Tax Productivity and Economic Development: A Quantitative Macroeconomic Analysis

Radhika Goyal
IMF

Anders Jensen
Harvard University and NBER

David Lagakos
Boston University and NBER

Abdoulaye Ndiaye
New York University

Yuntian Zhao
New York University

June 2026

Abstract

This paper examines the role of government capacity to collect taxes in driving economic development. We first draw on new evidence from 100 countries to document that tax collections relative to tax administration expenditures rise sharply with GDP per capita. We then interpret this finding in a neoclassical growth model with a tax administration sector in which labor inputs are required to raise tax revenues. Countries differ in their levels of tax productivity, capturing how effectively labor is used to raise taxes, as well as their levels of productivity in creating consumption and investment goods. Governments use tax revenues to provide public capital that complements private factors in production and to make transfers. Quantitatively, our model predicts that increases in tax productivity in low-income economies can raise GDP per capita substantially more than suggested by Hulten's theorem.

Email: rgoyal3@imf.org, Anders_Jensen@hks.harvard.edu, lagakos@bu.edu, an1490@stern.nyu.edu, and yz11552@stern.nyu.edu. For outstanding research assistance we thank Negin Jafari. For helpful comments we thank Augustin Bergeron, Corina Boar, Gustavo Ventura, Xavier Gabaix, Emilien Gouin-Bonenfant, Roozbeh Hosseini, Jordan Norris, Steven Pennings, Joachim Voth, James Robinson, Florian Scheuer, Tao Zha, and seminar/conference participants at Amherst College, ASU, the Atlanta Fed, Copenhagen, Houston, Nanyang Technical University, Nottingham, NYU, the 2025 NBER Summer Institute, Tufts, USC, UC Irvine, the World Bank, the Tax and Development Conference at Williams College. All potential errors are our own.

1 Introduction

One of the most prominent theories of long-run economic development places state capacity — and, in particular, the ability of governments to raise tax revenue — at the heart of the development process (see, e.g., Besley and Persson, 2011). Historical work by Mayshar, Moav, and Pascali (2022) shows that the most successful early states, such as Egypt, were those best able to raise revenue to fund public works and military ventures. Tilly (1990), Dincecco and Katz (2016), and Gennaioli and Voth (2015) provide evidence that the constant warfare of European states in the 16th and 17th centuries led them to build up state capacity, which may ultimately have laid the groundwork for their later development. Recent empirical work by D'Arcy, Nistotskaya, and Olsson (2024) ties the historical creation of public land records—the foundation of property taxation—to long-run growth.

By contrast, the link between a government's ability to collect taxes and its long-run level of development has received less attention in the macroeconomics literature. The foundational papers of Aschauer (1989), Barro (1990), and Baxter and King (1993), and incorporate public capital stocks into the production function but abstract from the costs of collecting taxes. Tax administration costs are likewise absent from nearly all analyses of Ramsey taxation (Judd, 1985; Chamley, 1986; Chari and Kehoe, 1999; Conesa, Kitao, and Krueger, 2009; Straub and Werning, 2020) and models of fiscal multipliers (Ramey, 2011; Farhi and Werning, 2016). A notable exception is the paper by Besley, Ilzetzki, and Persson (2013) which incorporates the political economy of tax capacity into a growth model. They show that more severe fiscal constraints can lead an economy to reach a less desirable steady state, although they do not pursue the quantitative implications of their model.

This paper assesses the quantitative importance of governments' capacity to collect taxes, which we refer to as *tax productivity*, for economic development. Methodologically, we follow the growing literature on macroeconomic development emphasizing the role of sector-specific productivity in long-run growth, such as agriculture (Restuccia, Yang, and Zhu, 2008; Caselli, 2005; Boppart, Kiernan, Krusell, and Malmberg, 2023), investment goods (Hsieh and Klenow, 2007; Herrendorf and Valentinyi, 2012), or electricity (Colmer, Lagakos, and Shu, 2024). The main novel feature of our model is a tax administration sector that requires labor inputs (e.g. "tax collectors") to raise tax revenues. Countries differ along two main dimensions. The first is economy-wide productivity, which governs their ability to produce consumption and investment goods. The second is tax productivity—the efficiency of labor in raising tax revenue—which determines the government's capacity to provide public capital, a complement to private factors of production. We measure tax productivity across countries and quantify the long-run aggregate effects of raising it.

We begin by creating a new dataset of tax collections and tax administration expenses in a broad set of countries. We hand-collect these statistics from national tax authorities—such as the U.S. Internal Revenue Service and Ghana Revenue Authority—for 100 countries over multiple years, yielding 625 country-year observations. Our outcome of interest is the ratio of total tax revenues to collection costs, representing the average amount collected per dollar spent on collection. This unit-free measure is intuitive and comparable both across countries and over time within countries. As we argue in the quantitative analysis below, it is also informative about the level of tax productivity in each country.

Our main empirical finding is that tax collections per unit of expenditure are strongly increasing in GDP per capita. The poorest countries in the world collect about $30 in revenue for every dollar spent on collection; this number rises to around $270 in the richest. We provide evidence that this pattern is not an artifact of differences in the scope of the revenue authorities' activities, such as whether they are responsible for VAT taxes or not. Exploiting the panel dimension of our data, we also examine how collections per expenditure vary with GDP per capita within countries over time. We find suggestive evidence that countries tend to increase their collections per expenditure as they grow, with an elasticity that is comparable to the cross-section. These patterns also hold for individual sub-national tax authorities over time in the United States, with less support from Indian states, based on newly constructed state-level panel data.

We interpret these facts using a neoclassical growth model that includes productive public spending and a tax administration sector. Following Baxter and King (1993) and related work in macro public finance, production combines labor, private capital, and public capital into a single output good usable for consumption or investment. Governments use tax revenues to invest in new public capital or to fund transfers. Households make standard consumption, savings, and labor-supply decisions subject to marginal taxation of their labor income and taken transfers as given.

We make three modeling choices about taxation that help the model capture public finance systems across different levels of development. First, rather than having the government simply announce a tax rate, we posit that the effective tax rate depends on both a country's tax collection productivity and the number of tax collectors employed. Second, rather than modeling each type of taxation explicitly, and the rising difficulty of collecting more from any single type, we allow for decreasing returns to scale in tax collection. This choice is consistent with evidence that expanding tax collection requires moving from easy-to-collect sources, such as taxes on international trade, to more challenging ones, such as property taxes and income taxes on the self-employed (e.g. Jensen, 2022). Third, we allow some fraction of tax collections to be lost as waste, with this fraction varying exogenously with the level of development to capture potentially higher waste in public expenditure in poorer economies.

We use the model to measure tax productivity in each country, and to simulate the long-run effects of raising it. The magnitude of these effects is not obvious: the tax collection sector employs well under one percent of the workforce in every country, and a naive application of Hulten's theorem would predict negligible aggregate gains from raising its productivity. To move beyond this back-of-the-envelope logic, we calibrate the model to match key features of the data, including GDP per capita, the ratio of tax collections to expenditure, the tax-to-GDP ratio, the employment share in tax collection, and the share of taxes devoted to transfers. One important parameter is the elasticity of output with respect to public capital, which we set in our main specification at 0.2, based on evidence from a meta-analysis. A second is the degree of decreasing returns to scale in tax collection, which we choose to match evidence on the returns to expanding collection efforts in the United States and several developing countries. As we show later, both parameter values strongly shape how much tax productivity matters for development.

We show that the model does reasonably well at matching the targets as well as several important moments not targeted in the parameterization. Private capital-to-labor ratios rise almost linearly with GDP per capita in both the model and data, with nearly identical levels. Empirical measures of public capital are far less common in macroeconomics and are complicated by a basic problem: real public investment is typically equated with public expenditure (Pritchett, 2000; Cubas, 2020). Our model reproduces the overall slope with respect to income per capita, though it under-predicts the measured stocks for lower-income countries. The model's investment share of GDP is increasing in development, as in the national account statistics, though with a higher level than in the data. Labor supply is modestly decreasing in development, consistent with the cross-country and time-series evidence (Ramey and Francis, 2009; Bick, Fuchs-Schündeln, and Lagakos, 2018; Boppart and Krusell, 2020).

The implied tax productivity estimates exhibit even more variation with income per capita than do aggregate productivity estimates. In our preferred estimates, a regression of tax productivity on log GDP per capita has a slope of around 0.66; by comparison, a similar regression yields a slope of 0.30 for aggregate productivity. We take this as evidence that the tax collection sector is particularly unproductive in poor countries, consistent with the broad claims of the state capacity literature. Indeed, our estimates are arguably more direct proxies for state capacity than the commonly used tax-to-GDP ratio.

The model predicts that a 50 percent increase in tax productivity raises GDP per capita in low-income countries by about 18 percent on average. Consumption rises by 16 percent, and labor supply falls by under one percent. Much of the gain comes from the direct increase in public capital per worker, and private capital accumulation also helps expand output. By contrast, Hulten's approximation theorem suggests GDP gains of less than 0.1 percent given the tax collection sector's very modest share of economic activity. The model predicts that middle-income countries gain far less from the same percent increase in tax productivity, and high-income countries actually lose in consumption terms. Two main forces explain this weaker response in more developed economies. First, public capital expands less in richer countries, which leads to smaller increases in wages. Second, the tax burden increase is far more onerous, lowering after-tax wages and depressing labor supply.

We conclude that improvements in tax collection productivity in the world's poorest economies likely lead to significant positive long-run gains in GDP and consumption. We also provide evidence that public capital stocks in lower-income countries are inefficiently low relative to what a benevolent planner would choose. Our results are sensitive to choices of two main parameters: the output elasticity of public capital and returns to scale in the tax administration sector. We are currently working on extensions to include a traditional sector, that cannot be taxed at all but does not benefit from public capital, and endogenous taxation a la Ramsey.

Related Literature. Our paper builds on the macroeconomic literature in which taxation funds productive public capital (see Glomm and Ravikumar, 1992; Jones, Manuelli, and Rossi, 1993, in addition to the references cited above). Our work is particularly inspired by Ramey (2021)'s insight that the long-run effects of public capital provision can be sizeable particularly when initial levels of public capital start below efficient levels. Whether this is the case in low-income countries today has received much less attention. Our paper is also closely related to work of Cubas (2020) arguing that public capital should be separated out from private capita in development accounting.

Most macro public finance work has been focused on advanced economies. Some of the more related papers include Leeper, Walker, and Yang (2010)'s neoclassical analysis of the U.S. American Recovery and Reinvestment Act, and Bhandari, Evans, Yao, and McGrattan (2024)'s model of tax evasion, calibrated to evidence on the impacts of tax audits. To our knowledge, no prior paper in this literature has modeled a tax administration sector; for studies of advanced economies, where tax collection comes at low cost, this may not be such a grave omission.

Finally, our paper adds to the largely microeconomic literature on public finance and development focused on tax capacity and its determinants (Gordon and Li, 2009). Paths to increasing tax capacity emphasized by this literature include incentive pay for tax collectors (Khan, Khwaja, and Olken, 2015), collaboration with traditional leaders, (Balan, Bergeron, Tourek, and Weigel, 2022), information technology (Hjort and Tian, 2024; Okunogbe and Tourek, 2024), including using georeferenced databases of taxpayers (Dzansi, Jensen, Lagakos, and Telli, 2025), and self-enforcing tax systems (Naritomi, 2019). Much of the recent literature has focused on the taxation of properties, which holds vast promise but is thought to be vastly under-collected in the developing world (Brockmeyer, Estefan, Arras, and Suárez Serrato, 2023). The same is true of taxes on labor income, particularly given the challenges in collecting from the self-employed (Jensen, 2022). Our study contributes by modeling and quantifying the long-run general equilibrium consequences of improving tax capacity, which is new to this literature.

2 Data and Empirical Findings

We begin by presenting the data we use in our analysis and defining the ratio of taxes collected to expenditures. We then investigate how this measure of collection per dollar of expenses varies with income per capita across space and time - between countries today, within countries over the long run, and at both the national and sub-national levels.

2.1 Data

In this subsection, we describe the data collection exercise undertaken to measure the ratio of tax revenue to tax administration costs in a large sample of countries, years, and levels of government. As the theory will make clear (Section 3 onward), this measure is closely related to, but also distinct from, tax productivity. It is best understood as the average (*not* marginal) tax collection per dollar of administrative expenditure for a given authority at a point in time. The measure is unit-free and comparable both across countries and, within countries, over time.

We first describe the numerator: taxes collected at the national level. The numerator captures taxes collected at the national level—typically taxes on personal and corporate income; property, wealth, and inheritance; and the consumption of goods and services.We also include taxes collected at the border, which are a major revenue source in many developing countries. Throughout, we use actual collections rather than projected or estimated figures, and, where the data permit, we exclude non-tax revenue.

The denominator measures the annual expenditure of the national tax authority—specifically, the sum of operating and capital expenditures. In the first category, by far the most important item is salary; other items include training and miscellaneous administrative expenses. The second category includes expenses that go towards purchasing and maintaining buildings and land, as well as other types of infrastructure. Expenses in information and communication technology (ICT) can be included in the first or second category, depending on whether they are long-term investments (e.g., purchase of a new server) or day-to-day expenses (e.g., software maintenance). Where this breakdown is available in our sample, salary and ICT are by far the largest cost categories. Whenever the data permit, we include as many of these cost categories as possible. Salaries are included in all settings; in many, though not all, the measure is more comprehensive and also captures capital expenses.

