STEG virtual course

Data in Macro Development: Public Finance and Expenditures

Anders Jensen
(HKS and NBER)

Outline of today's talk

1. Motivation

2. Relevant data for country-level research
   A. Tax collection
   B. Tax burdens: Statutory and Effective
   C. Equity
   D. Informality

3. Relevant data for within-country research

Outline of today's talk

1. **Motivation**

2. Relevant data for country-level research
   A. Tax collection
   B. Tax burdens: Statutory and Effective
   C. Equity
   D. Informality

3. Relevant data for within-country research

Tax capacity

- ***Tax capacity***: Government's ability to collect taxes from firms and citizens
  - At a reasonable administrative cost, with progressive distributional impacts and limited economic distortions

- Alleviating constraints on fiscal capacity has **multiple benefits**
  - Fund essential public goods and services
  - Curb alarming pre-tax income inequality
  - Reduce exposure to foreign economic and political influences
  - Strengthen state-citizen social contract

Tax capacity and development: Two complementary approaches

- Macro approach: What shapes tax capacity and tax policy in the long run (within countries and across countries)?
  - Devote most of today's talk to discuss data related to this Q

- Micro approach: Given initial (limited) capacity, what can governments do to incrementally improve enforcement, policy, administration and morale
  - Growing literature, based often on deep collaborations with tax authorities and emphasis on causal identification
  - Devote time at end of talk to discuss data related to these Qs

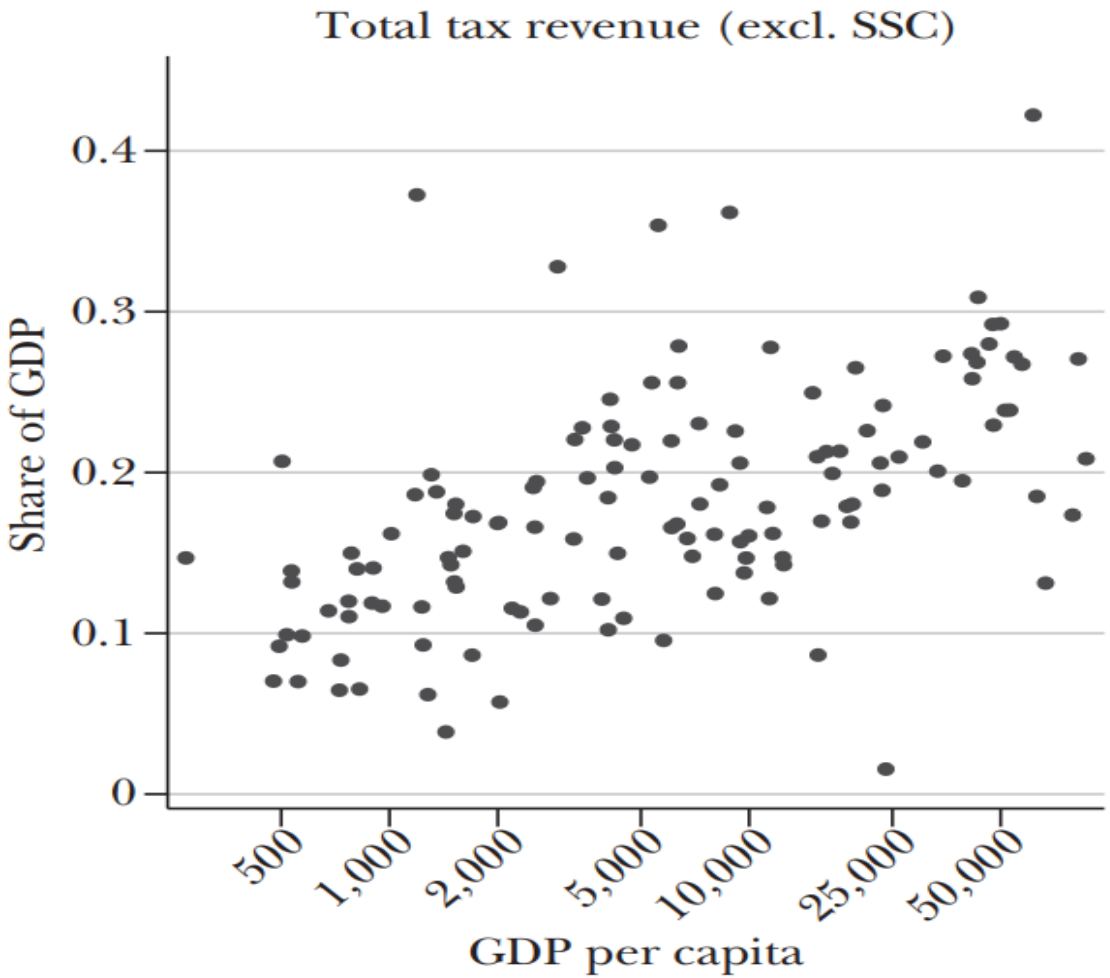
Outline of today's talk

1. Motivation

2. **Relevant data for country-level research**
   - **A. Tax collection**
   - B. Tax burdens: Statutory and Effective
   - C. Equity
   - D. Informality

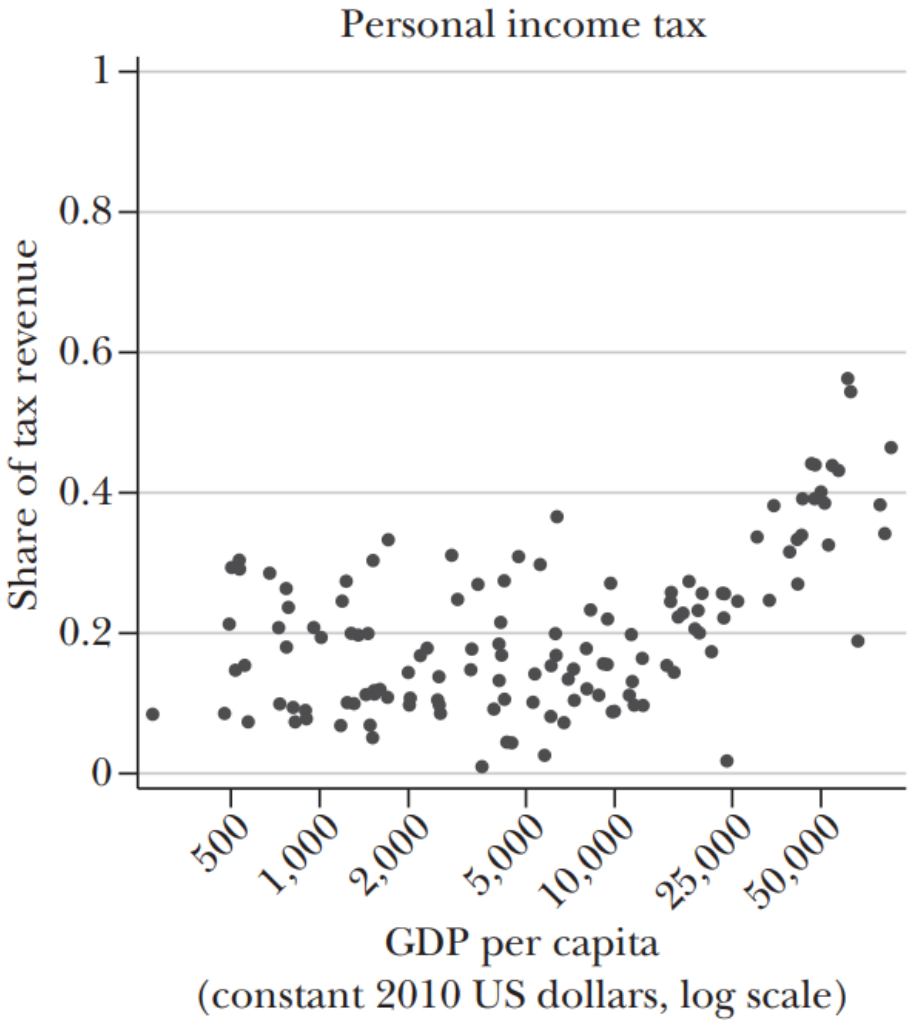
3. Relevant data for within-country research

