all continental states, between 1935 and 2010. I again interpolate both the numerator and denominator between data-years.

The earning structure is constructed for all states and all years between 1929 and 2001 by combining the two historical series, namely SA5H and SA5 'Personal Income by Major Components and Earnings by Industry' published by the US Bureau of Economic Analysis. The denominator for earnings-structure is line-item 45 'Net earnings by place of residence', which equals total earnings less contributions for government social insurance plus 'adjustment for residence'. The employee-share uses in the numerator line-item 90 'private non-farm earnings', while the self-employed share of income uses line-item 70 'proprietors' income'. The line-item 45 is also used as the denominator y to construct the ratio of the PIT-threshold K to average earnings, K/y . Importantly, this measure y of personal income excludes transfers from all levels of government, similarly to the gross income variable used in the cross-development sample.

B.1.2 Tax revenue

The tax-revenue sources by state and year are based on the historical series on the historical State Government Finances series. I proxy for government size by constructing the ratio of tax-revenue collected to total personal income, where the denominator is based on the BEA historical series of state personal income. This is consistent with previous studies of government size (Barro and Sala-i-Martin, 1992; Besley et al., 2010) that also use personal income in the denominator rather than GDP. GDP data at the state-year level in the US is only available from 1963 onward. T40 is the line-code corresponding to personal income tax; T41 corresponds to corporate net income tax; and, T09 corresponds to general sales tax.

B.1.3 Personal income tax structure: thresholds and rates

I use the state income tax calculator program from Bakija (2009). I thank Jon Bakija for kindly providing me access to the calculator, which is based on the legal tax codes from all states. I construct the income tax threshold K for an individual earner who files under the status of being single and who claims the standard deduction. This exemption thresh-

old is constructed in the same way as the threshold measure in the cross-development setting. Unlike the itemized deduction, the filer does not deduct state personal income tax from her federal income tax liability. This provides additional incentives for the filer to under-report state income taxes, consistent with the under-reporting model in Section 5. Standard deduction filers are systematically more prevalent at lower levels of gross income (IRS, 1982). I construct the ratio K/y where y is the state-year per capita personal income, extracted from the historical US BEA series. Finally, I use the state tax calculator to measure the marginal tax rate that applies on the first income tax bracket. These variables are used in Table 3.

B.1.4 Covariates

The poll tax and literacy test dummies are taken from Besley et al. (2010). The authors use variation in these dummies to instrument for political competition, which they find to have a significant impact on the share of non-farm income and tax revenues.

I construct proxies for the state-year policy environment. These different proxies are meant to capture variation in state-policies which may have affected location decisions of private firms. First, a dummy for the existence of a corporate income tax is constructed, which takes value 1 in all years in a state where there exists such a tax-base. The date of creation of state corporate income tax is taken from Table 4.1 of Newman (1984). The dummy for the existence of right-to-work laws is extracted from Besley et al. (2010). Right-to-work laws make it illegal to demand that employees join a union, or to automatically deduct union fees from wages. The continuous measure of state unemployment insurance firm-size coverage is taken from the historical publication series ‘Significant Provisions of UI State Laws’ published by the US Department of Labor. I download all publications between 1937 and 1979. In each state-year, I code the firm-size coverage, that is the firm-size above which an employee is entitled to receive state UI benefits. Federal-time varying regulation provided an upper-bound on the allowed firm-size, but states were free to legislate in order to define a firm-size below the Federally mandated size. I use these additional covariates in regression table B.3.

B.1.5 Additional outcome variables

I construct a proxy for tax administrative reforms based on the historical series of the Book of the State, published annually from 1993 until today by the Council of State Governments. I collect data at the state-year level on the number of agencies administering major taxes: property, income, sales, gasoline, motor vehicle, tobacco, death, liquor. I code the total number of state tax agencies in operation in every state-year. This variable is available from 1939 to 2009. This variable is intended to proxy for investments in enforcement capacity, through consolidation of the number of tax agencies, and is used

in Table 3. I also collected state-year data from the same source on the annual salaries of the chief state administrative official in different departments: revenue-collection and taxation; treasury; attorney general. I then constructed the ratio of the annual salary in revenue-taxation relative to the salary in the Treasury and to the salary as Attorney General. These ratios were meant to proxy for investment in enforcement capacity through funding higher wages to tax administrators (relative to other state administrators). In results not reported, I do not find an impact of the upholding event on this measure of relative pay. These variables represent, to my knowledge, the first long-run time-series on proxies for tax administrative capacity of individual states in the US. As an additional proxy for enforcement capacity, I code the year when each state adopted withholding of state personal income taxes by employers. Evidence shows positive impacts of withholding on state's income tax collection (Dusek & Bagchi, 2017). I use the historical IRS 'Annual Report' series to code the years of adoption. This variable is used in Table 3.