There are many similarities in tax authorities across countries, and this definition of collection cost allows us to capture most of them in a manner that is meaningfully comparable across borders and over time. Notwithstanding, there are limits to this measure. First, there are differences across countries in the range and nature of revenues administered. Here, the most significant distinction is probably between taxes administered by the customs authority (tariffs and other border taxes) versus taxes administered by the domestic tax authority. Because our measure captures the combined cost of collecting taxes both domestically and at the border, however, these differences in institutional assignment should not affect our ratio. In some countries, direct taxes are administered nationally while certain indirect taxes are administered sub-nationally—for instance, sales tax collection is devolved to states in both India and the United States. Second, some tax authorities are responsible for carrying out activities that are not directly related to taxation. For example, some tax authorities are also responsible for collecting social security contributions and disbursing transfers. Some, though not many, tax authorities are responsible for hosting and maintaining the population registry. We investigate the importance of these variations in institutional design across tax authorities in our regressions. Third, there are naturally other factors that influence this ratio and that are not related to tax productivity: for example, macroeconomic changes could, at least temporarily, create changes in the taxable base without corresponding adjustments to collection expenditures.

Cross-country data. We collect data from several sources in a cross-section of countries. The most important source is the International Survey on Revenue Administration (ISORA), which is administered collaboratively by multiple international agencies, including the OECD, the IMF, and the Inter-American Center of Tax Administrations. Whenever available, we use data on both the numerator and denominator from this source. We complement this with data from individual countries based on the yearly government budget and national reports from the tax authority and national statistics office. For each country, we seek to collect multiple years of data between 2014 and 2017; we report the yearly average in that range to smooth out fluctuations that occur in the ratio due to abnormal levels of expenses or collection in any given year.

Our resulting cross-country sample contains data from 100 countries. The countries span a broad range of development levels: from lower income countries including Malawi, Kenya, Tanzania, Nigeria, Ghana, Bangladesh, Pakistan and India, to middle income countries including Indonesia, Thailand, China, South Africa, Brazil, Colombia, Mexico and Turkey, to high-income countries including South Korea, Russia, UK, Germany, France, Scandinavian countries and the United

States. Together, these countries account for roughly 77 percent of the world's population and 67 percent of global GDP.

Historical country panel data. For a subset of countries, historical data exist that allows us to construct measures of tax productivity over the long run within countries. We follow the same guidelines for data collection as for the cross-section of countries.

Our longest time series comes from the United States, where we can measure collection per dollar of administrative expenses back to 1867. We collect data separately for domestic taxes and border taxes. In the former case, we rely on annual reports published on a yearly basis by the Commissioner of the Internal Revenue Service (IRS). In the latter case, we rely on the annual reports of the Treasury on the state of the finances and data from the US Customs Service and the US International Trade Commission. In the historical years, we digitize the relevant data and harmonize it with the pre-existing digital sources for years closer to the present time. In the US case, we are helped by the fact that the underlying sources are often published in a highly consistent manner across a large range of years. We measure tax collections as gross collections before refunds are issued, but after including penalties and interest paid towards outstanding taxpayer arrears. The expenditure measure is comprehensive and includes all expenditures associated with the process of collecting internal revenue. Throughout the entire period studied, personnel costs (including for central officers and field agents) have been an important cost-component. Over time, other cost components have become gradually more important, including physical infrastructure and ICT.

We collect historical data for a sample of other countries around the world. The resulting sample consists of an unbalanced panel of countries: Australia (1972-2017); Bangladesh (1964-2015); Botswana (1971-2005); Canada (1924-2008); Denmark (1959-2006); India (1946-2016); Kenya (1962-1995); Malawi (1964-1996); Pakistan (1949-2005); South Korea (1987-2016); Sierra Leone (1964-1996); South Africa (1960-2007); Thailand (1959-2017); United Kingdom (1993-2017); United States (1877-2016); Uruguay (1980-2017). The median number of years per country is 52, which constitutes a meaningfully long time series.

For the historical panel, we rely on sources that are specific to each country. In most cases, constructing the series for a country required us to scan, digitize, and clean historical publications that existed on paper.[1] For example, in countries including Denmark, Bangladesh, Kenya, and Thailand, we digitized the country's entire series of annual reports produced by the national statistics office or the ministries of planning and economic development. In countries including Canada, India, and Pakistan, we digitized the country's full set of annual government finance account pub-

[1]Most of these physical publications were available in the Lamont Library at Harvard University.

lications.

Historical sub-national panel. Up to this point, we have focused on collecting measures of tax collection and expenses at the national level for the national tax authority. However, many countries around the world implement a decentralized taxation system, where the mandate to design, implement, collect, and enforce certain taxes is devolved to the sub-national level. In these cases, the sub-national governments are also in charge of establishing their own tax authority, and, in principle, the concept of tax collection per dollar of expense applies at this level as well.

We create long-run panel data on tax productivity at the sub-national level for Indian and US states. We focus on the states because they are in charge of collection for the sales tax - a substantial source of taxation that is collected by the national tax authority in most other countries.[2] In the Indian case, we rely primarily on the 'Combined Finance and Revenue Accounts of the Central and State Governments' series published by the Comptroller and Auditor-General, which we digitize and harmonize for our purposes. In the US case, an important source of data was the historical series 'Book of the States' that is published annually by the Council of State Governments.

The resulting unbalanced panel dataset at the sub-national level contains measures of tax collection and expenses for 28 states in India between 1950 and 2017 and for all 50 U.S. states between 1993 and 2016.

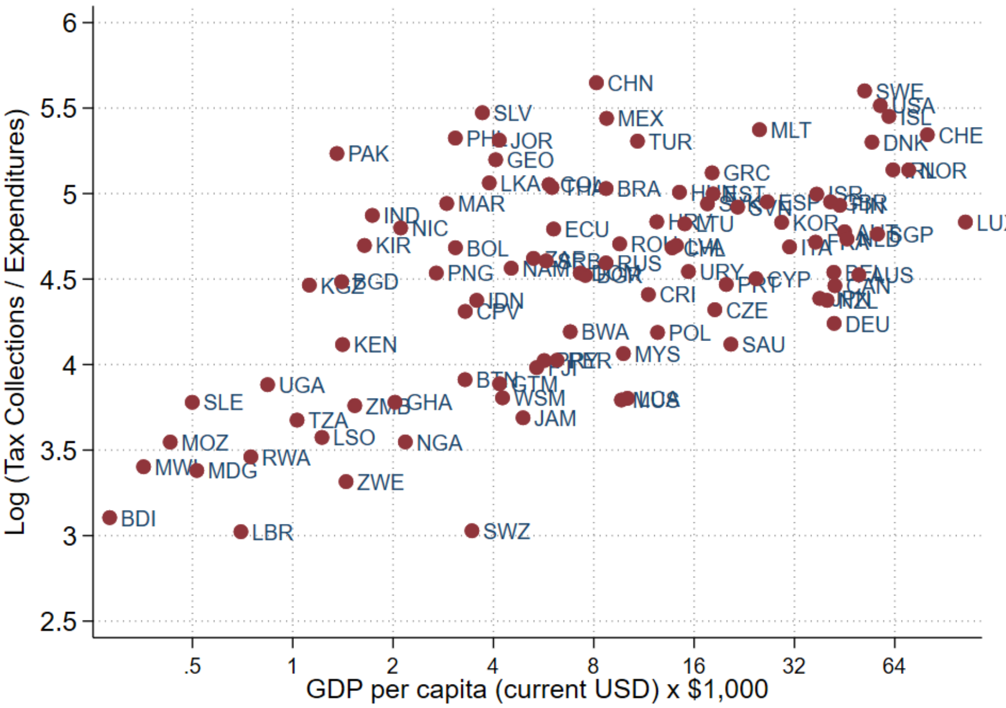
2.2 Results: Tax collections per expenditure and development

We begin by investigating the relationship between the average collection per expenditure and development in the cross-section. Figure 1 shows the association across countries between the (log) of GDP per capita in USD and the log of the ratio of taxes collected to collection expenses. There is a clear positive association whereby tax authorities in countries with higher GDP per capita have a higher average collection per dollar of tax expenditure.

The magnitude of the differences in collections per dollar of expenditure between the world's poorest and richest countries is substantial. Countries like Burundi, Liberia, and Malawi collect between \$20 and \$30 for every dollar they spend collecting taxes. The world's highest values are primarily in high-income countries such as Sweden and the United States, which collect between \$250 and \$270 in taxes for every dollar spent. China, which is situated around the middle of the

[2] Lower levels of sub-national government, such as the municipality or county, are also often in charge of implementing specific taxes (most importantly property tax) and implement their own tax authority to do so. Collecting measures of tax productivity at this level would be interesting but is also practically challenging since most countries do not systematically gather and centralize the required data from the (sometimes large) set of municipalities.

Figure 1: Tax Collections Per Expenditure in the Cross-Section of Countries



Notes: This figure plots the (log) ratio of taxes collected relative to expenses on tax collection against GDP per capita in a sample of 100 countries. Data is collected for each country between 2014 and 2017, and the average is taken over the years if multiple data points exist per country.

GDP per capita distribution, performs remarkably well ($283) for its level of development. Other middle-income countries typically collect in the range between $80 for every dollar spent, as in Indonesia, and $150, as in Brazil.[3]

Table 1 investigates extensions and robustness checks for the pattern uncovered in the cross-section. Column (1) reports the slope coefficient (0.248) that corresponds to the full sample in Figure 1. Columns (2) and (3) estimate the relationship separately in low and lower-middle-income countries and upper-middle and high-income countries, respectively. The slope between tax collection to expenses and (log) GDP per capita is larger at lower levels of development (0.654 versus 0.225), but it remains significant in both samples.

In column (4), we control for the countries' statutory tax rates, using data from the global tax rate database of Bachas, Fisher-Post, Jensen, and Zucman (2024). We include as controls the top personal income tax rate, the main corporate income tax (CIT) and value-added tax (VAT) rates,

[3]The relationship generally appears to hold within regions of the world, with stronger slopes in Sub-Saharan Africa and Europe and Central Asia, and flatter slopes in South Asia, and the Middle East and North Africa.

Table 1: Cross-Section of Countries: Tax Collections per Expenditure and Income Per Capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome	Log T/X	Log T/X	Log T/X	Log T/X	Log T/X	Log T/X	Log T/X
Log GDP p.c.	0.248***	0.654***	0.225***	0.221***	0.249***	0.260***	0.218***
	(0.0389)	(0.142)	(0.061)	(0.0586)	(0.0395)	(0.0398)	(0.0484)
N	100	35	65	99	97	90	73
R^2	0.293	0.390	0.174	0.321	0.296	0.326	0.222
Modif. to column (1)		Only low, lower-middle inc. countries	Only upper-middle, high inc. countries	Cont. for PIT, VAT, CIT, tariff	Excl. if no VAT or sales tax	Excl. if no customs	Excl. if include soc. sec.

Notes: This table provides regressions of the (log) ratio of taxes collected to expenses on collection on log GDP per capita. The sample is a cross-section of 100 countries between 2014 and 2017 (averaged at the country level when multiple data points exist within a country). Column 1 reports the basic OLS, which corresponds to Figure 1. Column 2 restricts the sample to low and lower-middle-income countries (with GDP per capita below $4,200). Column 3 restricts the sample to upper-middle and high-income countries (with GDP per capita above $4,200). Column 4 includes as controls the statutory tax rates for PIT, CIT, VAT, and tariff (Figure **??**). Column 5 excludes countries that do not have a VAT or general sales tax at the federal/national level. Column 6 excludes countries where the ratio of collection to expenses does not include customs. Column 7 excludes countries in which the tax authority is responsible for collecting and disbursing social security. Column 8 excludes countries in which the tax authority hosts (maintains and updates) the population registry. Robust standard errors are reported in parentheses. $*p < 0.10 **p < 0.05 ***p < 0.01$. The figure plots the (log) ratio of taxes collected relative to expenses on tax collection against GDP per capita in a sample of 100 countries.

and the average tariff rate. This specification investigates whether the variation in collection per dollar of expense is driven by variation in the amount that can be collected. It is based on the notion that countries at higher levels of development, due in part to stronger enforcement, can levy higher tax rates. Figure A.1 shows that there is some positive relation between tax rates and GDP per capita in our sample, but it is limited - in particular, it applies only PIT, with flat rates for both CIT and VAT and, as expected, decreasing rates for tariff. Moreover, the positive association between the PIT rate and development only exists for upper-middle and high-income countries, while the slope between tax collection per dollar of expense and GDP per capita is stronger in low and lower-middle-income countries. Consistent with these observations, column (3) shows that controlling for the statutory rates does not alter the slope coefficient substantially (0.221).

Columns (5) through (8) remove from the sample countries whose tax authorities differ in significant ways. In column (5), we remove countries where the national tax authority is not responsible for collecting the VAT or general sales tax. This leaves the main coefficient virtually unchanged (0.249). In column (6), we remove countries where the (domestic) tax authority is not also in charge of collecting border taxes – this decreases the sample to 90 countries and increases the slope coefficient slightly (0.260). In column (7), we remove countries where the tax authority is also in charge of collecting social security contributions - the reduction in sample size is significant (73), but the slope coefficient remains stable (0.218).[4] Overall, the sample restrictions of Table 1 suggest that the positive association between the average collection per expenditure and development is a general feature in the data that holds across a variety of institutional settings and at different levels of GDP per capita.