Cross-country total tax take



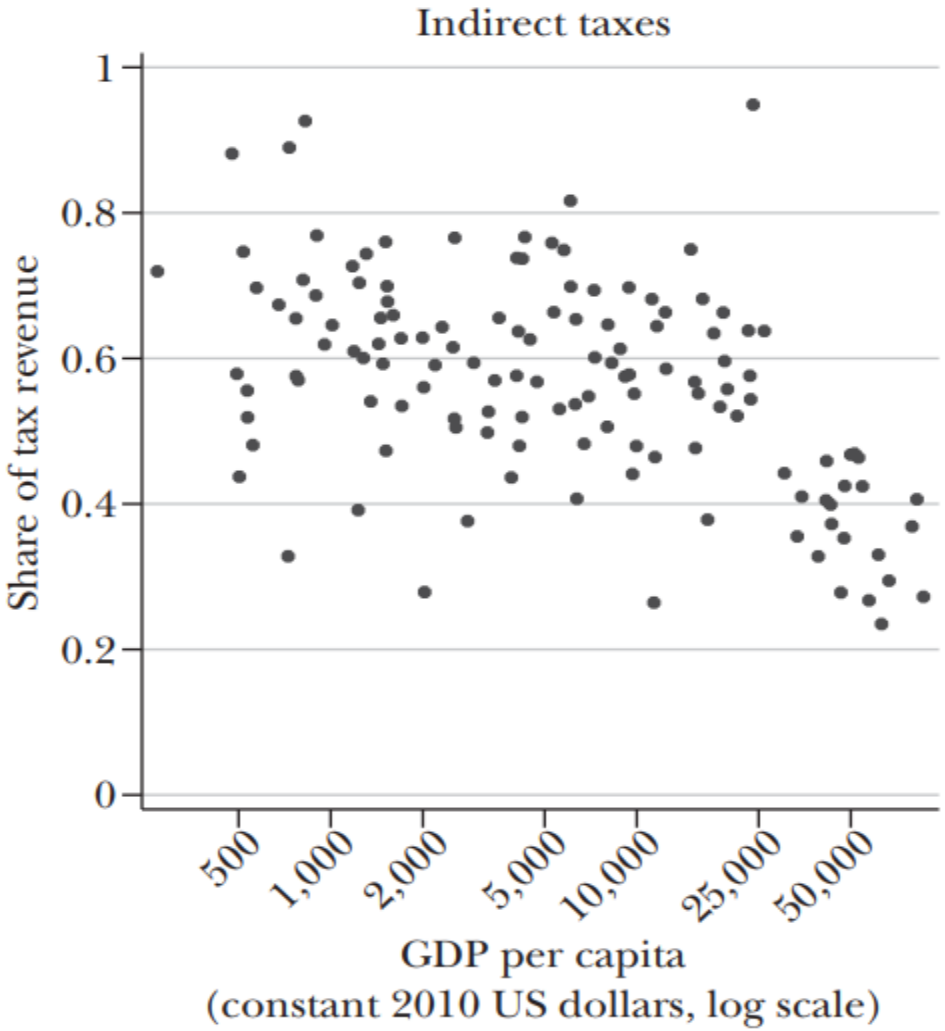
Source: Bachas, Gadenne and Jensen (2024)

Cross-country tax composition: Personal income taxes



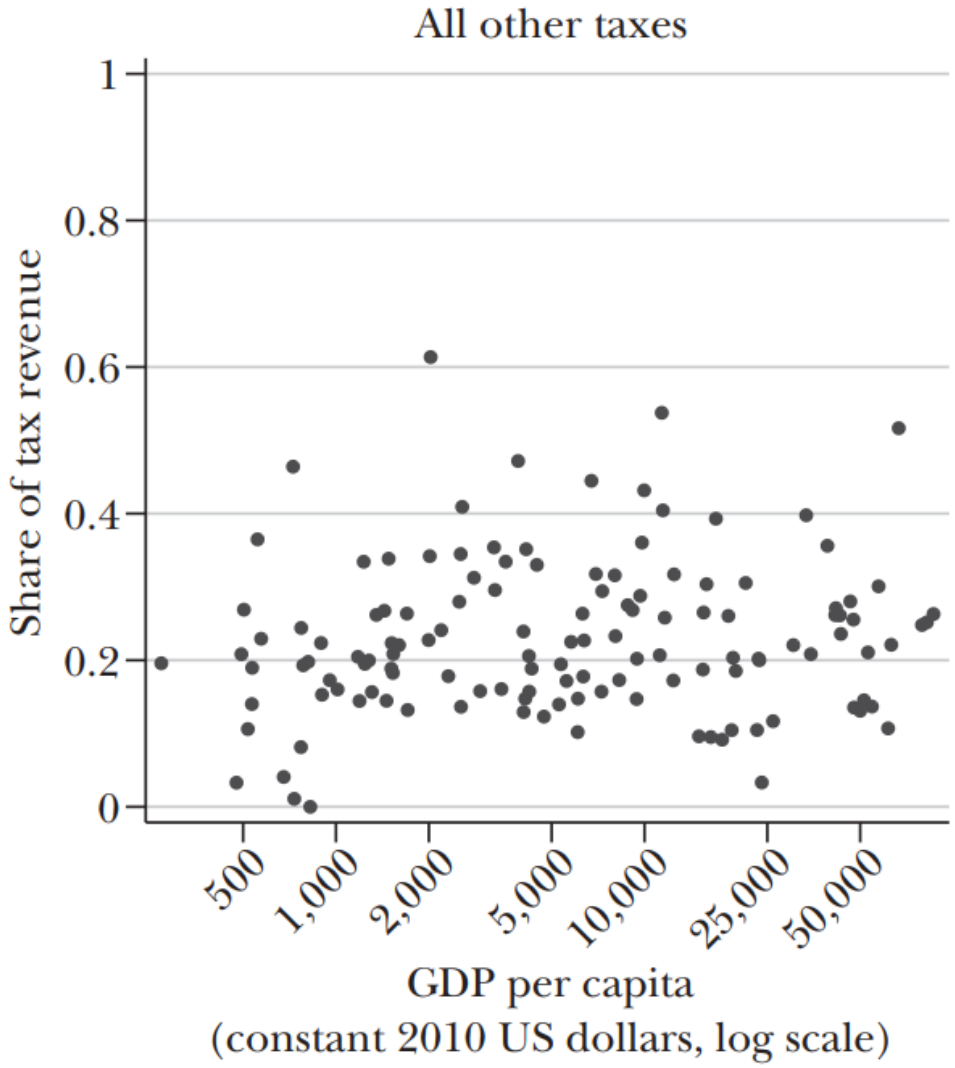
Source: Bachas, Gadenne and Jensen (2024)