I use data from Besley et al. (2010) to build proxies for political outcome-variables. I use their measure of party-neutral political competition and the Democratic vote-share averaged across all state-wide elections. These measures are used in Table 3.

In Table 3, I proxy for generosity of the state's unemployment benefits, using the measure of state maximum unemployment benefits. This variable is taken from the 'Correlates of State Policy' database (Jordan & Grossman, 2017).

Finally, in the robustness checks (Table 3), I study the impact of upholding on level of income, and income inequality. I use the 'net earnings' measure of income from BEA, and the top 1 percent income share from Frank et al. (2015).

B.1.6 Exchange of information agreements

I code the year of implementation of the agreement from the historical IRS series 'Annual Report'. The signature of the exchange of information acts has been found to increase income tax revenue (Troiano, 2017). Troiano's source for the year of implementation is Penniman (1980), which only differs slightly from my source. My results are robust to using this alternative measure of implementation dates. This variable is used in Table 2 and Table 3.

B.1.7 Cost of collection

I construct the measure of cost of collection from the Book of the States. The cost-components of collecting state taxes are capital outlays, operating costs, and payroll. I divide this total cost by the total collected tax revenue. This construction is similar Jensen & Lagakos (2019), which studies variation in cost of national tax administrations across levels of development. Interestingly, the cost of collection in the average US state in 1962 is slightly higher than the average current low-income country's tax administration.

B.2 Program details: Industrial Development Bonds

In this sub-section, I provide additional information on the Industrial Development Bonds program (IDB), and the legal uncertainty which generates variation at the state-level in the effective implementation date.

The IDB was a place-based local development program that sought to attract industrial facilities to predominantly rural areas characterized by 'surplus labor' concentrated among self-employed farmers (Advisory Commission on Intergovernmental Relations, 1962). The first state to implement an IDB program was Mississippi, when it launched 'Balancing Agriculture with Industry' in 1936. Practice of IDB did not, however, become a multi-State practice until the mid-1950s, when several other States decided to implement similar programs.

The official justification for government intervention was that these rural areas were "deficient in credit facilities" (ACIR, 1962), and capital for local firms was not readily available from conventional credit sources. Through the IDB, the local government therefore sought to relieve a local credit constraint. In the IDB program, sub-state government units (counties, boroughs, and cities) issue bonds to finance the acquisition or construction of facilities and equipment for lease to private firms. Importantly, IDB issuances were revenue bonds, which are secured exclusively by the revenues of the project. This is in contrast to general obligation bonds, which are secured by the credit of the issuer - in this case, the local government. This distinction removes the direct relationship between issuance of IDB and increased tax revenue to solidify credit capacity.

For identification purposes, I exploit the institutional features of implementation. In particular, the particular methods under IDB were unprecedented in the context of postwar state financing. The use of public credit for an otherwise private purpose was considered to be in direct violation of the public purpose doctrine, which prohibits such usage. Constitutions of many States explicitly contained such public purpose statutes. The implementation of IDB therefore required, in a first instance, a legislative vote of constitutional or statutory provision that authorizes industrial development financing.

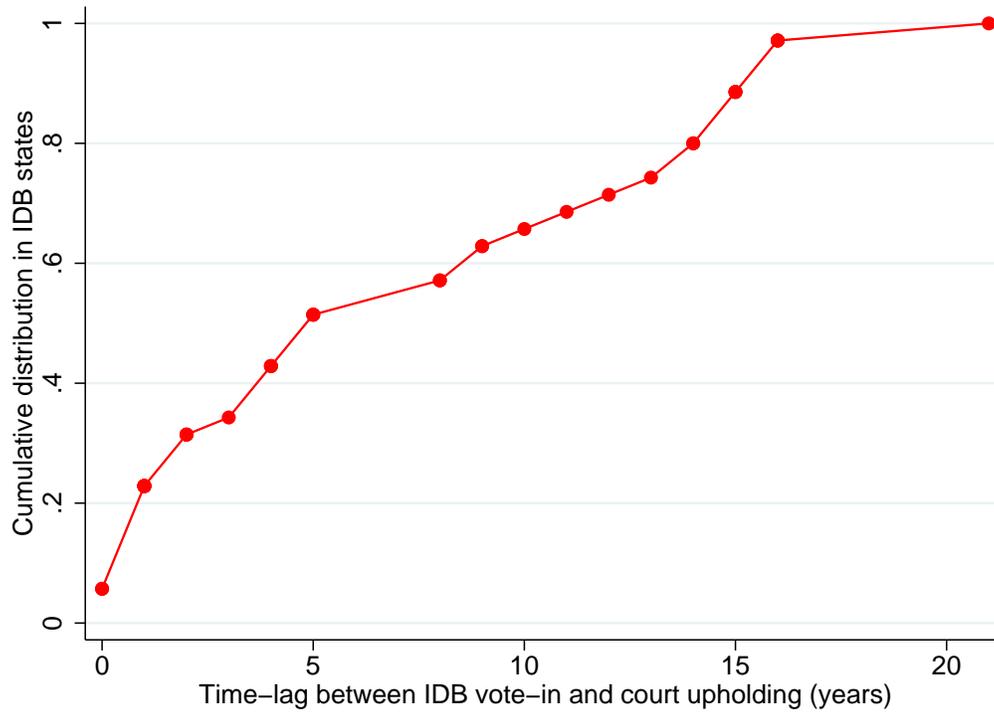
The lack of historical precedent, however, meant that the voted provision required judicial testing in order to be effectively implemented. Indeed, investors were reluctant to hold IDB securities in the period where the legality of the voted state provision had not been confirmed in the state's judicial system (Cobb, 1993). Judicial testing was most often delivered by a court case brought before the State's supreme court. This court case could be triggered in several ways. Most often, the issuance of an IDB required a significant amount of pre-issuance preparation, including a detailed description of the local workforce needs and a justification for why a particular candidate private firm would satisfy