We estimate the relationship between our measure of average collection per expenditure and (log) GDP per capita for the full panel of countries in Table 2a. Around half of these are currently advanced economies, specifically: Australia (1972-2017), Canada (1924-2008), Denmark (1959-2006), South Korea (1987-2016), the United Kingdom (1993-2017), and the United States (1877-2016). The other half are best described as low- and middle-income economies: Bangladesh (1964-2015), Botswana (1971-2005), India (1946-2016), Kenya (1962-1995), Malawi (1964-1996), Pakistan (1949-2005), Sierra Leone (1964-1996), South Africa (1960-2007), Thailand (1959-2017), Uruguay (1980-2017). All regressions include country and year fixed effects, such that the slope coefficient captures the average association based on changes within a country over time. To account for serial correlation in both the outcome and regressor, standard errors are clustered at the country level. Column (1) indicates a positive association, which is not significant but very sim-

[4]We also find similar patterns when removing tax authorities that are in charge of hosting and maintaining the national population registry. This only removes two countries, Sweden and Norway, and the slope coefficient is essentially unchanged (0.241). In our sample, no tax authority is directly in charge of conducting the national census, neither for citizens nor for firms.

ilar in magnitude (0.150) to the association across countries. Indeed, the middle row of Table 2a indicates that we cannot reject the equality of the within-country and the between-country coefficients. These patterns hold in the following columns, which remove country-year observations where the domestic tax authority is not responsible for collecting border taxes (column 2, coefficient of 0.144) or is responsible for collecting social security contributions (column 3, coefficient of 0.106).

To complete the investigation, we turn our focus to studying the relationship between collection per expenditure and GDP per capita at the sub-national level of U.S. and Indian states. Conceptually, individual state tax authorities within a country may differ in some characteristics, which impact collection per expenditure similarly to how these characteristics vary across countries. As such, it is a meaningful extension of our main investigation to investigate whether we observe similar patterns at the sub-national level. Naturally, focusing on variation within states over time also allows us to control for all unobserved variation at the country-year level, which may confound our interpretation of the results in both the cross-section and panel of countries.

Table 2b reports the results of a basic regression, which includes state and year fixed effects and clusters standard errors at the state level. Column 1 focuses on the sample of U.S. states (1993-2017), and column 2 focuses on Indian states (1950-2017). In both cases, we uncover a positive coefficient, but which is only significant in the US case. The coefficient within U.S. states over time (1.017) is larger than the country panel regression, while the coefficient within Indian states is much smaller (0.024).

In summary, based on novel data, we find that the collection per dollar of administrative expenditure, measured as the ratio of taxes collected to collection expenses, varies positively with development across space and time: we find a robust pattern across countries today, and suggestive results within countries over the long run, at both the national and sub-national levels of government.

Table 2: Panel Regressions: Tax Collections per Expenditure and Income per Capita

(a) Panel of Countries

Outcome	(1) Log T/X	(2) Log T/X	(3) Log T/X
Log GDP p.c.	0.150	0.144	0.106
	(0.216)	(0.243)	(0.223)
Test: Equal coef.	0.280	0.271	0.603
from cross-section	(0.599)	(0.606)	(0.439)
Country FEs	X	X	X
Year FEs	X	X	X
Countries	16	14	13
N	625	546	519
R^2	0.566	0.545	0.578
Sample Restrictions		Exclude if no customs	Exclude if soc sec incl.

(b) Panel of Sub-National States

Outcome	(1) Log T/X	(2) Log T/X
Log GDP p.c.	1.017***	0.024
	(0.098)	(0.096)
Test: Equal coef.	5.870	1.460
from cross-section	(0.015)	(0.227)
Sample	US states (1993-2017)	Indian states (1950-2017)
Number of states	50	28
State FEs	X	X
Year FEs	X	X
N	850	964
R^2	0.719	0.777

Notes: This table provides panel regressions of the (log) ratio of taxes collected to expenses on collection on log GDP per capita. Panel A consists of an unbalanced panel of countries: Australia (1972-2017); Bangladesh (1964-2015); Botswana (1971-2005); Canada (1924-2008); Denmark (1959-2006); India (1946-2016); Kenya (1962-1995); Malawi (1964-1996); Pakistan (1949-2005); South Korea (1987-2016); Sierra Leone (1964-1996); South Africa (1960-2007); Thailand (1959-2017); United Kingdom (1993-2017); United States (1877-2016); Uruguay (1980-2017). All columns include country and year fixed effects. Column 1 runs the basic OLS. Column 2 excludes country-years where customs are not included in the T/X ratio. Column 3 excludes countries where the tax authority is responsible for social security. Panel B consists of an unbalanced panel of sub-national states: US states between 1993 and 2016 (column 1) and Indian states between 1950 and 2017 (column 2). In both panels, the middle row reports the Wald test and p-value of the equality of coefficients between the specific column and column 1 of Table 1, using the 'suest' command in Stata and an asymptotic χ^2 distribution. All regressions include state fixed effects. Standard errors are clustered at the country or state level. * p<0.10 ** p<0.05 *** p<0.01.

3 Model of Tax Productivity and Development

We now turn to interpreting these patterns and exploring their implications for economic development. Our interpretation is based on a version of the neoclassical growth model in which tax revenue finances public capital, which enters the production function as in Baxter and King (1993). Unlike the main macro analyses of productive public capital, which focus on advanced economies and assume costless taxation, we model revenue as costly to raise. Instead, a tax administration sector operates a technology that converts household income into tax revenue for the government. Countries differ both in their productivity in creating consumption and investment goods and in their tax productivity, which governs how easily they raise revenue. Countries also differ in the fraction of revenue that is diverted and in the share they devote to public goods rather than transfers.

3.1 Environment

Each economy has a unit measure of households with preferences that are separable in consumption and labor supply:

$$\sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\gamma}}{1-\gamma} - \alpha \frac{L_t^{1+1/\varepsilon}}{1+1/\varepsilon} \right], \tag{1}$$

where C_t and L_t are consumption and labor at time t. All countries have the same discount factors, β, curvature of utliity of consumption, γ, disutility of labor supply, α, and Frsch elasticity of labor supply, ε. Households supply labor and capital in competitive markets and solve a standard consumption savings problem. Their budget constraint is:

$$C_t + K_{t+1} = (1 + r_t - \delta)K_t + (1 - \tau_t)w_t L_t + TR_t \tag{2}$$

where K_t is the household's capital stock, r_t and w_t are the competitive rental rates of capital and wage rates, τ_t is the effective tax rate on labor income, and TR_t represents transfers from the government. We will define the effective tax rate below. From the household's perspective, τ_t is taken as given and represents the fraction of labor income taken by the government. The household also takes transfers as given.

The production function for the final good – which can be used for either consumption or investment – is:

$$Y_t = F(K_t, K_{G,t}, L_{M,t}) = A_M(\theta K_{G,t})^{\zeta} K_t^{\mu} L_{M,t}^{1-\mu}, \tag{3}$$

where $K_{G,t}$ and $L_{M,t}$ are inputs of public capital and labor, A_M is the country's exogenous productivity level in final goods, and θ is the country's exogenous efficiency in turning public capital investments into real public capital. Parameter μ represents private capital's share in production and is common to all countries. The subscript M denotes labor supplied to this modern sector as distinct from labor employed in the government's tax administration sector (or, in an extension to follow, a traditional sector).

As in Baxter and King (1993) there are constant returns to scale in privately provided inputs: K_t and $L_{M,t}$. The power ζ governs the output elasticity of public capital, holding other inputs fixed. So long as ζ is positive, there are increasing returns to scale in all inputs (private and public combined).

The tax collection technology is given by:

$$\tau_t = G(L^i_{C,t}) = A^i_C (L^i_{C,t})^\eta, \tag{4}$$

where A^i_C is exogenous tax productivity in country i, $L^i_{C,t}$ is labor input into tax collection (e.g. the number of "tax collectors"), and $0 < \eta < 1$ governs the returns to scale in tax collection. If $\eta = 1$ there is constant returns to scale in tax collection, and each additional tax collector raises the effective tax rate by as much as the last. If $\eta < 1$ there will be decreasing returns to scale, and each additional tax collector will prove less effective than prior ones. Later we argue that the latter is the empirically relevant case, and that it materially affects the paper's main conclusions.

For simplicity we assume that government's exogenously decide the number of tax collectors each period and these choices are fully known by the households. Governments thus move first in determining the full schedule of tax rates each period, and households respond by solving their dynamic problems taking government behavior as given, as in a Stackelberg game. In the extended model to follow we let the government's tax-collection inputs be chosen endogenously, though at some cost in transparency relative to this simpler setup.

The government budget constraint each period is:

$$TR_t + X_t + K_{G,t+1} = T_t + (1-\delta_G)K_{G,t}. \tag{5}$$

where T_t represents total tax collection:

$$T_t = \tau_t w_t (L_{M,t} + L^i_{C,t}), \tag{6}$$

and X_t represents expenditures on tax collections:

$$X_t = w_t L_{C,t}. \tag{7}$$

The government thus uses its current resources each period to finance transfers, payments to tax collectors, and public capital for the next period. Its current resources are its tax revenues from market labor and government tax collectors plus the current un-depreciated stock of public capital. Governments exogenously split their gross tax revenues between transfers and public capital investments:

$$TR_t = \xi T_t, \tag{8}$$

implying that:

$$K_{G,t+1} = w_t \Big[(1-\xi)\tau_t (L_{M,t} + L_{C,t}) - L_{C,t} \Big] + (1-\delta_G) K_{G,t}. \tag{9}$$

Economies thus exogenously differ in four parameters: final-goods productivity, A_M, tax productivity, A_C, public spending effectiveness, θ, and the transfer share of gross tax revenues, ξ. They also differ in their number of tax collectors, L_C, each period.

3.2 Equilibrium and Tax Collections per Expenditure

Our equilibrium concept is a sequence-of-markets equilibrium in which households solve their dynamic problem taking prices, tax rates, and transfers as given. Firms solve their static problems taking prices as given and all markets clear. Tax collections per expenditure in equilibrium are:

$$T_t / X_t = A_C (L_{C,t})^{\eta} (L_{M,t} + L_{C,t}) / L_{C,t}. \tag{10}$$

There is a direct relationship, in other words, between tax collections relative to expenditures, the statistic we reported in the previous section, and tax productivity, A_C, in the model. Both are increasing in the wage rate of the economy, but these cancel out in the ratio. Labor in the modern sector depends on the level of income and the effective tax rate in equilibrium, and these depend endogenously on other parameters of the model.

3.3 Steady State

We characterize the steady state for a single economy. The effective tax rate is:

$$\tau = A_C (L_C)^{\eta}, \tag{11}$$

which will be positive so long as A_C and L_C are positive but could be greater than one in principle. In our quantitative work to follow, tax rates stay strictly between zero and one. A steady state (C, K, K_G, L_M, w, r) satisfies the household optimality conditions, government budget constraint, and resource constraint in steady-state:

$$1 = \beta(1 + r - \delta), \tag{12}$$
$$\alpha(L_M + L_C)^{1/\varepsilon} = C^{-\gamma}(1 - \tau)w, \tag{13}$$
$$\delta_G K_G = w\Big[(1 - \xi)\tau(L_M + L_C) - L_C\Big], \tag{14}$$
$$C + \delta K + \delta_G K_G = Y, \tag{15}$$

where wage and rental rates satisfy the firm optimality conditions in steady state:

$$Y = A(\theta K_G)^{\zeta} K^{\mu} L_M^{1-\mu}, \tag{16}$$
$$r = \mu A(\theta K_G)^{\zeta} K^{\mu-1} L_M^{1-\mu}, \tag{17}$$
$$w = (1 - \mu)A(\theta K_G)^{\zeta} K^{\mu} L_M^{-\mu}. \tag{18}$$

It is convenient to recover the steady state from the pair (L_M, K). For a given (L_M, K), public capital is:

$$K_G = \left[\frac{(1 - \mu)A(\theta)^{\zeta} K^{\mu} L_M^{-\mu}\Big[(1 - \xi)\tau(L_M + L_C) - L_C\Big]}{\delta_G}\right]^{\frac{1}{1-\zeta}}. \tag{19}$$

Then

$$w = (1 - \mu)A(\theta K_G)^{\zeta} K^{\mu} L_M^{-\mu}, \tag{20}$$
$$C = \left(\frac{(1 - \tau)w}{\alpha(L_M + L_C)^{1/\varepsilon}}\right)^{1/\gamma}. \tag{21}$$

Hence the steady state is characterized by two equations in two unknowns:

$$\mathscr{R}_1(L_M, K) \equiv 1 - \beta\left[1 + \mu A(\theta K_G)^{\zeta} K^{\mu-1} L_M^{1-\mu} - \delta\right] = 0, \tag{22}$$
$$\mathscr{R}_2(L_M, K) \equiv C + \delta K + \delta_G K_G - A(\theta K_G)^{\zeta} K^{\mu} L_M^{1-\mu} = 0. \tag{23}$$

4 Quantitative Analysis

We now draw on our cross-country evidence on tax collections relative to tax administration costs and other salient statistics to recover tax productivity in each country and simulate the long-run effects of raising it. To do so we parameterize the model in the 97 countries for which we have complete data.

4.1 General Parameters

We begin by assigning values to general parameters that apply to all countries. We deliberately take standard choices wherever they are available. Each period represents one year and thus the discount factor, β, is set at 0.96. The depreciation rates of physical and public capital, δ and δ_G, are taken to be 0.05. Following the evidence of Gollin (2002) we assume that capital's share in the production function is 0.33, consistent with a labor share that is roughly two-thirds across countries of all income levels.

We pick a Frisch elasticity of labor supply, ε, of 0.5, which is somewhere between the smaller values preferred by microeconomic estimates and the larger values favored by most macroeconomic estimates. It is also the long-run value favored in the recent study by Kleven, Kreiner, Larsen, and Søgaard (2025). We normalize the disutility weight on labor supply, α, to be 1 as this value affects only the level of labor supply, which is not relevant in our analysis. We pick a coefficient of relative risk aversion, γ, of 1.1, which implies that income effects are modestly stronger than substitution effects in labor supply decisions. This is the value used by Bick, Fuchs-Schündeln, Lagakos, and Tsujiyama (2022) for example in trying to match cross-country differences in hours worked, and the decreasing patterns of hours with development.