Cross-country tax composition: Indirect taxes



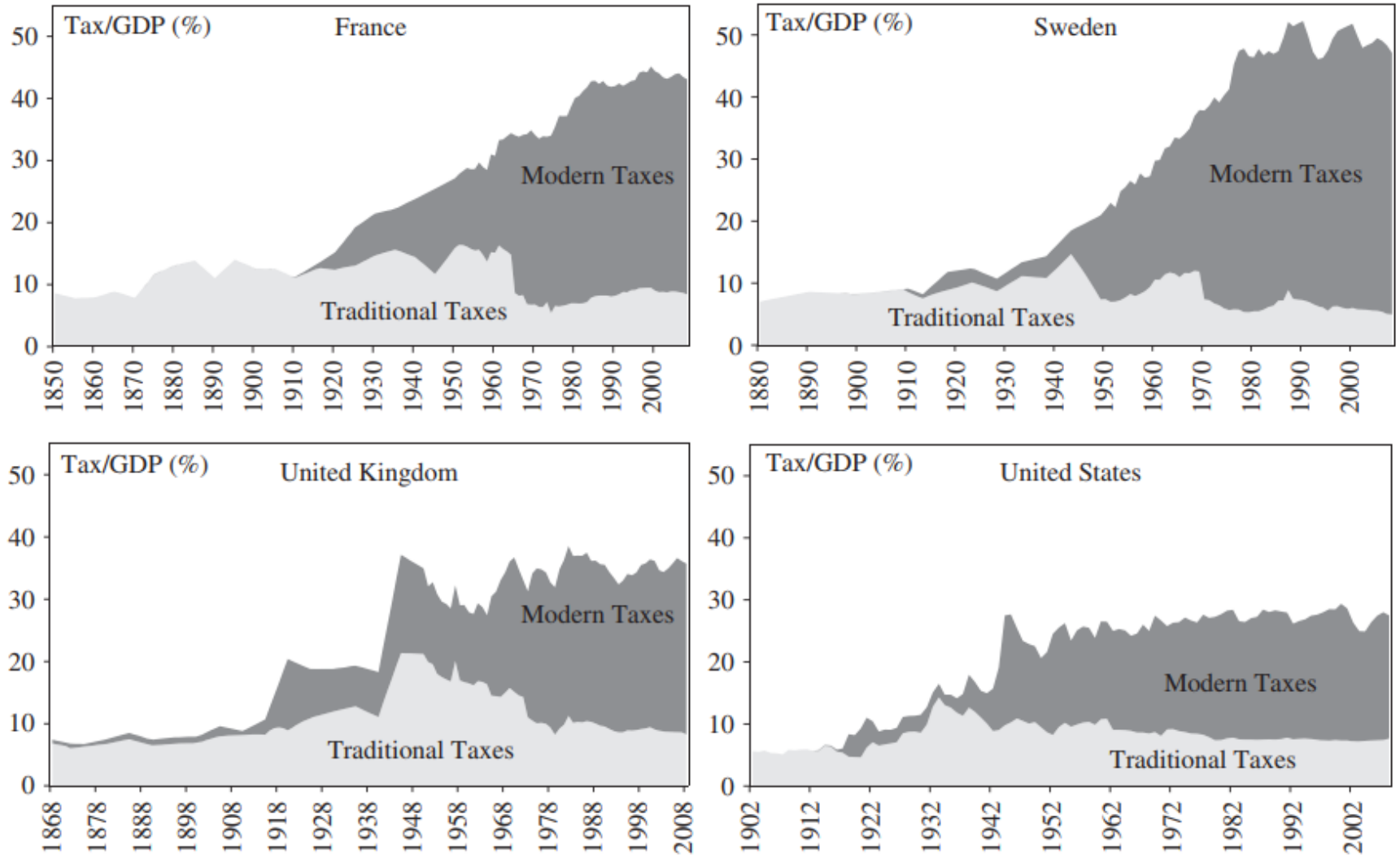
Source: Bachas, Gadenne and Jensen (2024)

Cross-country tax composition: Other taxes



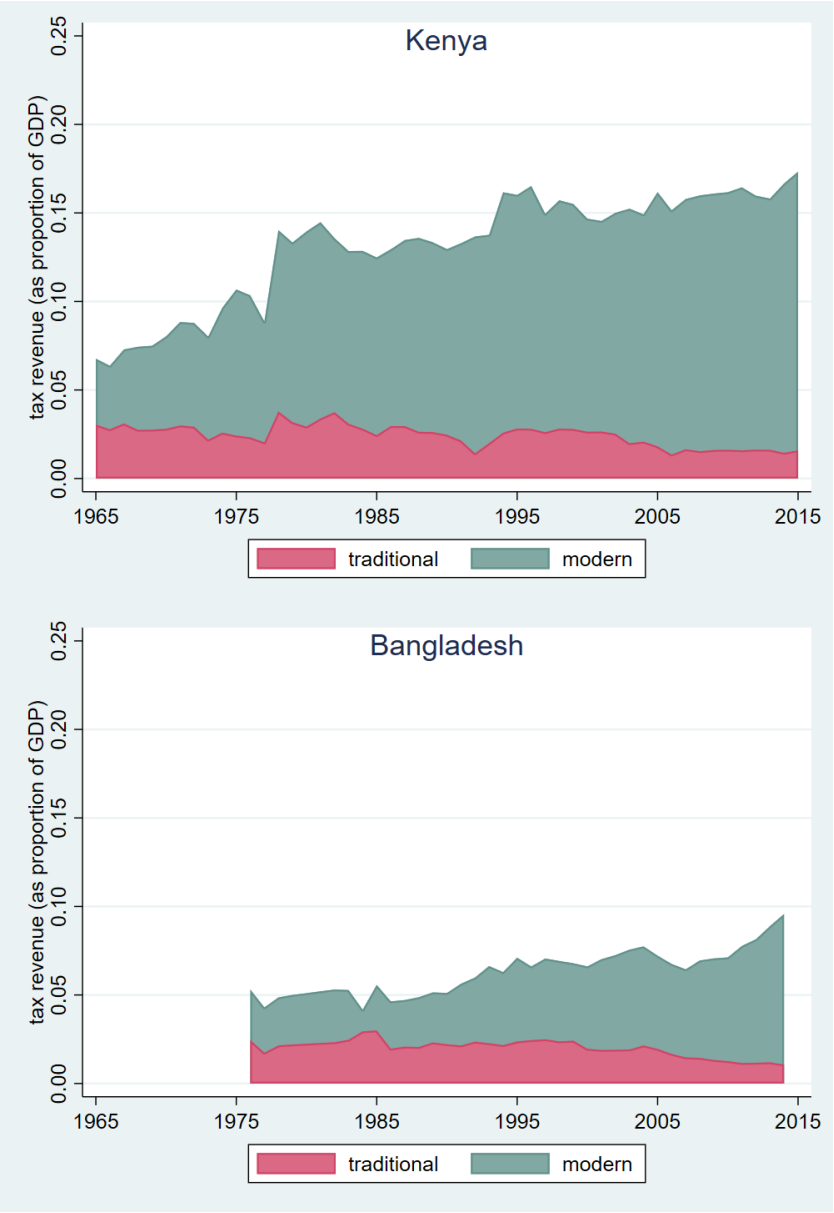
Source: Bachas, Gadenne and Jensen (2024)