those needs. These preparations were often done by a local government agency, created specifically for this purpose. The case would then be brought against the legality of this local development agency. More generally, any legal step required to issue IDB could be targeted in a court case. In several States, including Tennessee, the IDB statute featured the requirement of a vote of approval by the relevant electors as a special municipal election. The court case could also directly involve the issuance of an IDB bond itself. But as the graphical evidence in Panel A of Figure 7 shows, this was only very rarely the case. In several instances, the fact that IDB were issued as revenue, rather than general obligation, bonds, was the basis of the argument for not violating the 'credit for public purpose' doctrine. In the case of *Wayland v. Snapp*, the Arkansas Supreme Court "(...) chose to uphold the issuance of the revenue bonds by invoking the doctrine that revenue bonds do not violate a credit clause because they are retired through lease revenues of the project, not out of tax funds" (Yale Law Journal, 1961).

I collect information on the dates of the legislative vote and the upholding from several sources. Importantly, I collect information from both administrative sources and legal reviews: Abbey (1965), ACIR (1963), Pinsky (1972), and Economic Development Administration (1978). The date for the vote is the year of appearance of the constitutional statute or provision authorizing local development financing. The date for the upholding event is the publication year of the leading case that upholds the constitutionality of the statute or provision. There is only little conflict in the reported dates of the vote and the upholding between the administrative and legal sources. In the case of upholding, there are sometimes several leading cases, when the first case upholds the constitutionality of the statute allowing cities to issue development bonds, and the later cases extend these powers to counties. I always choose the earliest date across sources for both the vote and the upholding events.

Table B.1 provides information for each IDB state in my time-period of study. The table reports the year of vote and the year of upholding that I use in the main analysis. In Figure B.1, I plot the cumulative distribution of the time-lag between the vote-year and the upholding-year. The average lag is 6.67 years, with a standard deviation of 6.77. In just under 40 percent of States, the time-lag exceeds 10 years. In the main analysis, my estimation is helped by the existence of a significant lag between the vote and the upholding events within state; the variance in lag across States; and, the differential timing of court upholding decisions across States.

Figure B.1: Time-Lag Between Vote and Upholding Events Across IDB States



Notes: This graph displays the empirical cumulative distribution function of the time-lag in years between the vote event and the upholding event within each State that has upheld IDB by 1980. The year of the vote is the year where the State legislature voted in a statute or provision authorizing IDB. The year of the uphold event is the year where the State supreme court upheld the legality of the voted IDB statute or provision through a leading court case. The time-lag is defined as the difference in years between these two events. Source: Section 4.1 and Appendix Section B.2.