The remaining two general parameters are somewhat less standard in macroeconomics and, as we show below, quite important for our quantitative conclusions. These are: the returns to scale in the tax administration sector, η, and the output elasticity of public capital, ζ. We discuss each of these in turn.

We argue that η is likely to be far below one, and we set $\eta = 0.05$ in our quantitative analysis based on our reading of the evidence. In the model, η corresponds to the ratio of the marginal product of tax collections to the average product (e.g., total collections per tax administration expenditure, the subject of our empirical analysis). The evidence points to marginal products well below average products, and it is well established that tax authorities organize their resources and activities to prioritize the more easily taxable segments of the economy (Sugin, 2014; Christians,

Table 3: General Parameters

Description	Parameter	Value
Discount factor	β	0.96
Risk aversion / inverse IES	γ	1.1
Frisch elasticity	ε	0.5
Labor disutility weight	α	1
Capital share in modern production	μ	0.33
Output elasticity of public capital	ζ	0.20
Tax-administration curvature	η	0.05
Private capital depreciation	δ	0.05
Public capital depreciation	δ_G	0.05

Notes: This table reports the parameters and values used in the quantitative analysis for all countries. See the text for rationales for each parameter choice.

2013; OECD, 2024; IMF, 2015; EU, 2015). For example, wage employees are easier to tax than the self-employed (Kleven, Knudsen, Kreiner, Pedersen, and Saez, 2011), larger firms are easier to tax than smaller firms (Pomeranz, 2015; Kleven, Kreiner, and Saez, 2016), and consumers in large, modern retail stores are more easily taxed on their expenditure than consumers in small, traditional stores (Bachas, Gadenne, and Jensen, 2024).[5]

Average products of tax collection come directly from our data on T/X. Direct evidence for the marginal cost of collection is harder to come by, but can be proxied using several recent micro studies. In many low income countries, property taxes are an important new margin that governments are trying to pursue to increase tax revenues. In Ghana, Dzansi, Jensen, Lagakos, and Telli (2025) estimate that the ratio of collections to revenue collector salaries across local governments is 1.56. They find similar numbers in an RCT within one large local government that experimentally increased both collections and expenses to hire tax collectors, which implied a marginal collection per dollar of expenses of 1.62. Our cross-country data for Ghana show that total taxes relative to total tax administration expenses equals 43 (see Figure 1). This implies a ratio of marginal to average of 0.04 (=1.6/43).

Moving to a middle-income country, in Costa Rica, Brockmeyer, Smith, Hernandez, and Kettle (2019) uses an experiment to estimate the impact of increased enforcement on nil filers (registered

[5] Many tax codes exempt firms below certain thresholds (of size, assets, or number of employees) and in certain sectors (often agriculture) from the main tax bases. Manuals for practitioners explicitly justify this policy design as driven by administrative cost-benefit considerations: in *The Modern VAT*, IMF (2001) states that “for the great majority of small traders, the cost of taxing them would exceed the VAT that could be collected, so an exemption threshold is efficient and sensible.”

taxpayers who declare a value of 0 for their taxable activities). The paper finds each additional dollar of enforcement expenditure leads to around \$4 in additional tax revenue. Our cross-country data report average collections per unit of expense in Costa Rica of 82, which implies that the ratio of marginal to average returns is 0.05 (=4/82).

Finally, evidence from a high-income country can be found in Boning, Hendren, Sprung-Keyser, and Stuart (2025), which estimates the marginal collections from random audits of individual income taxpayers in the United States. Randomized audits are not the only margin that the IRS can use to increase tax collections, but it is likely to be one of their most important margins. The authors find that each additional dollar spent on audits generates \$13 of tax payments. With an average collection of 260 in the United States in our cross-country data, the ratio of marginal to average returns is 0.05 (=13/260). Our choice of a constant $\eta = 0.05$ is consistent with these three pieces of evidence, drawn from low-, middle-, and high-income countries, which point to remarkably similar ratios.

The elasticity of output to public capital, ζ, has received more attention in the literature, though no clear consensus has emerged on its value – especially in a cross-country setting and for a broad notion of public capital. In our main analysis, we choose a value of $\zeta = 0.2$,, but consider sensitivity to lower and higher values later. Our read of the literature is that the wide range of estimates for ζ reflects both difficulties in estimation as well as the natural variation in the actual value of public spending, depending on the project and location. Aschauer (1989) estimated a ζ of 0.39 in his famous study of U.S. infrastructure spending. Others found far lower values for transportation infrastructure and Ramey (2021) favors a value of 0.05 in her analysis of transportation infrastructure. Cubas (2020) uses national income accounts data to infer a value of 0.09.

The World Bank draws on estimates of ζ to discipline its Long-Term Growth Model (see Devadas and Pennings, 2018), drawing on a well-known meta-analysis by Bom and Ligthart (2014) based on studies that mostly pre-date the ‘credibility revolution’ in economics. Figure A.2 plots the estimated elasticities from the meta-analysis; our benchmark value of 0.20 falls well within their wide range. One potential limitation of these studies for our purposes is that 93 percent of them come from advanced economies. Another is that they are primarily focused on transportation infrastructure. To the extent that public capital represents broader types of public services – like functioning court systems or police departments – these estimates could significantly underestimate (or overestimate) the true return to public capital.

Several more recent studies in less developed settings are informative about the output elasticity of public capital. However, these studies also do not point to a consensus value for ζ. Donaldson and Hornbeck (2016) study the expansion of the U.S. railroad network to estimate the contribution of transport infrastructure to agricultural land values. Their estimates imply a long-run output elasticity of public capital of around 0.6. By contrast, Faber (2014) examines the construction of China's national highway system and finds that while highway access generates gains for directly connected areas, the implied output elasticity for peripheral regions is close to zero. Both of these studies are again focused on transportation infrastructure. Taken together, they reinforce the need to consider a range of plausible ζ values.

4.2 Cross-Country Parameterization and Predictions

We take the other parameters to be country-specific. One of these is θ, which governs the fraction of government tax revenue that translates into real public capital. Our strategy is to take three separate targets for low-, middle-, and high-income economies, since it is widely agreed that government corruption is more prevalent in less developed economies. We target values of 0.58 for the low-income economies, 0.66 for the middle-income countries, and 0.84 for the high-income ones. These values are drawn from the World Bank's Infrastructure Efficiency Index (IEI); see Devadas and Pennings (2018). To our knowledge there are no country-specific values of this parameter that have been constructed.[6]

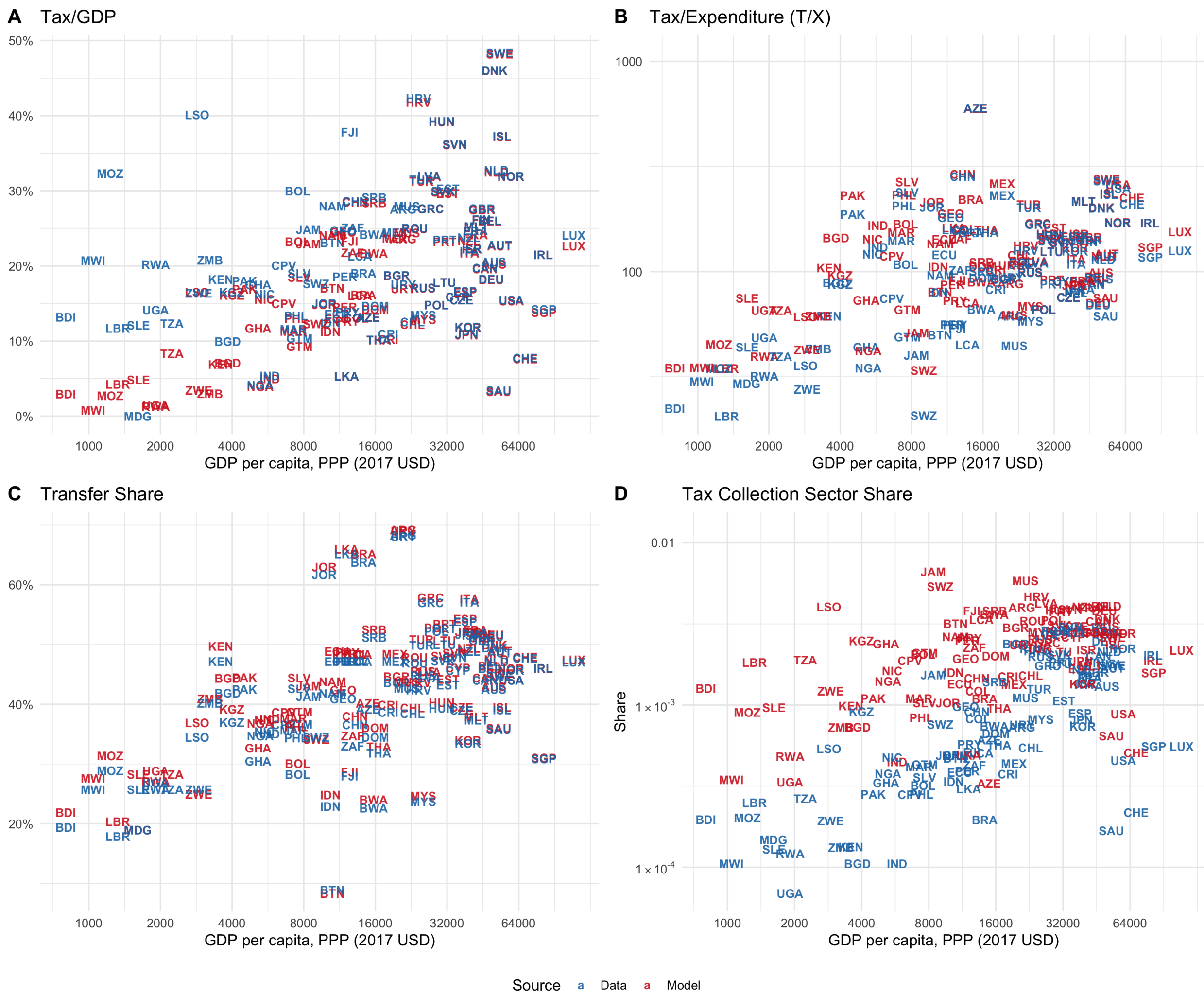
Four parameters remain to be set for each country: general productivity, A_M, tax productivity, A_C, the fraction of public revenues that are redistributed as transfers, ξ, and the size of the tax collection sector, L_C. We target these jointly to match four targeted moments in each country: the tax to GDP ratio, the T/X ratio from the empirical section above, the transfer share of public expenditures, and tax collection employment as a fraction of total employment.

Figure 2 plots each of the targeted moments in the model and data against GDP per capita. The model generally succeeds in matching the increasing patterns in taxes to GDP (Panel A), taxes per unit of administrative expenditure (Panel B), and the transfer share of public expenditures (Panel C). Some lower income countries are outliers in taxes to GDP, such as Mozambique and Lesotho, and these are not well captured by the model. These deviations may reflect revenue sources our model omits, such as offshore gas leases in Mozambique and Lesotho's unique customs union with its richer neighbor, South Africa. Tax collectors as fraction of total employment (Panel D) overshoots the data for most countries, though the magnitudes are small throughout and well below one percent of all employment in both the model and data. Both are modestly increasing in income per capita.

The model also does reasonably well in matching cross-country moments that are not targeted

[6] Nor do we view these as definitive estimates, as it is very difficult to accurately measure how effectively governments turn nominal public expenditures into real investment (Pritchett, 2000). We are working on additional sensitivity analyses to these parameters.

Figure 2: Model's Predictions for Targeted Moments



Notes: This figure plots the model's predicted tax-to-GDP ratios (Panel A), taxes per unit of expenditure on collection (Panel B), share of tax revenues used for transfers rather than public capital (Panel C), and employment share of tax collectors (Panel D). In each graph the red dots represent the model's predictions and the blue dots represent their empirical counterparts. Total taxes collected, tax administration expenditures, and the size of the tax collection sector by employment come from our own calculations described in the text. The transfer share comes from the IMF.

in the calibration. Figure 3 plots the model's predictions for private capital-labor ratios, public capital-labor ratios, investment shares of GDP, and labor supply (relative to the world average). Private capital-labor ratios are readily observable in the data and we take these directly from the Penn World Tables. Panel A shows that our model gets an almost identical slope in income per capita. This is reassuring, though perhaps not terribly surprising, given the successes of the neo-classical growth model in this dimension. Public capital-labor ratios are more difficult to measure in practice, though they can proxied by investment data from the IMF's Government Financial Statistics. Panel B plots the IMF's public capital per worker, scaled by our θ values. In practice the real mapping between nominal public expenditure and real government capital stocks could be far lower or higher, and the choice of only three θ values to cover all low-, middle-, and high-income countries is obviously a crude one. It is still useful to verify that in the model and data, public capital stocks per worker rise nearly linearly in income per capita.