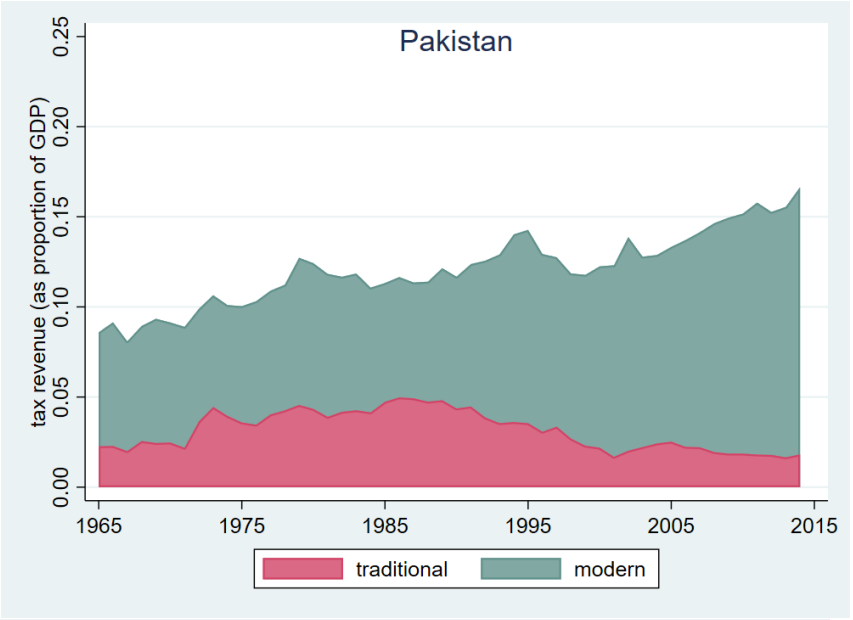
Within country tax take and composition (i)

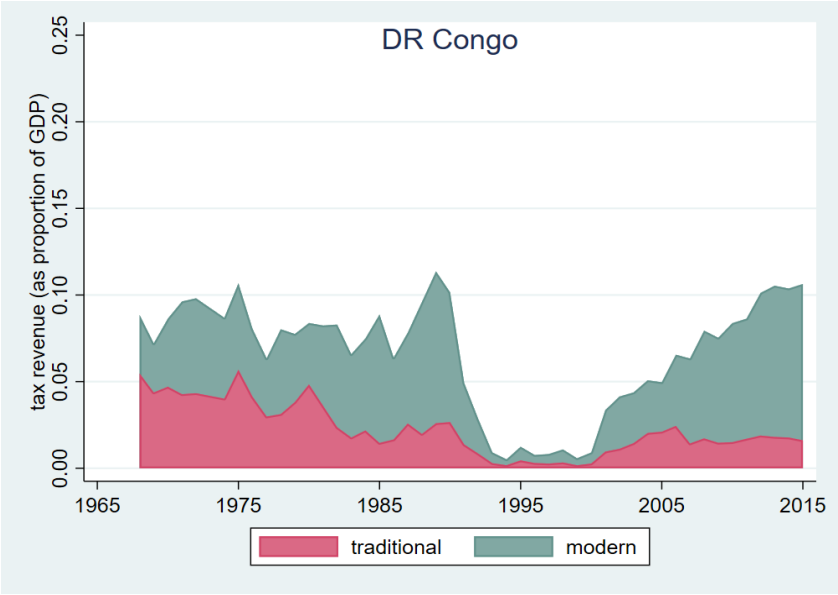


Source: Kleven, Kreiner and Saez (2016)

Within country tax take and composition (ii)







Source: globaltaxation.world

Tax collection: Main data sources

- There exist several publicly available databases created by international organizations
  - OECD Revenue Statistics (link)
  - UNU Wider Government Revenue Dataset (link)

- Bachas et al. (2024) supplement these data with digitized archival data from developing countries
  - Global, historical coverage in 154 countries: 1965-2018
  - Available at https://globaltaxation.world/

Tax collection: Data challenges (i)

- Challenge: Differences in reporting standards
  - Data ultimately derives from individual countries' ministries (often the ministry of finance or the tax authority)
  - Differences in reporting may persist
- Challenge: Social security contributions
  - Coverage in OECD data is of high quality
  - Coverage in globaltaxation.world is meaningful, but questions remain on a country-by-year basis
- Challenge: Sub-national taxes
  - Coverage of sub-national taxes varies with extent of country's fiscal decentralization
  - Common for strongly decentralized countries to not feature a centralized data collection process
  - On-going data efforts: World Observatory on Subnational Government Finance (SNG-WOFI; link)

Tax collection: Data challenges (ii)

- Challenge: Non-tax sources of revenue
  - Resource revenues: On-going data effort by Natural Resource Governance Institute (link)

- Challenge: Long-run coverage
  - Going back historically, focus remains on collecting information on main types of taxes
  - See, for example, recent work on long-run taxation in Africa (1900-2018) by Cogneau et al. (2021)

Outline of today's talk

1. Motivation

2. **Relevant data for country-level research**
   A. Tax collection
   B. Tax burdens: Statutory and Effective
   C. Equity
   D. Informality