Table B.1: Industrial Development Bonds Program Legal Timing

State	Year vote	Year uphold	Leading court case
Alabama	1949	1950	<u>Newberry v. City of Andalusia</u> , 257 Ala. 49, 57 So. 2d 629
Arizona	1963	1973	<u>Industrial Development Authority of Pinal County v. Nelson</u> , 109 Ariz. 368, 509 P. 2d 705
Arkansas	1958	1960	<u>Wayland v. Snapp</u> , 232 Ark. 57, 334 S.W. 2d 633
Colorado	1955	1970	<u>Allardice v. Adams County</u> , 173 Colo. 133, 476 P. 2d 982
Delaware	1961	1962	<u>In re Opinion of the Justices</u> , 177 A. 2d 205
Georgia	1957	1970	<u>In re Opinion on Sub. H. B. 24</u>
Illinois	1951	1972	<u>People ex rel. City of Salem v. McMackin</u> , 53 Ill. 2d 347
Iowa	1963	1964	<u>Green v. City of Mount Pleasant</u> , 131 N.W. 2d 5
Kansas	1961	1962	<u>State ex rel. Ferguson v. Pittsburgh</u> , 364 P. 2d 71
Kentucky	1946	1950	<u>Faulconer v. City of Danville</u> , 313 Ky. 468, 232 S.W. 2d 80
Louisiana	1952	1954	<u>Miller v. Washington Parish</u> , 75 Southern So. 2d 394
Maine	1958	1966	<u>Northeast Shoe Company v. Industrial and Recreational Finance Approval Board</u> , 233 A. 2d 423
Maryland	1960	1974	<u>Wilson v. Board of County Commissioners of Allegheny County</u> , 273 Md. 30, 327 A. 2d 488
Michigan	1963	1966	<u>City of Gaylord v. Beckett</u> , 144 N.W. 2d 460
Minnesota	1961	1970	<u>City of Pipestone v. Madsen</u> , 178 N.W. 2d 594
Mississippi	1936	1944	<u>Albritton v. City of Winona</u> , 178 So. 799
Missouri	1960	1975	<u>Atkinson v. Planned Industrial Expansion Authority of St. Louis</u> , 517 S.W. 2d 36
Montana	1965	1970	<u>Fickles v. Missoula County</u> , 470 P. 2d 287
Nebraska	1960	1962	<u>State ex rel. Meyer v. County of Lancaster</u> , 113 N.W. 2d 63
Nevada	1959	1973	<u>State ex rel. Brennan v. Bowman</u> , 512 P. 2d 1321
New Hampshire	1955	1971	<u>Opinion of the Justices</u> , 278 A. 2d 357
New Mexico	1955	1956	<u>Village of Deming v. Hosdreg Co.</u> , 62 N.M. 18, 303 P. 2d 920
North Dakota	1955	1964	<u>Gripentrog v. City of Wahpeton</u> , 126 N.W. 2d 230
Ohio	1955	1966	<u>State v. Greater Portsmouth Growth Corporation</u> , 218 N.E. 2d 446
Oklahoma	1960	1961	<u>Application of The Oklahoma Industrial Financial Authority</u> , 360 P. 2d 720

Table B.1: Industrial Development Bonds Program Legal Timing (end)

State	Year vote	Year uphold	Leading court case
Pennsylvania	1956	1968	<u>Basehore v. Hampden IDA and Walker v. Butler County IDA</u> , 248 A. 2d 212
Rhode Island	1958	1974	<u>In re Advisory to Governor</u> , 324 A. 2d 641
South Carolina	1962	1967	<u>Elliott v. McNair</u> , 156 S.E. 2d 421
South Dakota	1964	1968	<u>Clem v. City of Yankton</u> , 160 N.W. 2d 125
Tennessee	1951	1952	<u>Holly v. Elizabethon</u> , 241 S.W. 2d 1001
Utah	1953	1968	<u>Allen v. Toole County</u> , 445 P. 2d 994
Virginia	1962	1967	<u>Industrial Development Authority of the City of Chesapeake v. Suthers</u> , 208 Va. 51 155 S.E. 2d 326
West Virginia	1963	1964	<u>State ex rel. Marion County v. Demus</u> , 135 S.E. 2d 352
Wisconsin	1957	1973	<u>Hammermill Paper Co. v. LaPlante</u> , 205 N.W. 2d 784
Wyoming	1963	1967	<u>Uhls v. State</u> , 429 P. 2d 74

Notes: This table provides details on the legal timing of the IDB program in all States. The year of the vote is the year where the State legislature voted in a statute or provision authorizing IDB. The year of the uphold event is the year where the State supreme court upheld the legality of the voted IDB statute or provision through a leading court case. This leading court case is indicated in the final column of the table. The years of the vote and upholding event are drawn from administrative and legal reviews. Source: Section 4.1 and Appendix Section B.2.

B.3 Robustness of main regression results

In this sub-section, I provide robustness checks for the main results in the US states regression setting. In Table B.2, I maintain the same specification as in the main text, but consider alternative measures of the main outcome variables. In the main text, I studied the impact of the upholding event on employment structure using the employee-share of the active workforce. The disadvantage of this variable is that it is interpolated between Census years. In Column 1, as an alternative I use the employee-share of income. This variable is drawn from the SA5H BEA series and is continuous throughout the sample period. The variable is constructed as the ratio of total wages and salaries to total resident income.² Column 1 indicates that the upholding event led to a large and significant increase in the employee-share of income, in line with the main finding of an increase in the employee-share in employment.