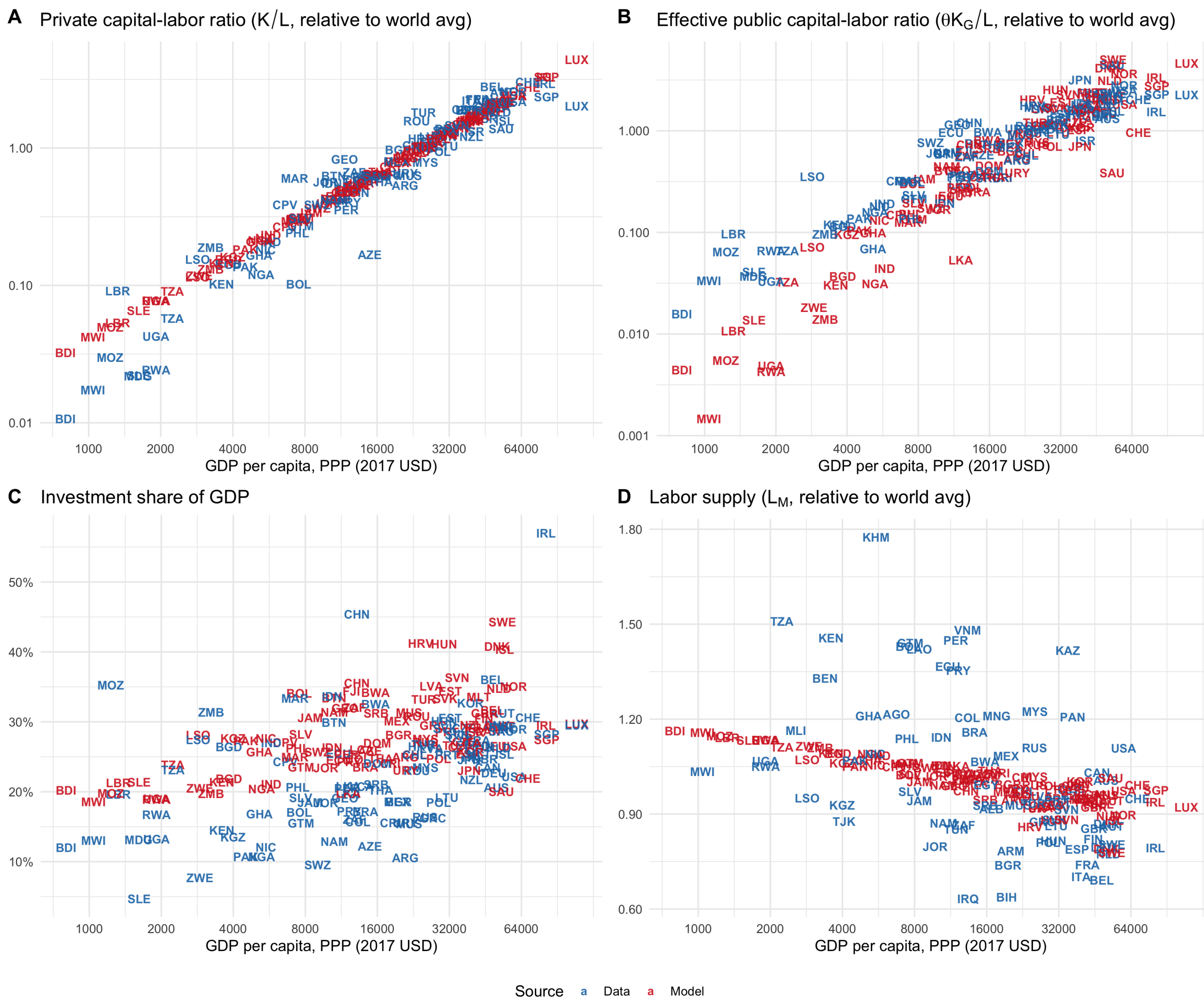
Panel C reports the investment share of GDP at domestic prices, again taken from the Penn World Tables. Both in the data and model these are modestly increasing in GDP per capita, though our model over-predicts the level of investment to GDP on average. Panel D plots average hours worked per adult from Bick, Fuchs-Schündeln, and Lagakos (2018), expressed relative to the world average. Both in the model and data, hours exhibit a modest decline in income per capita. Matching the long-run decline in hours is reassuring, particularly in light of the labor-supply responses we report below. It is also expected, given that our calibration of preferences makes income effects modestly dominate substitution effects.

One of the goals of the parameterization was to recover the values of A_C and A_M needed to match the cross-country evidence. Figure 4 plots these for each of the 97 countries. Two patterns can be seen from these two plots. First, A_C exhibits a steeper slope with GDP per capita (Panel A) than does A_M (Panel B). Second, there is more variation conditional on income for A_C than A_M. We are mostly interested in the first pattern from a development perspective. A regression of A_C on log GDP per capita yields a slope of 0.666 (s.e. 0.148). For A_M, the slope coefficient is 0.364 (s.e. 0.036). As a frame of reference, the slope coefficient of the Penn World Tables' measure of TFP is 0.300 (s.e. 0.020), which is quite similar to (indeed, statistically indistinguishable from) the slope in A_M. The upshot is that A_C varies more with development than A_M, and our best estimate is that A_C varies around twice as much with income per head as a more general measure of TFP.

4.3 Quantitative Counterfactuals

Once the model is calibrated country by country, our main quantitative exercise simulates the long-run general equilibrium effects of raising tax productivity, or A_C. We compute the effects in

Figure 3: Model's Predictions for Non-Targeted Moments



Notes: This figure plots the model's predicted capital-labor ratios (Panel A), public capital-labor ratios (Panel B), investment shares of GDP (Panel C), and labor supply (Panel D). In each graph the red dots represent the model's predictions and the blue dots represent their empirical counterparts. GDP per capita, employment, and private capital stocks come from the Penn World Tables version 11. Public capital stocks come from International Monetary Fund. Labor supply is proxied by hours worked per adult come from Bick, Fuchs-Schündeln, and Lagakos (2018).

Figure 4: Tax Productivity and Aggregate Productivity

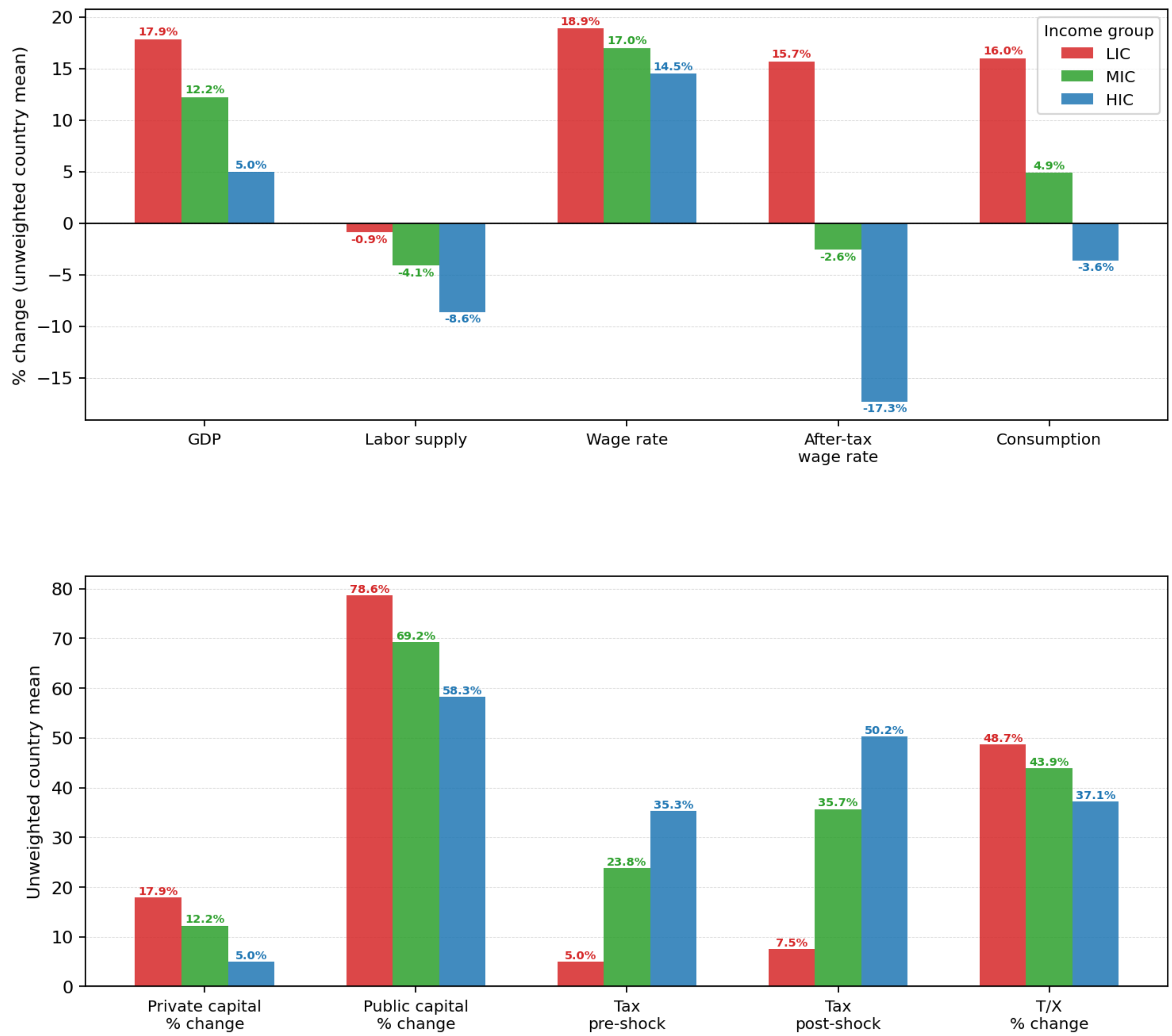
Notes: This figure plots the model's predicted tax productivity levels (Panel A) and aggregate productivity terms (Panel B) in each country. Both series are estimated using the model and data described in the text.

particular of an increase in A_C of 50 percent, which, we argue, should be considered a transformative change in how efficient tax collection is. We note that productivity increases in this range are eminently feasible given existing policy levers available to developing economies. For example, in an experiment by a subset of the authors (Dzansi, Jensen, Lagakos, and Telli, 2025), better databases of property owners linked to hand-held tablets with GIS capabilities roughly doubled tax collections in Ghana. Although these results do not literally map into changes in A_C, they suggest that large changes in the productivity of tax collectors are not purely hypothetical.

Figure 5 reports the changes in GDP and other key macroeconomic variables in the new steady state resulting from a permanent increase in A_C of fifty percent. The red bars report the average percent changes across all low-income countries. GDP increases by 17.9 percent on average. Middle-income and high-income countries (green and blue, respectively) see somewhat smaller gains of 12.2 percent and 5 percent. This boost in output is accompanied by consumption gains of 16 percent in the low-income countries, 4.9 percent in the middle-income income countries, and a consumption loss of 3.6 percent in the high-income ones. Thus, the consumption effects are large and positive in the poorest countries but turn negative in the richest.

Why? To understand the mechanics of the model's predictions, note that an increase in A_C raises effective tax rates by 50 percent in all economies, since labor in the tax collection sector is fixed: $\bar{\tau} = A_C \bar{L}_C^{\eta}$. From the government budget, one can see that the higher tax rate raises k_G directly, and

Figure 5: Long-Run Effects of Raising Tax Collection Productivity



Notes: This figure plots the model's predicted percent changes in macroeconomic variables from a permanent 50-percent increase in tax productivity, A_C. The three sets of bars represent low-income countries (red); middle-countries (green); and high-income countries (blue), using the World Bank country categories.

further raises it through a higher equilibrium wage; it is at least partly offset as higher taxes reduce labor supply:

$$k_G = \delta_G^{-1} w_M \left[(1-\xi)\bar{\tau}(L_M + \bar{L}_C) - \bar{L}_C \right]. \tag{24}$$

From the production function, $y_M = A_M(\theta k_G)^{\zeta} k^{\mu} L_M^{1-\mu}$, one can see that output is affected directly through increases in public capital, with a short-run output elasticity of public capital equal to ζ.

Output is also affected indirectly through changes in private capital and changes in labor supply. From the Euler equation, the rental rate of capital, r, depends only on the rate of time preference and depreciation rate in steady state, and is hence constant. The marginal product of capital, and capital-output ratios, are thus also constant in steady state: $r = \mu y_M / k$. Any changes in output are thus the same in percentage terms as the changes in private capital. The wage rate, and marginal product of labor, increase in output and decrease in labor: $w_M = (1-\mu) y_M / L_M$.

From the labor-leisure household optimality condition, $L_M^{1/\varepsilon} = c^{-\gamma}(1-\bar{\tau})w_M$, higher wage rates raise labor supply while higher tax rates lower it. Consumption depends on after-tax wages and labor supply: $c = [(1-\bar{\tau})w_M / L_M^{1/\varepsilon}]^{1/\gamma}$.

With these equations in mind, one can see more clearly how the low-income countries benefit more from increases in tax productivity in these counterfactual simulations. In low-income countries, the gains in public capital are 78.6 percent. These drive a 17.9 percent increase in private capital, which is the same as the GDP increase. Wages increase by 18.9 percent, benefiting workers. The starting tax rate is so low that a 50 percent increase in taxes takes the low-income economies from a tax rate of around 5 percent to around 7.5 percent. After-tax wages rise by 15.7 percent, which certainly still leaves workers better off. Labor supply falls by just 0.9 percent. The resulting change in consumption is 16 percent.

The calculus is different in upper-middle income countries and high-income countries. Public capital increases substantially in those countries too, at 69.2 and 58.3 percent respectively. Private capital stocks rise but by less, at 12.2 and 5 percent, with wages rising significantly in both cases. The downside is that tax rates rise from 23.8 percent to 35.7 in the middle income countries and from an already robust 35.3 percent to over 50 percent in the high-income ones. These leads to declines in after-tax wages, and quite significant ones in the richest countries. This explains the more modest 4.9 percent increase in consumption in the middle-income nations and the consumption drop in the richest ones.