3. Relevant data for within-country research

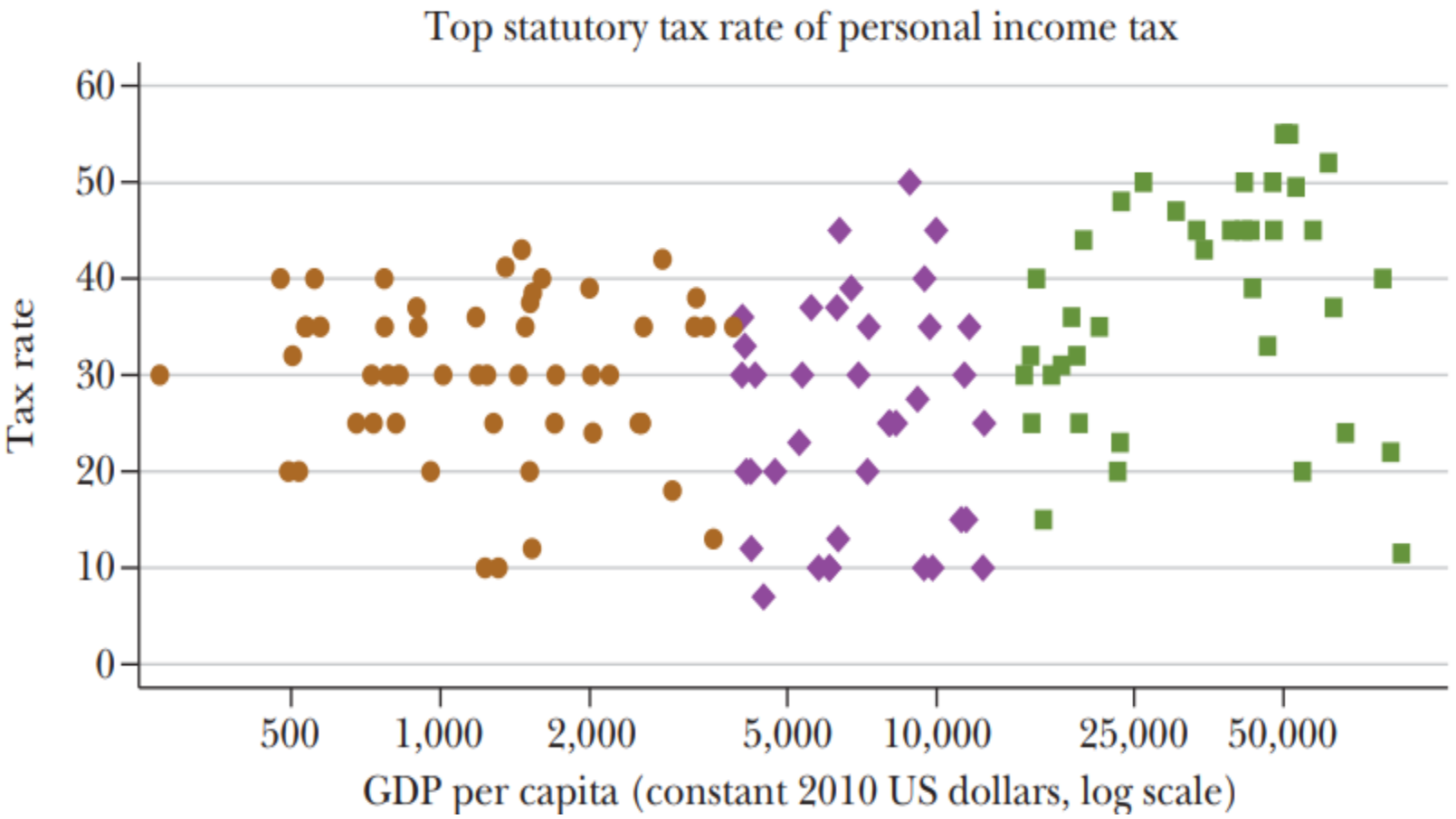
Tax burdens: Statutory and effective measures

There are two relevant types of tax burdens

1. Statutory:
   - Based on tax code
   - Applies rates, deductions and exemptions based on relevant characteristics

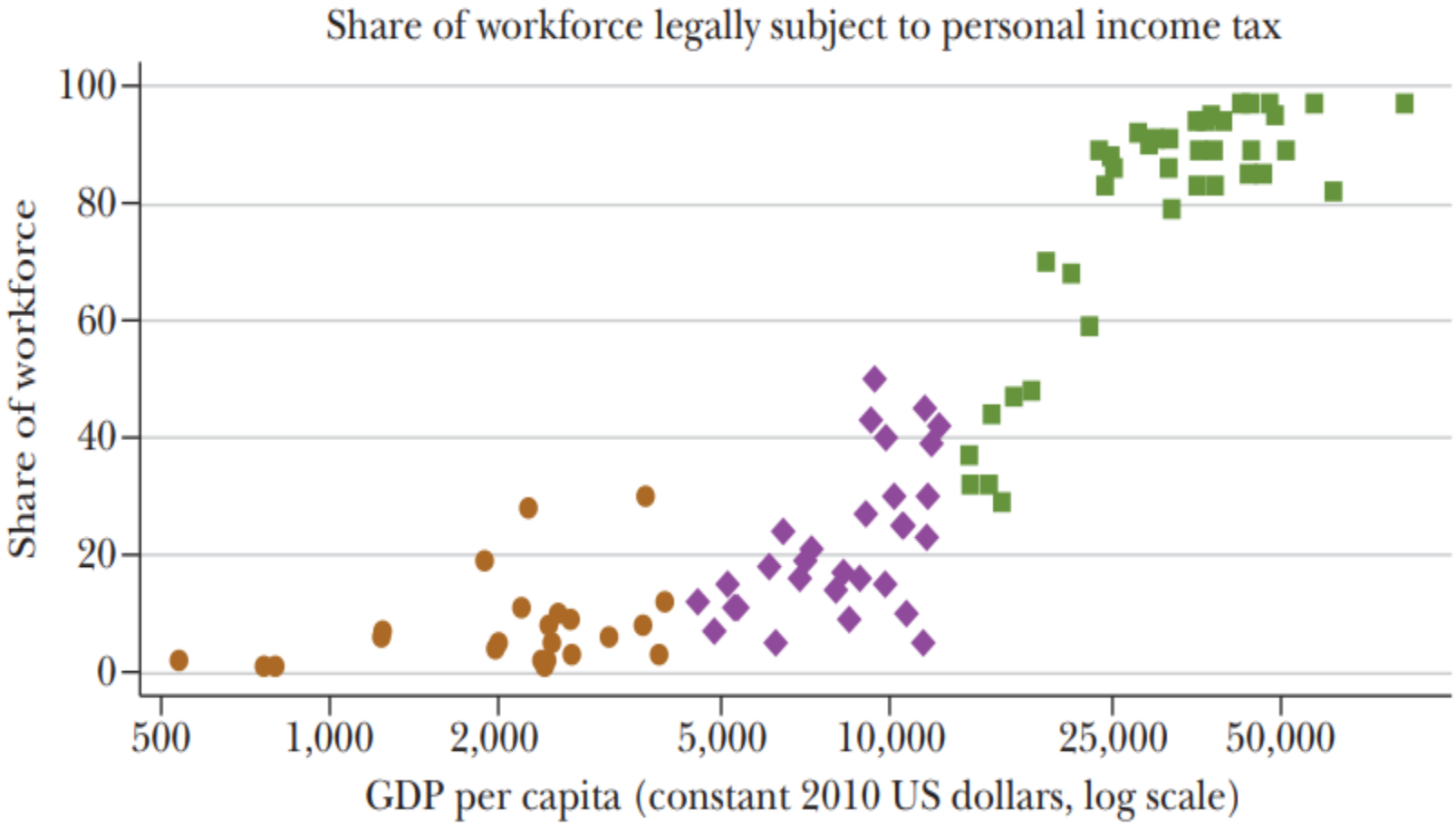
2. Effective:
   - Accounts for tax evasion and avoidance
   - Relates actual taxes paid to taxable base
   - Potential distinction between ‘true tax base’ and ‘reported tax base’