The final three columns of Table B.2 investigate the robustness of the absence of a per capita income effect. In column 2, I use the BEA 'net earnings by place of residence', which equals total earnings less contributions for government social insurance plus 'adjustment for residence'. In Column 3, I use the Census-based measure of total personal income. This income measure is interpolated between Census years. In Column 4, I use the IRS-based measure of income, adjusted gross income (AGI), drawn from the top-income share series in Frank et al. (2015). I find an insignificant impact of both the upholding event and the vote-in event across these three alternative measures of income. Both the BEA and Census measures suggest an insignificant positive impact, while the IRS measure suggests an insignificant, but negative impact. The absence of an impact on income at the state-level using various measures is consistent with the regressions in Section B.4 which also fail to detect a per capita income impact, but at the local county-level. The absence of an impact on income in this context is not inconsistent with other place-based program evaluations which have found positive development impacts. Indeed, those findings, including Kline & Moretti (2014), are based on long-run estimates, while my estimates only capture the short-run program impacts.

In Table B.3, I consider the robustness of the impact on employment-structure to alternative specifications. Column 1 replicates the result from the main specification. In Column 2, I remove the vote-in dummy from the main specification. The counterfactual is now entirely built from states that uphold IDB at a later date. This has only a minor impact on the estimated coefficient, which changes from 1.7 percentage points to 1.5 percentage points. In Column 3, I remove the main covariates from the main specification. This has no effect on the estimated coefficient. In Column 4, I augment the main specifica-

²It was also used in the graphical evidence in Figure 7.

tion with additional controls, but this has no impact on the estimates. These additional controls are plausible determinants of employment-structure, but were not included in the main specification because of their potential endogeneity. They are: log per capita income; an indicator for the existence of right to work laws; an indicator for the existence of a corporate income tax; and, a firm-size measure of the state's unemployment insurance. In Column 5, I allow for the determinants of the time-lag between the vote-in and upholding event to have an independent impact. Table 1 showed that the time-lag was shorter in states with civil law origins, and longer in states that had witnessed defaults for a historical public-private funding initiative. While these are state-specific but time-invariant characteristics, they may nonetheless be correlated with state-time varying determinants of employment structure. I therefore allow civil law states and historical default states to be on fully non-parametric time-paths throughout the sample period. Formally, I estimate

$$y_{st} = \beta + \alpha \mathbf{1}(\text{Vote-in})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \gamma_t (\mathbf{1}(\text{Civil Law})_s \times \gamma_t) + \beta_t (\mathbf{1}(\text{Hist Default})_s \times \gamma_t) + \lambda X_{st} + \mu_s + \gamma_t + \varepsilon_{st}$$

where all variables are defined as in the main text, and $\mathbf{1}(\text{Civil Law})_s$ and $\mathbf{1}(\text{Hist Default})_s$ are indicators taking a value of 1, respectively, if a state is has civil law origins or has experienced a historical default. The inclusion of these time-paths only marginally reduces the estimated coefficient on the upholding event.

In Column 6, I investigate the possibility that my main control specification does not adequately capture differential convergence patterns in employment structure over time across states. Indeed, the IDB implementation period was characterized by rapid structural convergence for the less-developed states in the US (Barro & Sala-i-Martin, 1992; Caselli and Coleman, 2001). To investigate this, I augment the main specification with an interaction between a linear time-trend and the cross-sectional level of state GDP per capita in 1940. That is, I estimate

$$y_{st} = \beta + \alpha \mathbf{1}(\text{Vote-in})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \lambda X_{st} + \mu_s + \gamma_t + \beta (\text{Initial_Income}_{s1940} \times [t - 1940]) + \varepsilon_{st}$$

where variables are defined as in the main text, and where $\text{Initial_Income}_{s1940}$ is the cross-section of initial GDP per capita in 1940, which is interacted with a linear time trend $[t - 1940]$. This leads to only a very marginal reduction in the magnitude of the estimated impact of upholding.

Finally, in Column 7 I show that the results are robust to excluding the set of IDB states which were implementing the Tennessee Valley Authority program (Kline and Moretti, 2013), a concurrent Federal place-based development program.

Table B.2: Alternative Measures of Employment and Income

	E-share of income	Avg Income (BEA)	Avg Income (Census)	Avg Income (IRS)
	(1)	(2)	(3)	(4)
1(Vote)	.000 (.004)	23.369 (29.475)	38.807 (43.544)	-684.822 (978.986)
1(Uphold)	.013 (.006)**	.424 (33.114)	59.052 (55.131)	-624.991 (1115.546)
Mean outcome variable	.707	1016	2003	1596
State FE	x	x	x	x
Year FE	x	x	x	x
State-year controls	x	x	x	x
States	28	28	28	28
State-year Obs	466	466	466	466