In reality, middle-income and rich countries would likely reduce the sizes of their tax collection sectors in response to an increase in tax productivity. This would in turn keep effective tax rates closer to their original levels. In that case, the consumption and labor-supply changes would be closer to zero than these simulations show. The point is still that lower-income countries respond quite differently to tax productivity changes than do richer economies. These endogenous changes are discussed further below, particularly in the extension to Ramsey taxation.

Table 4 puts the long-run GDP changes for low-income countries in perspective by considering Hulten's approximation as a simple benchmark. Hulten's approximation says that the aggregate impact of a 50 percent TFP increase in a sector is roughly equal to 50 percent times the sector's

Table 4: Model Predictions vs Hulten in Low-Income Economy

Variable	Percent Change
Model output	17.9%
Hulten's prediction (wage share: $0.5 \times w_M L_C / Y$)	0.03%
Hulten's prediction (tax share: $0.5 \times T/Y$)	1.7%
Public capital effect ($\zeta \times \%\Delta k_G$)	15.7%

Notes: This table reports the average GDP gain from a 50 percent increase in A_C in the model (first row) relative to the predicted gain according to Hulten's Approximation (second and third rows), and the pure effect of public capital (fourth row). The gains are an average for all low-income countries.

Domar weight (defined as its sales relative to GDP). The tax collection sector does not have sales per se, so we consider its wage bill share and also total tax collections to GDP as weights.

In both cases the model's GDP gain of 17.9 percent far exceeds the simple Hulten approximation. Using the wage bill share, the A_C gains should boost the aggregate economy by a minuscule 0.03 percent. With the tax-to-GDP ratio as a weight, the Hulten approach predicts a 1.7 percent GDP increase. The model's prediction is still an order of magnitude larger. The fourth row isolates the direct public-capital channel: a 78.6 percent increase in k_G scaled by the output elasticity $\zeta = 0.2$ yields a 15.7 percent GDP gain, most of the model's total. The rest comes from the induced rise in private capital and change in labor supply.

How does the model predict such outsized long-term gains from upgrading productivity in the (tiny) tax administration sector? The answer comes not from within the sector itself, but from the resulting increase in public and private investment illustrated above in Figure 5. Hulten's theorem presumes efficient allocation without externalities; here public capital acts as a production "externality" financed by distortionary taxation, so the first-order approximation misses the reallocation toward public and private capital.

For additional perspective, Table 5 reports the effects of other counterfactual changes one could imagine in low-income countries. The first column reproduces the effects of a fifty percent increase in A_C (reported as a fraction of baseline, for expositional purposes). The second column simulates raising the efficiency of public expenditure from the 0.58 up to 0.84, the level computed for high-income countries by the World Bank. This would result in a 16 percent increase in GDP, both public and private capital stocks, and consumption. The tax rate and T/X ratio would not change in any important way. In all other aspects this experiment looks remarkably like the higher A_C scenario. The takeaway is that low-income can, in principle, achieve the same gains from improving their tax collection capacity by investing their current tax proceeds more effectively.

Low-income countries also do not need better tax productivity to increase tax revenues and public

Table 5: Effects of Other Counterfactuals on Low Income Countries

	$A_C \times 1.5$	$\theta \to 0.84$	$L_C \times 1.5$	$A_M \times 1.5$
GDP	1.18	1.16	1.00	2.28
Private capital	1.18	1.16	1.00	2.28
Public capital	1.79	1.16	1.01	2.28
Consumption	1.16	1.16	1.00	2.28
Transfers	1.77	1.16	1.02	2.28
After-tax wage	1.16	1.17	1.00	2.34
Labor supply	0.99	0.99	1.00	0.97
T/X ratio	1.49	0.99	0.68	0.97
Tax rate	1.50	1.00	1.02	1.00

Notes: This table reports the effects of other counterfactual parameter changes on low-income countries in the long run. Effects are reported as the new long-run level of the variable in each row relative to the baseline calibration. The numbers reported are averages across all low-income economies.

capital: they can simple hire more tax collectors using the current systems. The third column of Table 5 simulates raising the number of tax collectors by fifty percent. Unfortunately, the model predicts virtually no gains from such a change. Tax rates rise only modestly due to the strong decreasing returns to scale in that activity. Intuitively this would be like sending more tax collectors into areas with self-employed family farms and business or small firms more generally. In the model, public capital increases by around 1 percent and GDP increaeses by even less. The tax to expenditure ratio falls significantly to 68 percent of the previous level, as expenditures rise significantly while tax collections increase only very modestly.

The final column of Table 5 simulates raising A_M by fifty percent. This is no "policy counterfactual" to be sure: few concrete policy options identified by economists correspond to such massive TFP increases. It is still useful to simulate this scenario as a frame of reference to the similar sized increases in tax productivity. Countriues make enormous gains from such a change, with GDP per capita increasing by a factor of 2.28, along with public and private capital stocks, reflecting neoclassical forces. Labor supply falls (only slightly more than when raising A_C) due to income effects. Tax rates do not change, and the T/X ratio changes only modestly, reflecting the earlier intuition that richer households make for juicier targets for tax collection, but also more expensive tax collectors.

4.4 Efficient Benchmark

Are increases in public capital desirable from an efficiency perspective? To get insight into this question we follow Ramey (2021) and ask how the competitive outcomes above compare to the efficient allocation that would have been chosen by a benevolent social planner. Following tradition, we shut down markets and allow the planner to direct inputs to production as she chooses, and to allocate output to consumption, investment in private capital, or investment in public capital. In other words, we eliminate the restriction that households supply labor in markets and that the government has to use up resources in order to take goods from the household and turn them into public investment.

Formally, the planner chooses $\{C_t, L_t, Y_t, K_{t+1}, K_{G,t+1}, \Lambda_t, L_{M,t}, L_{C,t}\}_{t=0}^{\infty}$ to maximize

$$\mathscr{L} = \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\gamma}}{1-\gamma} - \frac{L_t^{1+1/\varepsilon}}{1+1/\varepsilon} + \Lambda_t \Big(Y_t + (1-\delta)K_t + (1-\delta_G)K_{G,t} - C_t - K_{t+1} - K_{G,t+1} \Big) \right]$$

where $L_t = L_{M,t} + L_{C,t}$ and $Y_t = A_M \left(\theta K_{G,t}\right)^{\zeta} K_t^{\mu} L_{M,t}^{1-\mu}$ each period and given K_0 and $K_{G,0}$. The planner's optimality conditions are:

$$C_t^{-\gamma} = \beta C_{t+1}^{-\gamma} \left(\mu \frac{Y_{t+1}}{K_{t+1}} + 1 - \delta \right),$$

$$C_t^{-\gamma} = \beta C_{t+1}^{-\gamma} \left(\zeta \frac{Y_{t+1}}{K_{G,t+1}} + 1 - \delta_G \right),$$

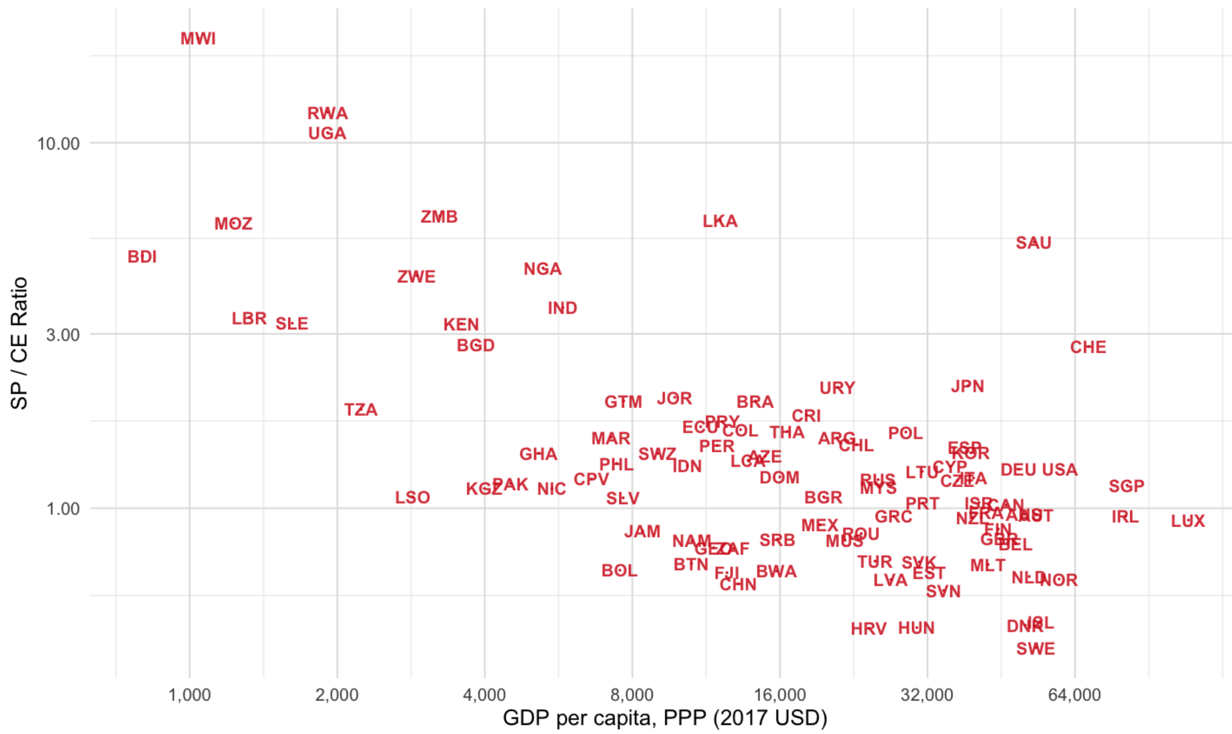
$$L_t^{1/\varepsilon} = C_t^{-\gamma}(1-\mu)\frac{Y_t}{L_{M,t}},$$

and govern optimal investment in private capital, public capital, and deployment of labor inputs. In steady state, the efficient public capital-output ratio is then:

$$\frac{K_G^{SP}}{Y^{SP}} = \frac{\zeta}{\beta^{-1} - 1 + \delta_G}, \tag{25}$$

which depends critically, and quite intuitively, on the parameter ζ. The planner will want a higher stock of public capital relative to output the higher is ζ. The rate of time preference and depreciation rate also factor, and for our calibrated parameters the efficient public capital-output ratio is around two. Interestingly, this applies to all economies: neither A_M nor A_C features in optimal public capital stocks relative to output.

Figure 6: Public Capital: Efficient Benchmark Relative to Market Allocation



Notes: This figure plots the efficient public capital to output ratio relative to the market's ratio.

The competitive equilibrium delivers entirely different values of the public capital-output ratio:

$$\frac{K_G^{CE}}{Y^{CE}} = \frac{(1-\mu)}{\delta_G} \cdot \frac{(1-\xi)\tau(L_M+L_C)-L_C}{L_M}. \tag{26}$$

These depend on the labor's share in the production function, effective tax rates (which depend on A_C), the fraction of tax revenues invested rather than transferred, the total number of tax collectors and the total labor supply. The output elasticity of public capital is not present at all. These equations highlight how this neoclassical model is not wired to reach an efficient steady state.

We solved the social planner's steady state for each country given the same parameter values in the competitive equilibrium. Figure 6 plots the ratio of the planner's capital-output ratio to the one present in the competitive equilibrium. A value of one, in other words, means that the public capital-output ratios are the same under the planner and market outcomes.

We find that in low income countries, the public capital-output ratios are significantly higher under planner's outcome than in the market allocation. On average, the planner has around three times the ratio of the market, suggesting underprovision of public capital. Middle- and high-income countries are closer to the efficient level, though there is significant variation country to country. On average, countries above around $8,000 per capita income per year have a ratio around one. This analysis points in the direction of low public capital stocks in less developed countries being an inefficient outcome relative to this idealized scenario when output could be costlessly redirected toward investments in "public" and "private" types of capital. To the extent that low-income countries could reduce the costs of collecting taxes, in other words, the resulting increases in capital stocks would move them closer to the efficient benchmark.

4.5 Sensitivity

We are still working on sensitivity analysis. For now we report the effect of two obvious changes: lower and higher values of ζ, and lower and higher values of η. The first three columns of Table 6 takes a value of 0.10, 0.20 (the benchmark), and 0.30 for ζ. Slopes of A_C are similar in all cases, but higher values of ζ lead to a more modest slope of A_M in income per capita.

Gains in GDP, consumption and after tax wages are still present when zeta takes on the smaller value but significantly reduced, at only a 6.9 percent incresae in GDP for example. At the higher value, GDP gains come out at 36.7 percent, with robust growth in consumption and after tax wages. If this case does not represent what is most likely given the literature's knowledge about the importance of public capital, at least it represents what is possible.

Alternative values of tax curvature of 0.10 and 0.20 (compared to 0.05 in the benchmark) have little bearing on the quantitative conclusions. Neither the inference about the variation of A_C in development nor the impacts of raising A_C in low-income countries change to any relevant degree with these modestly higher values. We are working on cases with even higher η values, and this should certainly become material for large enough values. When η approaches 1, the fact that A_C is low ceases to become an important constraint on investing more in public capital. Poor economies can simply hire more and more tax collectors – with no significant productivity loss – and reach higher levels of public capital stocks. We will have this illustrated in subsequent iterations of this section.