Statutory variation in PIT: Top marginal rate



Source: Bachas, Gadenne and Jensen (2024)

Statutory variation in PIT: Size of tax base



Source: Bachas, Gadenne and Jensen (2024)

Statutory tax burden: Data sources for PIT

- Some dimensions of the statutory tax burden (e.g. top marginal rate) are covered in publicly available datasets
  - See tax summaries by consulting firms (e.g. PwC; link)

- Many statutory dimensions are not systematically measured in developing countries (let alone over a long time-period)
  - E.g. data on PIT base in Jensen (2022) was collected based on individual country reports
  - Good starting point: International Bureau of Fiscal Documentation (link) [requires subscription]

- Research papers have built statutory tax simulators
  - Including Duncan and Peter (2016), Egger et al. (2019)
  - See also Commitment to Equity's Data Center (link)

Statutory tax burden: Data sources for CIT

- Similar to PIT, some dimensions of the statutory corporate income tax (CIT) burden are well captured across countries
  - See Deloitte's International Tax Source database (link)

- Significant research effort has been devoted to comprehensively measure the legal tax burdens by constructing forward-looking tax rates on capital based on the simulated present value of returns and costs of a new investment
  - See Devereux and Griffith (1999)
  - Data that covers developing countries: World Bank (2022), Steinmüller, E., Thunecke, G. U., & Wamser, G. (2019)

Statutory tax burden: Data sources for VAT

- The value-added tax (VAT) is the most important source of indirect taxation in developing countries

- Some of the high-level statutory information is available publicly
  - See PwC's worldwide summary of the VAT (link)

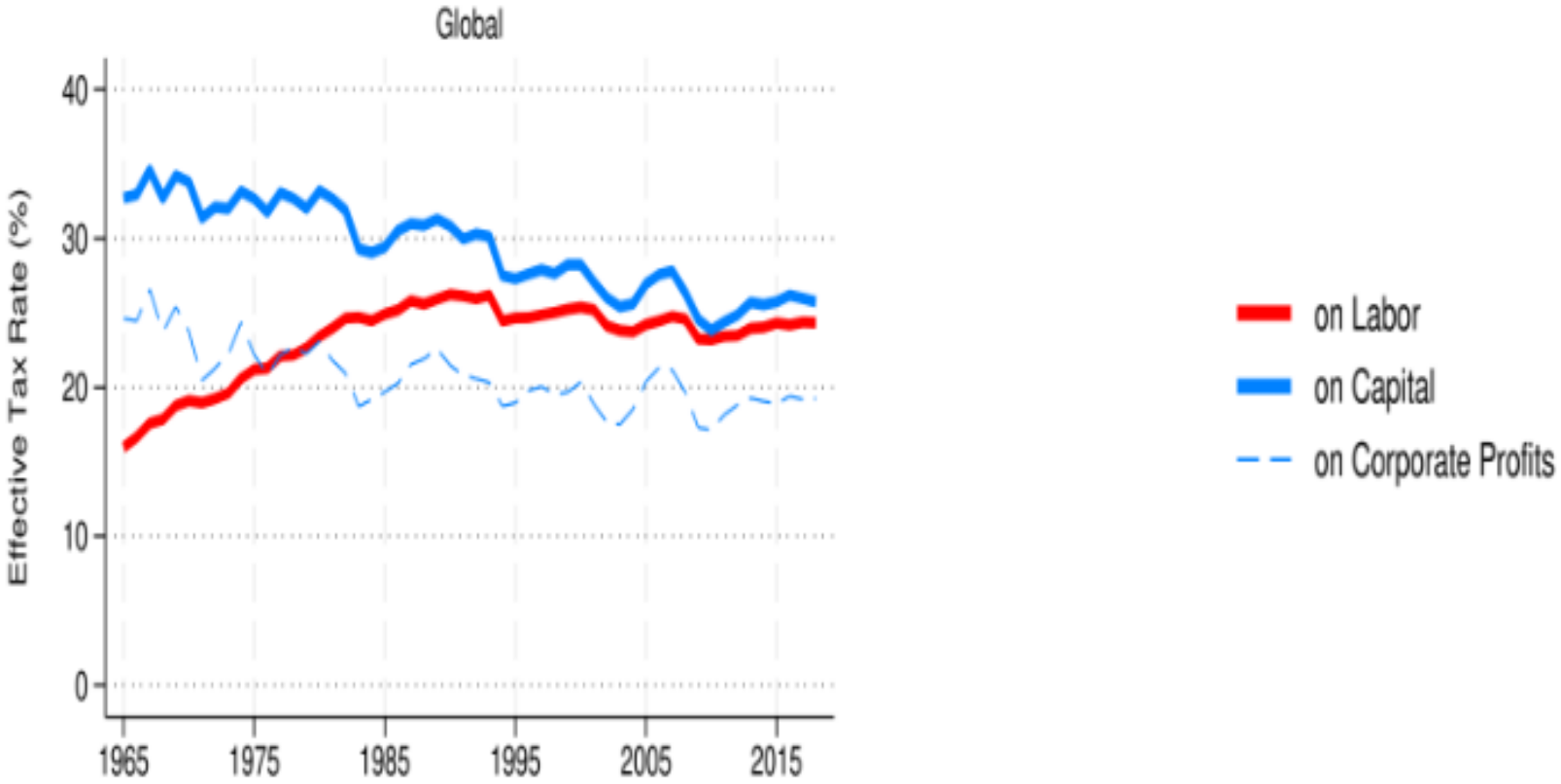
- Like for PIT and CIT, the VAT features a host of additional statutory dimensions with limited cross-country data coverage
  - Multiple rates, exemptions, refund mechanisms
  - Deviate from 'textbook VAT' (single rate, broad base)
  - See Brockmeyer et al. (2024) for detailed discussion

Effective tax burden: Measurement

- Starting from the statutory tax burden, the effective tax burden accounts for all behavioral responses (both real and evasion-driven) and captures how much a person or firm has actually paid in taxes
  - Total tax take is an effective tax burden at the country-level