Notes: This table reports results from estimating the following regression

$$y_{st} = \beta + \alpha \mathbf{1}(\text{Vote-in})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \beta \mathbf{1}(\text{Upheld})_{st} \times \mathbf{1}(\text{EoI})_{st} + \lambda \mathbf{X}_{st} + \mu_s + \gamma_t + \varepsilon_{st}$$

where s denotes state and t denotes time. $\mathbf{1}(\text{Vote-in})_{st}$ indicates whether a vote has occurred in the state-house to allow issuance of IDB but the IDB has not yet been upheld, $\mathbf{1}(\text{Upheld})_{st}$ indicates whether the State court system has upheld the constitutionality of the voted IDB statute or provision. The vote-in and upholding events are mutually exclusive. $\mathbf{1}(\text{EoI})_{st}$ is an indicator variable taking a value of 1 when a State has passed an exchange of information agreement with the Federal Internal Revenue Service. In Column 1, the outcome variable is the wage and salary share of individual income, drawn from BEA historical data. In Column 2, the outcome variable is 'net earnings by place of residence', which is the BEA concept of personal income. It is equal to total earnings less contributions for government social insurance plans plus a residence adjustment. In Column 3, I use the measure of gross personal income from the decennial Census. This measure is interpolated between Census year, using a natural cubic spline. In Column 4, I use the IRS measure of income, adjusted gross income, which is drawn from the top income share series (Frank et al., 2015). The state-year controls, \mathbf{X}_{st} , are indicator variables for election year, and indicator variables for the existence of voting restrictions in the form of poll tax and literacy tests. These are the first stage instruments used by Besley et al. (2010) to study political competition and policy-making in US states. *, **, *** denote significance at the 10%, 5%, 1% level. Robust standard errors clustered at the state level in parentheses. Source: Section 4.5 and Appendix Section B.3.

Table B.3: Alternative Specifications

	Employee-share of employment						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Vote)	.003 (.005)		.004 (.005)	.003 (.005)	.004 (.006)	.004 (.005)	.004 (.006)
1(Uphold)	.017 (.005)***	.015 (.005)***	.017 (.005)***	.017 (.005)***	.015 (.006)**	.016 (.005)***	.018 (.006)***
Specification	Main	Cross-sectional only	No controls	Extensive controls	Time path civil law states	Initial income time-trend	Exclude TVA states
Mean outcome variable	.707	.707	.707	.707	.707	.707	.777
State FE	x	x	x	x	x	x	x
Year FE	x	x	x	x	x	x	x
States	28	28	28	28	28	28	24
State-year Obs	466	466	466	466	466	466	409

Notes: this table reports results from estimating alternative specifications, described in detail in Section B.3. In all regressions, the outcome variable is the employee-share of the economically active workforce. Column 1 replicates the central finding from estimating the main specification (1). Column 2 removes the indicator variable for the vote event from the main specification. Column 3 removes the controls from the main specification. Column 4 augments the main specification with additional controls: log per capita income; an indicator for the existence of right to work laws; an indicator for the existence of a corporate income tax; and, a firm-size measure of the state's unemployment insurance. Column 5 augments the main specification with a full set of year indicator interactions with both the indicator for civil law origins and the indicator for historical rail default. Column 6 augments the main specification with an interaction between a linear time-trend and the cross-section of GDP per capita in 1940. Finally, Column 7 estimates the main specification, but on a reduced sample which excludes the four States (Alabama, Kentucky, Mississippi, Tennessee) which were part of the Tennessee Valley Authority development program. *, **, *** denote significance at the 10%, 5%, 1% level. Robust standard errors clustered at the state level in parentheses. Source: Section 4.5 and Appendix Section B.3.

B.4 Local IDB impacts: evidence from county-level regressions

In this subsection, I study the impact of the IDB program at the county-level. While the effective start of the program is triggered at the level of the state supreme court (Sections 4.1- 4.2), the decision itself to issue IDB is predominantly made by counties within the state.³ As such, a county-level analysis provides an assessment of the direct local economic impacts of the program. At the same time, the county-level analysis at the level of local implementation helps to shed light on the absence of an economically meaningful impact of IDB on non-employment outcomes.

In order to study the county-level impacts of IDB, I rely on two main data-sets. The first is the comprehensive county-level panel data-set ICPSR 2896 'Historical, Demographic, Economic, and Social Data: The United States, 1790-2002.' This data-set has been used in other studies of long-run impacts of local place-based development programs, including the Tennessee Valley Authority program (Kline & Moretti, 2013). I interpolate values between data-points in the ICPSR data-set. Note that during my period of interest, the primary source in ICPSR is the County Data Book. Since this source delivers data every five years, the interpolation period is smaller than between decennial Census which is used in the main analysis. Nonetheless, the data has the disadvantage that it does not contain a continuous measure of per capita income, which is the main object of interest in this county-level analysis. I therefore supplement it with a second county-year panel data-set. This data-set is the combination of the continuous BEA county-level per capita income data, which exists from 1969 onward; and, the 1959 Census module which measures per capita income in the cross-section of all counties. While I do have to interpolate per capita income between 1959 and 1969, this data-set nonetheless gives me a more naturally continuous measure of per capita income than the ICPSR data-set. The only disadvantage is that my sample only starts in 1959, while 7 states have voted in the IDB program before that date. The county-level analysis is therefore limited to the counties in the 21 states that vote in after 1959.