Table 6: Sensitivity Analysis

	Output Elasticity of K_G (ζ)			Tax Curvature (η)	
	0.10	0.20	0.30	0.10	0.20
Cross-Country Productivity Slopes					
Slope of A_M in GDP/capita	0.515	0.364	0.196	0.367	0.318
Slope of A_C in GDP/capita	0.666	0.666	0.662	0.652	0.624
Counterfactual: $A_C \times 1.5$ *(low-income countries)*					
Δ GDP	6.9%	17.9%	36.7%	17.9%	17.9%
Δ Consumption	5.2%	16.0%	34.8%	16.0%	16.0%
Δ After-Tax Wage	4.6%	15.7%	35.2%	15.7%	15.7%

5 Extensions

5.1 Extend the Model with a Traditional Sector

We extend the baseline model by adding a nontaxable traditional sector. Household utility becomes:

$$U(C, L_M + L_C, L_T) = \frac{C^{1-\gamma}}{1-\gamma} - \alpha_M \frac{(L_M + L_C)^{1+1/\varepsilon}}{1+1/\varepsilon} - \alpha_T \frac{L_T^{1+1/\varepsilon}}{1+1/\varepsilon}$$

The traditional sector produces $Y_T = A_T L_T$ using only labor, with no capital input. The wage is $w_T = A_T$, and since the sector is untaxed, households earn the full marginal product. We fix $\alpha_M = 1$ and $\alpha_T = 75$, and calibrate A_T alongside the existing parameters (L_C, A_M, A_C, ξ) to match each country's agricultural employment share.

Figure 7 shows the calibrated productivities for all three sectors.

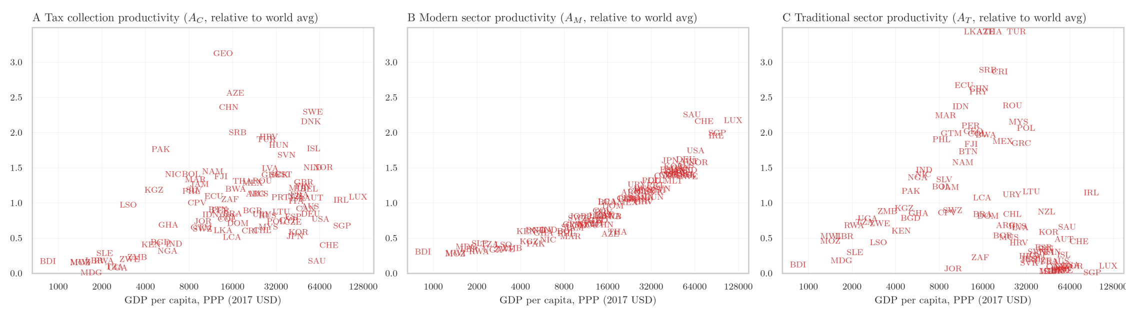


Figure 7: Productivities

Table 7 reports the counterfactuals by income groups.

Table 7: Effects of Policy Counterfactuals by Income Bracket

Shock	Income group	GDP	Consumption	K_G	After-tax wage	T / X	$\Delta\tau$ (pp)
$A_C \times 1.5$	Low income	1.070	1.056	1.998	1.206	1.602	0.018
$A_C \times 1.5$	Middle income	1.052	0.991	1.604	0.869	1.380	0.153
$A_C \times 1.5$	High income	1.047	0.960	1.579	0.822	1.369	0.169
$A_M \times 1.5$	Low income	1.518	1.484	4.050	2.743	1.421	-0.000
$A_M \times 1.5$	Middle income	2.084	1.946	2.744	2.468	1.103	-0.000
$A_M \times 1.5$	High income	2.274	2.269	2.286	2.344	0.976	-0.000
$L_C \times 1.5$	Low income	1.000	1.000	1.004	1.001	0.680	0.001
$L_C \times 1.5$	Middle income	0.997	0.997	1.004	0.979	0.673	0.007
$L_C \times 1.5$	High income	1.000	0.997	1.013	0.990	0.678	0.007
$\theta = 0.84$	Low income	1.064	1.060	1.285	1.203	1.065	-0.000
$\theta = 0.84$	Middle income	1.079	1.066	1.137	1.115	1.019	-0.000
$\theta = 0.84$	High income	1.000	1.000	1.000	1.000	1.000	-0.000

Table 8 shows the prediction of a 50% A_C shock to lower income countries under Hulten's Theorem. Under the calculation of Hulten's Theorem, the tax collection sector is small measured by its wage input $w_M L_C/Y$ or its product T/Y. Consequently, the predicted gain from the change in tax sector productivity is minimal.

Table 8: Hulten's Theorem vs Model predictions for $+50\%$ A_C shock (LIC)

Prediction Method	Δ GDP (%)
Model output	6.973
Hulten (wage share: $0.5 \times w_M L_C/Y$)	0.018
Hulten (tax share: $0.5 \times T/Y$)	0.799
Public capital effect ($\zeta \times \%\Delta K_G \times Y_M/GDP$)	7.231

Notes: Entries are LIC means for the $A_C \times 1.5$ shock among solved countries. The two Hulten rows are direct accounting terms evaluated at baseline quantities. The public-capital row is the first-order modern-sector contribution $\zeta \times \%\Delta K_G$, scaled by the mean baseline modern-sector GDP share Y_M/GDP, so that it is comparable to aggregate GDP.

5.2 Unmechanized Agriculture: A Robustness Check

Agricultural employment data presents a measurement challenge: in rich countries, agriculture is mechanized and pays taxes, making it a poor proxy for the untaxed traditional sector we aim to model. We therefore construct an alternative target based on unmechanized agricultural employment:

$$\text{Unmechanized Agricultural Employment}_i = \text{Agricultural Employment Share}_i \times (1 - \text{Mechanization}_i)$$

Mechanization is modeled as linear in log GDP per capita, with the four-country average for Malawi, Nigeria, Tanzania, and Uganda set to 31.9%, and the richest country set to 100%. We then recalibrate A_T to match this constructed moment.

Figure 8 shows the recalibrated productivities.

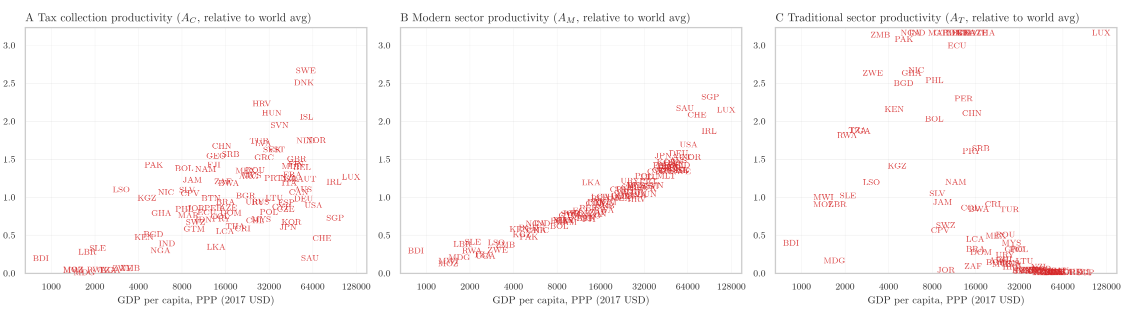


Figure 8: Productivities with traditional sector targeted to unmechanized agricultural sector

Table 9 reports the policy counterfactuals under the unmechanized calibration.

Table 9: Effects of Policy Counterfactuals by Income Bracket

Shock	Income group	GDP	Consumption	K_G	After-tax wage	T / X	$\Delta\tau$ (pp)
$A_C \times 1.5$	Low income	1.058	1.046	2.025	1.215	1.616	0.014
$A_C \times 1.5$	Middle income	1.106	1.028	1.696	0.957	1.440	0.127
$A_C \times 1.5$	High income	1.051	0.962	1.585	0.828	1.373	0.167
$A_M \times 1.5$	Low income	1.448	1.406	4.151	2.774	1.450	-0.000
$A_M \times 1.5$	Middle income	2.215	2.175	2.414	2.379	1.011	-0.000
$A_M \times 1.5$	High income	2.278	2.278	2.277	2.342	0.973	-0.000
$L_C \times 1.5$	Low income	1.000	1.000	1.004	1.000	0.680	0.001
$L_C \times 1.5$	Middle income	1.002	1.000	1.016	0.997	0.679	0.005
$L_C \times 1.5$	High income	1.000	0.997	1.014	0.991	0.678	0.007
$\theta = 0.84$	Low income	1.053	1.048	1.299	1.207	1.072	-0.000
$\theta = 0.84$	Middle income	1.092	1.087	1.115	1.108	1.005	-0.000
$\theta = 0.84$	High income	1.000	1.000	1.000	1.000	1.000	-0.000

5.3 Endogenize Tax Collection

The baseline calibration matches the size of the tax-collection sector, L_C, to its observed employment share. This treats the government's choice of collection capacity as given by data rather than as a policy decision. Alternatively, we can ask: if the government chose L_C to maximize steady-state household welfare rather than inheriting the observed level, how far would optimal collection labor deviate from the calibrated value, and what are the welfare consequences?

We optimize L_C for each calibrated country separately, holding all other parameters fixed at their baseline calibrated values. For each candidate L_C, we resolve the full competitive-equilibrium steady state and evaluate one-period household utility.

Table 10 reports the results by income group.

Table 10: Optimal collection labor by income group: calibrated vs welfare-maximizing L_C

Income group	L_C/L_C^*	L_C/L (%)	Δ Utility	Δ Consumption (%)	Δ GDP (%)	Δ K_G (%)
Low income	0.6703	0.0719	0.0033	0.1779	0.1946	0.8396
Middle income	0.7158	0.2147	0.0052	0.2673	0.4230	2.1460
High income	1.7457	0.2721	0.0079	0.8296	0.3563	-0.5355

Figure 9 visualizes the divergence between calibrated and optimal collection employment share, L_C/L versus L_C^*/L^*, plotted against GDP per capita.

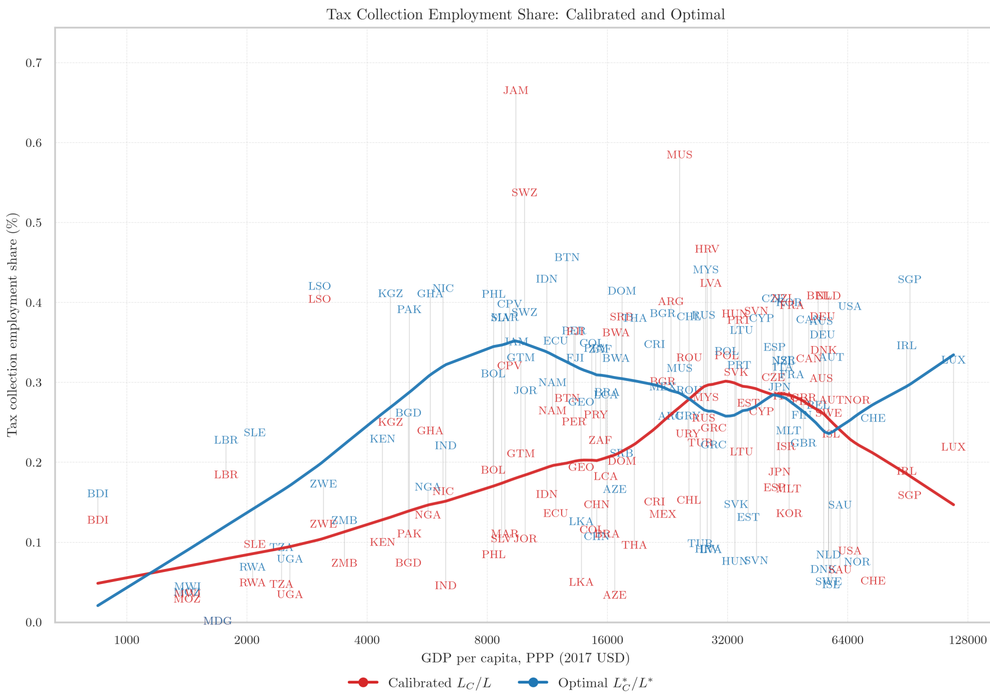


Figure 9: Optimal tax collection share

6 Conclusion

This paper develops the concept of tax productivity – the efficiency with which governments convert administrative expenditures into tax revenues – and studies its relationship with economic development. Using comprehensive data covering 100 countries across a wide range of income levels and spanning multiple decades, we document a robust positive association between tax collections per expenditure on collections and GDP per capita, observed consistently in both cross-sectional and longitudinal analyses. We interpret these empirical patterns through a quantitative growth model that incorporates tax administration into neoclassical and Ramsey frameworks. The model highlights that increased tax productivity enhances a government's ability to finance productive public investments, thus contributing to higher levels of economic output. Quantitative findings so far indicate that improvements in tax productivity could substantially impact aggregate outcomes, and significantly more than suggested by the tax administration sector's rather modest share of total employment. Low-income countries today, according to our model's calculations, likely have inefficiently low levels of public capital to output due in large part to their low levels of tax productivity.

References

ASCHAUER, D. A. (1989): "Is Public Expenditure Productive?," *Journal of Monetary Economics*, 23(2), 177–200.

BACHAS, P., M. FISHER-POST, A. JENSEN, AND G. ZUCMAN (2024): "Globalization, Capital Taxation, and Development: Evidence from a Macro-Historical Database," Unpublished Working Paper, Harvard University.

BACHAS, P., L. GADENNE, AND A. JENSEN (2024): "Informality, Consumption Taxes, and Redistribution," *Review of Economic Studies*, 91(5).

BALAN, P., A. BERGERON, G. TOUREK, AND J. L. WEIGEL (2022): "Local Elites as State Capacity," *American Economic Review*, 112(3).

BARRO, R. J. (1990): "Government Spending in a Simple Model of Endogeneous Growth," *Journal of Political Economy*, 98(5).