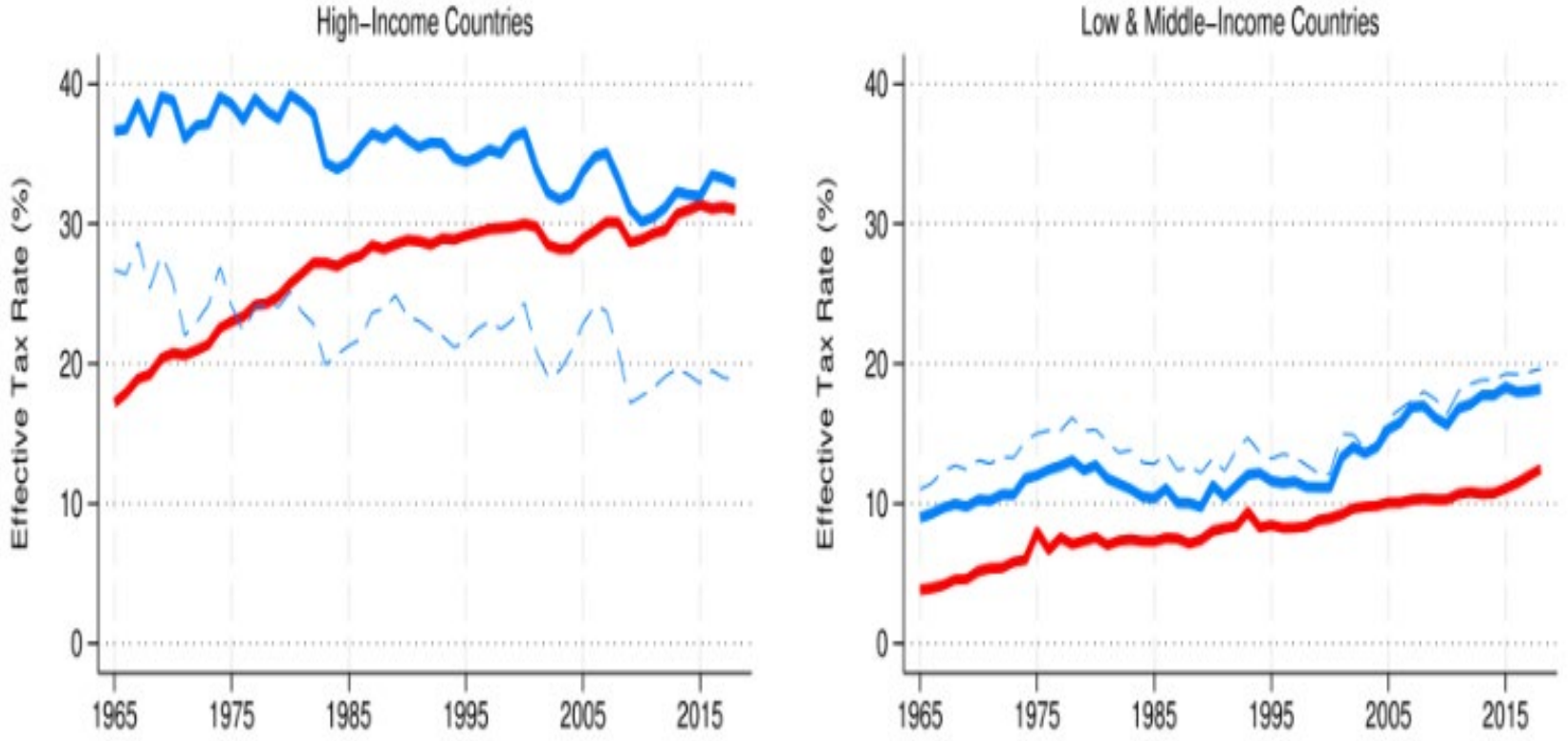
- Recent data-work in Bachas et al. (2024) sheds light on the effective tax burden on capital and labor
  - The effective tax rate on capital divides all taxes collected on capital by the national income that accrues to capital
  - Requires additional data to measure factor shares of national income: UN's System of National Accounts (link)
  - Database of effective tax rates on capital and labor are available at https://globaltaxation.world/ (154 countries, 1965-2018)

Globally, effective tax rates on labor and capital have converged…



Source: Bachas, Fisher-Post, Jensen and Zucman (2024)

(…) but there is significant heterogeneity by development level



Source: Bachas, Fisher-Post, Jensen and Zucman (2024)

Outline of today's talk

1. Motivation

2. **Relevant data for country-level research**
   A. Tax collection
   B. Tax burdens: Statutory and Effective
   C. Equity
   D. Informality

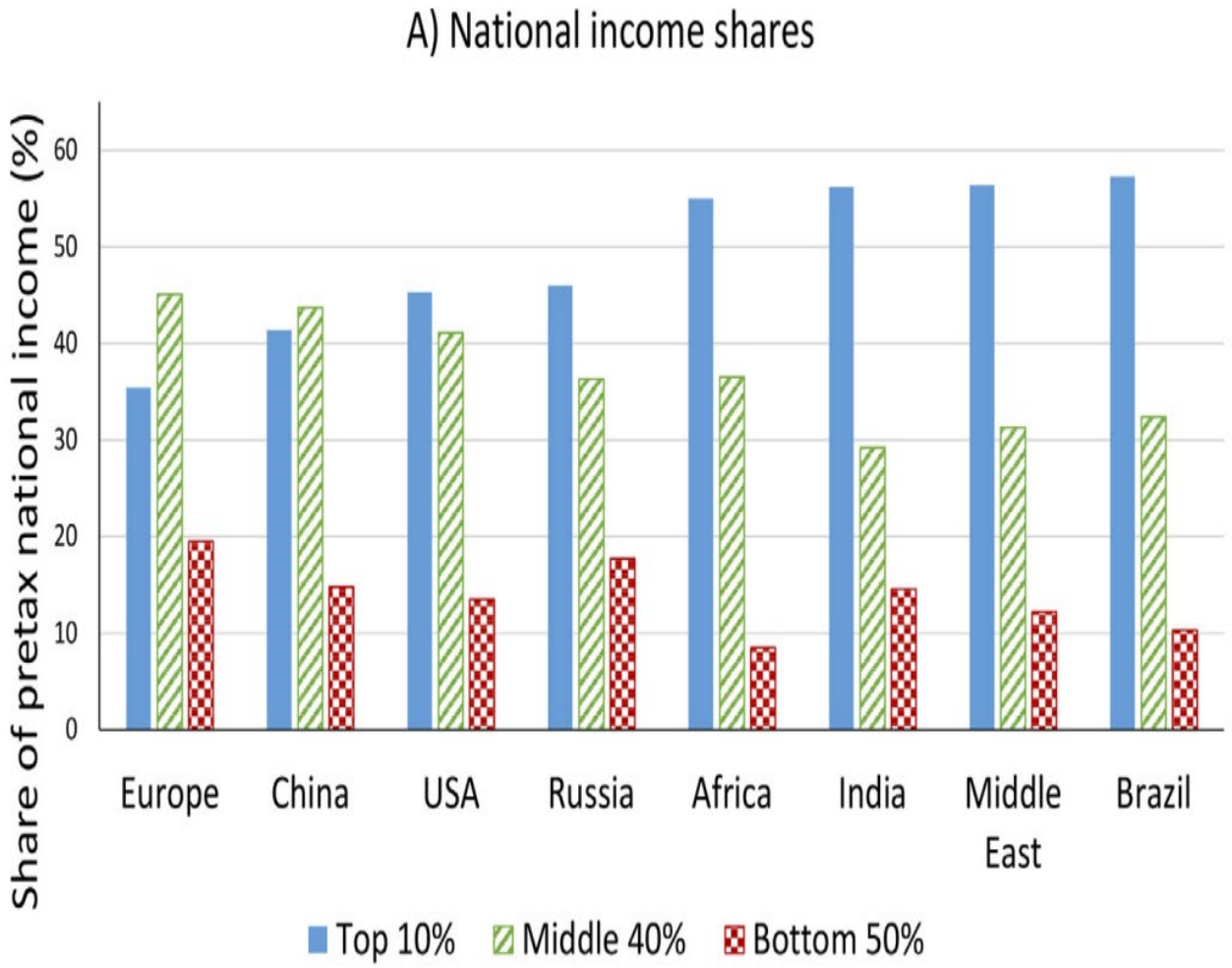
3. Relevant data for within-country research