The aim is to investigate the impacts of the IDB by comparing counties with the program to counties without the program. The basis for this exercise remains the specification used in the main text, which assesses impacts by comparing changes before and after the upholding event, while controlling for any impact occurring during the vote in event. But without any additional modifications, this specification would rely on counties in different states as a counterfactual. Instead, I want to create a control county within the same state. This is a meaningful exercise since the IDB program was initiated at the

³The decision could also be made by higher tiers of government, such as the the state government, or lower tiers of government, including cities. Data from Moody's Investor Service (1974) suggests that in practice, actual issuance was predominantly carried out by counties.

county level and only a subset of counties in a given state would initiate IDB.

I assign treatment at the county-level within the state based on a list created by the federal government before IDB had become widely implemented. The Area Redevelopment Administration (ARA) is a federal agency that was created in 1961, with the aim of providing technical (data-driven) assistance to state and local governments to implement local development financing. For this purpose, the ARA created criteria that defined ‘redevelopment areas.’ These were predominantly rural geographical areas, characterized by “structural underemployment”, where the encouragement of new industries was perceived as a solution to the stagnant levels of development (ARA, 1961). This characterization is effectively identical to the characterization of counties that IDB was targeted towards. In every state, the ARA compiled data from Census and the Departments of Health, Education, Welfare, and Agriculture, to establish a statistical profile of every county in 1961, and classify a subset of those as ‘redevelopment areas.’

I digitize the list of ‘redevelopment’ counties based on the ‘Statistical Profiles’ in every state (ARA, 1961), and merge it with the main county panel data-set. This list has the appealing feature that it was created by a government entity which was not responsible for implementing IDB in the pre-IDB period. As such, the selection of counties into the list may be considered plausibly exogenous to unobservable county-time varying confounding determinants of local development.

I augment the empirical specification used in the main text with this list to create a difference-in-differences design. More specifically, I consider the ARA ‘redevelopment’ status to be a county-specific time-invariant assignment to program treatment. Since there exists counties on the ARA list that do not take up IDB, and there exists counties not on the ARA list that can take up IDB, this is an intent-to-treat design. The diff-in-diff evaluation will compare changes in outcome in ARA counties before and after the court upholding event to changes in outcomes in non-ARA counties within the same state, while controlling for any impacts that occurred during the vote-in event. Formally, I estimate

$$y_{cst} = \beta + \alpha \mathbf{1}(\text{Vote-in})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \pi (\mathbf{1}(\text{Vote in})_{st} \times \mathbf{1}(\text{ARA})_c) \quad (1)$$

$$+ \phi (\mathbf{1}(\text{Upheld})_{st} \times \mathbf{1}(\text{ARA})_c) + \mu_c + \gamma_t + \varepsilon_{cst}$$

where y_{cst} is the outcome of interest in county c , in state s , at time t , $\mathbf{1}(\text{Vote-in})_{st}$ and $\mathbf{1}(\text{Upheld})_{st}$ indicate whether a state has, respectively, voted in but not upheld or upheld the IDB program. $\mathbf{1}(\text{ARA})_c$ is a county-specific, time-invariant indicator that takes value 1 if a county is on the ARA federal list of redevelopment areas, and μ_c and γ_t are county and year fixed effects, respectively. I cluster the standard error at the state level, to allow

for spill-over between ARA and non-ARA counties within the same state. The time-window is identical to the one used in the main estimation (Tables 2-3): in every IDB state, I consider the time-period that ranges from 5 years before the vote-in event to 5 years after the upholding event.

The results are displayed in Table B.4. In the first column, I study the employee-share of the active workforce as the outcome variable. I find that the large, positive impact is concentrated in the ARA counties in the upholding period. In the following two columns, I find no overall impacts on the level of employment and urbanization. The absence of impacts on these two outcomes is consistent with the interpretation that the IDB program achieved its stated objective of reducing underemployment in specifically targeted rural areas. Issuance of IDB required documenting the specific local industrial needs of a county and a justification for why the size and characteristics of the proposed IDB facility would achieve this local need. In comparison to other place-based development programs, IDB was therefore highly targeted in nature and narrow in scope, aiming to finance industrial development commensurate with the specific local workforce needs. The increase in the employee-share and the absence of an impact on the size of the workforce suggests IDB primarily provided a transition into employee-jobs of 'underemployed', self-employed farmers. The absence of any change in urbanization suggests that workers did not migrate to the predominantly rural areas where the IDB facilities were being opened. Consistent with the absence of generalized economic impacts, the final two columns find no statistically significant impacts per capita income. The fourth column uses family income, measured in the ICPSR data-set, while the fifth column uses the continuous BEA per capita income measure. The impact on the continuous measure of per capita income is particularly insignificant, both statistically and economically.