BAXTER, M., AND R. G. KING (1993): "Fiscal Policy in General Equilibrium," *American Economic Review*, 83(3), 315–334.

BESLEY, T., E. ILZETZKI, AND T. PERSSON (2013): "Weak States and Steady States," *American Economic Journal: Macroeconomics*, 5(4), 203–235.

BESLEY, T., AND T. PERSSON (2011): *Pillars of Prosperity*. Princeton University Press.

BHANDARI, A., D. EVANS, Y. YAO, AND E. R. MCGRATTAN (2024): "Business Income Underreporting and Public Finance," Federal Reserve Bank of Minneapolis Working Paper.

BICK, A., N. FUCHS-SCHÜNDELN, AND D. LAGAKOS (2018): "How Do Hours Worked Vary with Income? Cross-Country Evidence and Implications," *American Economic Review*, 108(1), 170–199.

BICK, A., N. FUCHS-SCHÜNDELN, D. LAGAKOS, AND H. TSUJIYAMA (2022): "Structural Change in Labor Supply and Cross-Country Differences in Hours Worked," *Journal of Monetary Economics*, 130, 68–85.

BOM, P. R., AND J. E. LIGTHART (2014): "What Have We Learned From Three Decades of Research on the Productivity of Public Capital?," *Journal of Economic Surveys*, 28(5), 889–916.

BONING, W. C., N. HENDREN, B. SPRUNG-KEYSER, AND E. STUART (2025): "A Welfare Analysis of Tax Audits Across the Income Distribution," *Quarterly Journal of Economics*, 140(1).

BOPPART, T., P. KIERNAN, P. KRUSELL, AND H. MALMBERG (2023): "The Macroeconomics of Intensive Agriculture," NBER Working Paper 31101].

BOPPART, T., AND P. KRUSELL (2020): "Labor Supply in the Past, Present, and Future: A Balanced-Growth Perspective," *Journal of Political Economy*, 128(1), 118–157.

BROCKMEYER, A., A. ESTEFAN, K. ARRAS, AND J. C. SUÁREZ SERRATO (2023): "Taxing Property: Theory and Evidence from Mexico," *Working Paper*.

BROCKMEYER, A., S. SMITH, M. HERNANDEZ, AND S. KETTLE (2019): "Casting a Wider Tax Net: Experimental Evidence from Costa Rica ," *American Economic Journal: Economic Policy*, 11(3).

CASELLI, F. (2005): "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf.

CHAMLEY, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, pp. 607–622.

CHARI, V. V., AND P. J. KEHOE (1999): "Optimal Fiscal and Monetary Policy," in *Handbook of Macroeconomics, Volume 1C*, ed. by J. B. Taylor, and M. Woodford, pp. 1671–1745. Elsevier, Amsterdam.

CHRISTIANS, A. (2013): "Drawing the boundaries of tax justice," *Working Paper*.

COLMER, J. M., D. LAGAKOS, AND M. SHU (2024): "Is the Electricity Sector a Weak Link in Development?," NBER Working Paper 32041.

CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, 99(1), 25–48.

CUBAS, G. (2020): "Public Capital and Economic Development," *Economic Journal*, 130(632), 2354–2381.

D'ARCY, M., M. NISTOTSKAYA, AND O. OLSSON (2024): "Cadasters and Economic Growth: A Long-Run Cross-Country Panel," *Journal of Political Economy*, 132(11).

DEVADAS, S., AND S. PENNINGS (2018): "Assessing the Effect of Public Capital on Growth: An Extension of the World Bank Long-Term Growth Model," Policy Research Working Paper 8604, World Bank, Development Economics Development Research Group.

DINCECCO, M., AND G. KATZ (2016): "State Capacity and Long-Run Economic Performance," *Economic Journal*, 126(590).

DONALDSON, D., AND R. HORNBECK (2016): "Railroads and American Economic Growth: A "Market Access" Approach," *Quarterly Journal of Economics*, 131(2), 799–858.

DZANSI, J., A. JENSEN, D. LAGAKOS, AND H. TELLI (2025): "Technology and Tax Capacity," *NBER Working Paper No. 29923*.

EU (2015): "Tax Revenue Mobilization in Developing Countries," *Technicsl Report*.

FABER, B. (2014): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 81(3), 1046–1070.

FARHI, E., AND I. WERNING (2016): "Fiscal Multipliers: Liquidity Traps and Currency Unions," in *Handbook of Macroeconomics, Volume 2*, ed. by J. B. Taylor, and H. Uhlig, pp. 2417–2492. Elsevier, Amsterdam.

GENNAIOLI, N., AND H.-J. VOTH (2015): "State Capacity and Military Conflict," *Review of Economic Studies*, 82(4).

GLOMM, G., AND B. RAVIKUMAR (1992): "Public Versus Private Investment in Human Capital: Endogenous Growth andIncome Inequality," *Journal of Political Economy*, 100(4), 818–834.

GOLLIN, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110(2), 458–474.

GORDON, R., AND W. LI (2009): "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93(7-8).

HERRENDORF, B., AND Á. VALENTINYI (2012): "Which Sectors Make Poor Countries So Unproductive?," *Journal of the European Economic Association*, 10(2), 323–341.

HJORT, J., AND L. TIAN (2024): "The Economic Impact of Internet Connectivity in Developing Countries," *Annual Review of Economics*, Forthcoming.

HSIEH, C.-T., AND P. J. KLENOW (2007): "Relative Prices and Relative Prosperity," *American Economic Review*, 97(3), 562–585.

IMF (2001): "The Modern VAT," *IMF Publications*.

——— (2015): "Current Challenges in Revenue Mobilization: Improving Tax Compliance," *Staff Report*.

JENSEN, A. (2022): "Employment Structure and the Rise of the Modern Tax System," *American Economic Review*, 112(1), 213–234.

JONES, L. E., R. E. MANUELLI, AND P. E. ROSSI (1993): "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, 101(3), 485–517.

JUDD, K. L. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28(1), 59–83.

KHAN, A., A. KHWAJA, AND B. OLKEN (2015): "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics*, 131(1).

KLEVEN, H., C. KREINER, K. LARSEN, AND J. SØGAARD (2025): "Micro versus Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort," *American Economic Review*, 115(9), 2849–90.

KLEVEN, H. J., M. B. KNUDSEN, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): "Unwilling or Unable to Cheat?," *Econometrica*, 79(3).

KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2016): "Why can Modern Governments Tax so Much?," *Economica*, 83(330).

LEEPER, E. M., T. B. WALKER, AND S.-C. S. YANG (2010): "Government Investment and Fiscal Stimulus," *Journal of Monetary Economics*, 57(8), 1000–1012.

MAYSHAR, J., O. MOAV, AND L. PASCALI (2022): "The Origin of the State: Land Productivity or Appropriability?," *Journal of Political Economy*, 130(4).

NARITOMI, J. (2019): "Consumers as Tax Auditors," *American Economic Review*, 109(9), 3031–72.

OECD (2024): "Tax Administration 2022: Comparative Information on OECD and Other Advanced and Emerging Economies," *Annual Report*.

OKUNOGBE, O., AND G. TOUREK (2024): "How Can Lower-Income Countries Collect More Taxes?," *Journal of Economic Perspectives*, 38(1).

POMERANZ, D. (2015): "No Taxation without Information," *American Economic Review*, 105(8).

PRITCHETT, L. (2000): "The Tyranny of Concepts: CUDIE (Cumulated, Depreciated, Investment Effort) Is Not Capital," *Journal of Economic Growth*, 5(4), 361–384.

RAMEY, V. A. (2011): "Can Government Purchases Stimulate the Economy?," *Journal of Economic Literature*, 49(3), 673–685.

——— (2021): "The Macroeconomic Consequences of Infrastructure Investment," in *Economic Analysis and Infrastructure Investment*, ed. by E. L. Glaeser, and J. M. Poterba, pp. 219–268. National Bureau of Economic Research, Cambridge, MA.

RAMEY, V. A., AND N. FRANCIS (2009): "A Century of Work and Leisure," *American Economic Journal: Macroeconomics*, 1(2), 189–224.

RESTUCCIA, D., D. T. YANG, AND X. ZHU (2008): "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis," *Journal of Monetary Economics*, 55, 234–250.

STRAUB, L., AND I. WERNING (2020): "Positive Long-Run Capital Taxation: Chamley-Judd Revisited," *American Economic Review*, 110(1), 86–119.

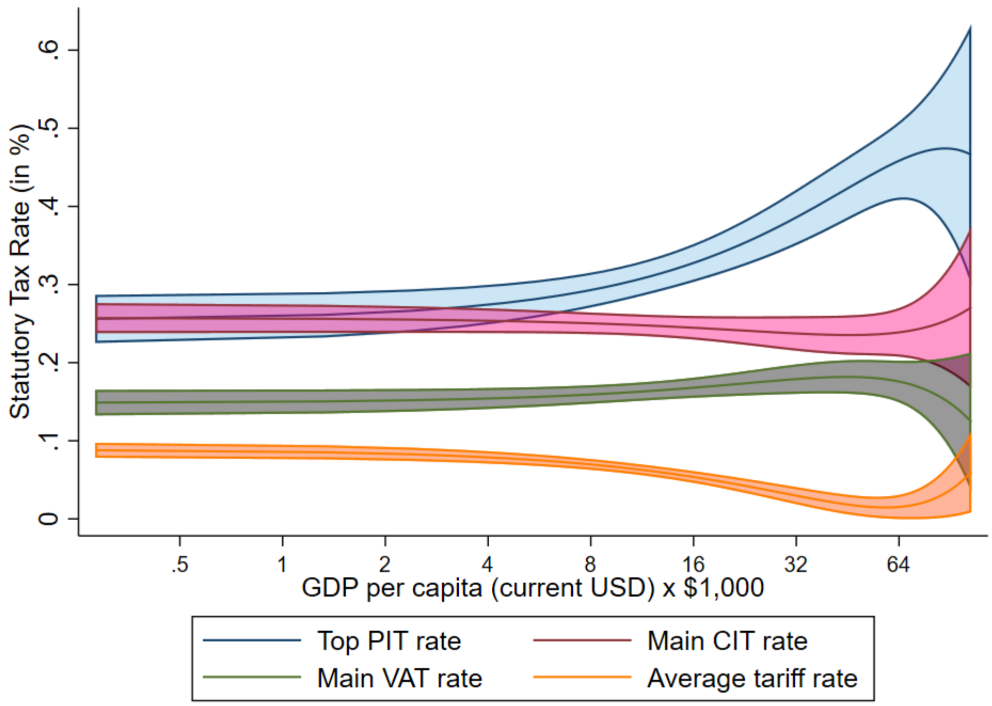
SUGIN, L. (2014): "Payroll Taxes, Mythology, and Fairness.," *Harvard Journal on Legislation*, 51.

TILLY, C. (1990): *Coercion, Capital, and European States, AD 990–1990*. Blackwell.

A Online Appendix

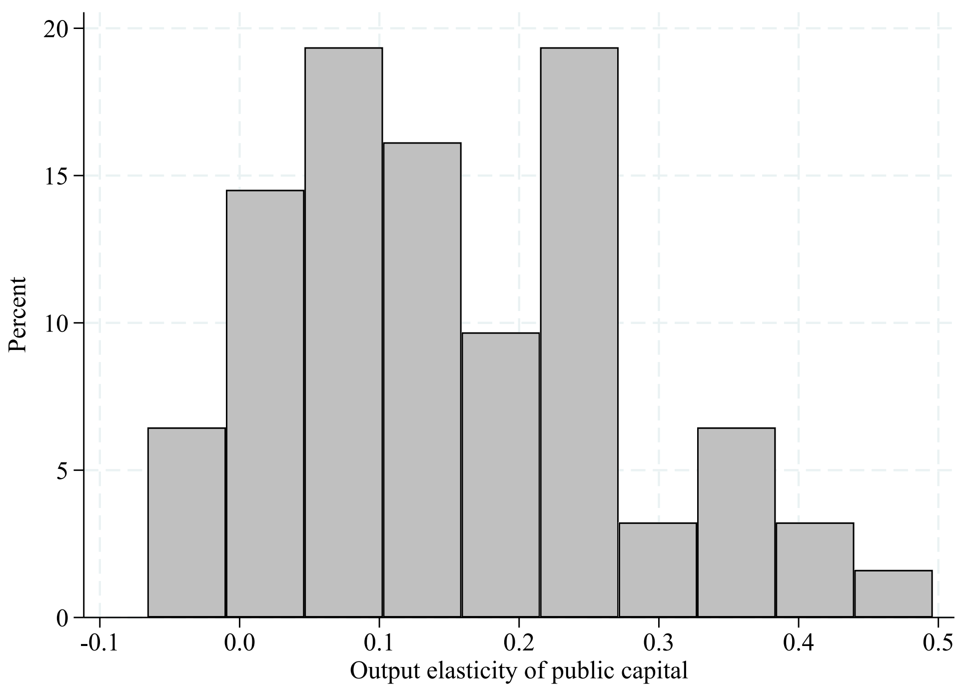
A.1 Appendix Figures and Tables

Figure A.1: Statutory Tax Rates by GDP per Capita



Note: This figure plots predicted average statutory rates for personal income taxes (PIT), corporate income taxes (CIT), value added taxes (VAT), and tariffs by GDP per capita. Source: Bachas, Fisher-Post, Jensen, and Zucman (2024).

Figure A.2: Bom and Ligthart (2014) Output Elasticities of Public Capital



Notes: This figure plots a histogram of output elasticities of public capital from the meta-analysis of Bom and Ligthart (2014).