Taken together, these county-level results provide additional evidence to support the absence of any meaningful non-employment development impacts, in the short-run 5-year window considered in this estimation strategy. In particular, the IDB program seems to have led to a significant transition from self-employment to employee-jobs but only locally in the specifically targeted IDB counties. The absence of any spill-over to work structure or workforce attachment in non-treated counties suggests sectoral re-allocation was limited. The absence of any change to levels of urbanization suggests migration from non-treated to treated counties was also limited. The compensation for migration costs and the efficiency gains from sectoral re-allocation are two of the main mechanisms through which previous studies have found long-run positive income impacts of place-based programs. These mechanisms seem to not be significant forces in the IDB context in the 5-year short run.

Table B.4: County-Level Evidence on Local Impacts of IDB

	E-share	Employment	Urbanization	Log(Family Income)	Log(Personal Income)
	(1)	(2)	(3)	(4)	(5)
1(Vote)	-0.007 (.006)	-0.003 (.002)	-0.003 (.003)	-0.016 (.030)	.031 (.043)
1(Vote)*1(ARA)	.003 (.007)	-0.001 (.003)	-0.001 (.009)	-0.007 (.033)	-0.019 (.038)
1(Uphold)	-0.009 (.008)	-0.004 (.003)	-0.003 (.005)	-0.044 (.047)	-0.003 (.072)
1(Uphold)*1(ARA)	.031 (.010)***	-0.002 (.004)	.000 (.010)	.044 (.040)	-0.008 (.051)
County FE	x	x	x	x	x
Year FE	x	x	x	x	x
States	21	21	21	21	21
County-year obs	5140	5140	5140	5140	5140

Notes: This table reports the results from estimating the following regression

$$y_{cst} = \beta + \alpha \mathbf{1}(\text{Vote in})_{st} + \theta \mathbf{1}(\text{Upheld})_{st} + \pi (\mathbf{1}(\text{Vote in})_{st} \times \mathbf{1}(\text{ARA})_c) + \phi (\mathbf{1}(\text{Upheld})_{st} \times \mathbf{1}(\text{ARA})_c) + \mu_c + \gamma_t + \varepsilon_{cst}$$

where y_{cst} is the outcome of interest in county c , in state s , at time t , $\mathbf{1}(\text{Vote in})_{st}$ and $\mathbf{1}(\text{Upheld})_{st}$ indicate whether a state has, respectively, voted in but not upheld or upheld the IDB program. $\mathbf{1}(\text{ARA})_c$ is a county-specific, time-invariant indicator that takes value 1 if a county is on the ARA federal list of redevelopment areas, and μ_c and γ_t are county and year fixed effects, respectively. The outcome variables are: employee-share of employment; economically active employment share of population; urbanization share of population; log family income; and, log personal income. The first 5 outcomes are drawn from ICPSR 2896, while the final outcome variable is constructed from historical BEA series and the 1959 Census. The sample is limited to counties in the 21 IDB States that vote in IDB statutes or provisions after 1959. *, **, *** denote significance at the 10%, 5%, 1% level. Robust standard errors clustered at the state level in parentheses. Sources: Appendix Section B.4.

References

- [1] Area Redevelopment Administration [1961]. *Statistical Profiles for Rural Redevelopment Areas. All states*. Washington D.C: National Archives.
- [2] Barro, R. J., X. Sala-i-Martin [1992]. "Convergence", *Journal of Political Economy*, 100(2).
- [3] Caselli, F., J. Coleman [2001]. "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation", *Journal of Political Economy*, 109(3).
- [4] Council of State Governments [years between 1939 and 2009]. *Book of the States, all Parts*. Washington DC: The Council of State Governments.
- [5] Dusek, L., S. Bagchi [2017]. "Third-Party Reporting, Tax Collections, and the Size of Government: Evidence from Withholding", *Working Paper*.
- [6] Economic Development Administration [1978]. *The Industrial Revenue Bond as a Financial Attraction Device*. Office of Economic Research, Washington D.C.
- [7] Frank, M. E. Sommeiller, M. Price, E. Saez [2015]. "Series for Top Income Shares by US States since 1917", *WID Technical Note Series*, No.2015-7.
- [8] Jordan, M., M. Grossman [2017]. *The Correlates of State Policy v.2.1*. East Lansing, MI: Institute for Public Policy and Social Research.