Tax equity measured across countries

- Two dimensions of equity that are important in public finance:
  - Pre-tax and transfer income inequality
  - Post-tax and transfer income inequality

- Several databases measure pre-tax income inequality across a broad set of countries and over time
  - The World Inequality Lab measures inequality by combining multiple sources: national accounts, survey data, fiscal data, and wealth rankings (link)
  - Many inequality measures used to be based on household survey data, with challenges to measure top incomes
  - Combining multiple data-sources allows WIL to measure income at the very top more accurately and to allocate 100% of national income to individuals (with some assumptions)

Pre-tax income inequality in developing countries is high and persistent



Source: Word Inequality Lab

Measuring inequality post-tax and transfer

- Fiscal incidence analysis is designed to measure who bears the burden of taxes and who receives the benefits of government spending (in particular, social spending)
  - In practice, this allows a comparison of incomes before taxes and transfers versus after them

- Assumptions about behavioral responses can be integrated (or abstracted from in the 'accounting approach)

- The Center for Commitment to Equity (link) makes their fiscal incidence analyses available in many developing countries

Citizen preferences over redistribution

- The extent to which governments implement tax and transfer policies may reflect, at least in part, citizens' preferences for redistribution

- The World Values Survey contain some questions that are related to preferences for redistribution (link)

- More directly, researchers are building evidence on preferences for tax redistribution across developing countries
  - Using recent survey methods and information experiments: see the work by Christoper Hoy (link)
  - Process to implement online surveys is described in detail by Stantcheva (2022)

Outline of today's talk

1. Motivation

2. **Relevant data for country-level research**
   A. Tax collection
   B. Tax burdens: Statutory and Effective
   C. Equity
   D. Informality

3. Relevant data for within-country research

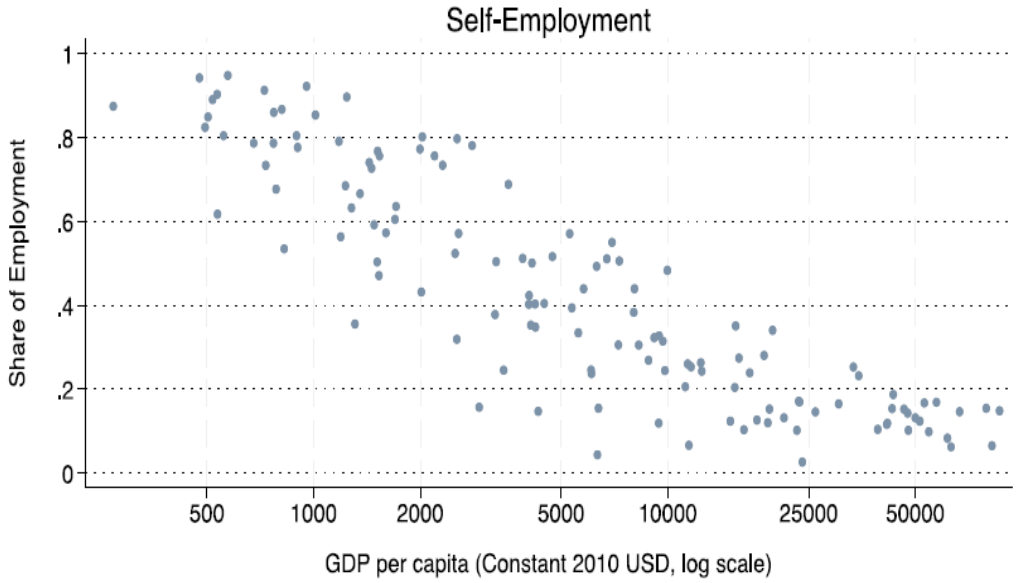
Informality and development

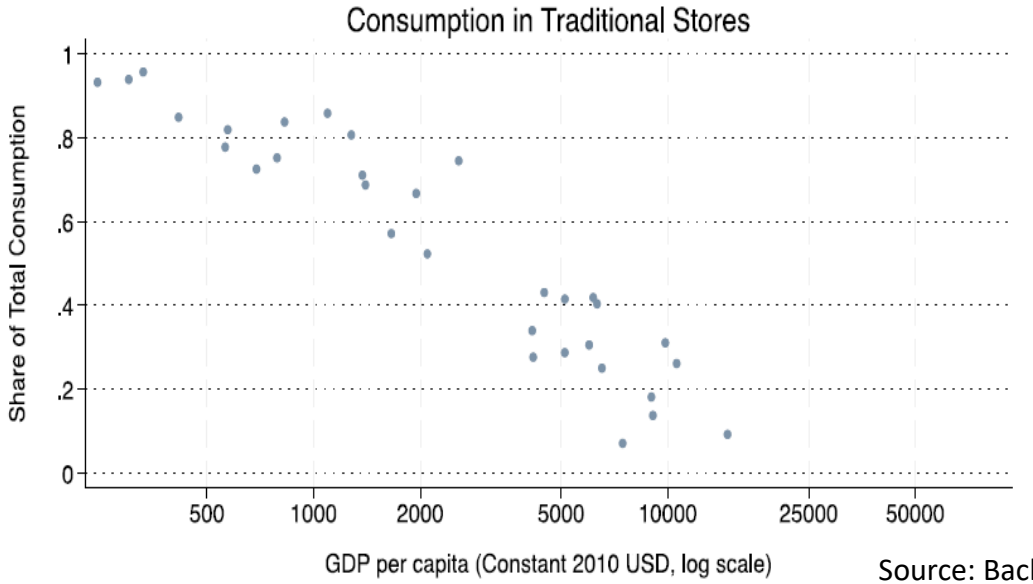
- A key feature in developing countries is that many firms and citizens operate informally outside the tax net
  - In the sense that they pay no or little taxes
  - In principle, this can occur both because of evasion or because they are legally exempt

- Numerous, complementary ways to measure informality within countries and across countries (Shleifer and LaPorta, 2014)
  - Moreover, informality at the firm level and at the worker level overlap only partly (Ulyssea, 2018)

- In cross-country setting, helpful to build proxies for informality
  - Employment: Self-employed share of workforce (Jensen, 2022)
  - Consumption: Consumption share in traditional stores (Bachas et al., 2023)

In developing countries, a lot of activity operates informally outside tax net





Source: Bachas, Gadenne and Jensen (2023)

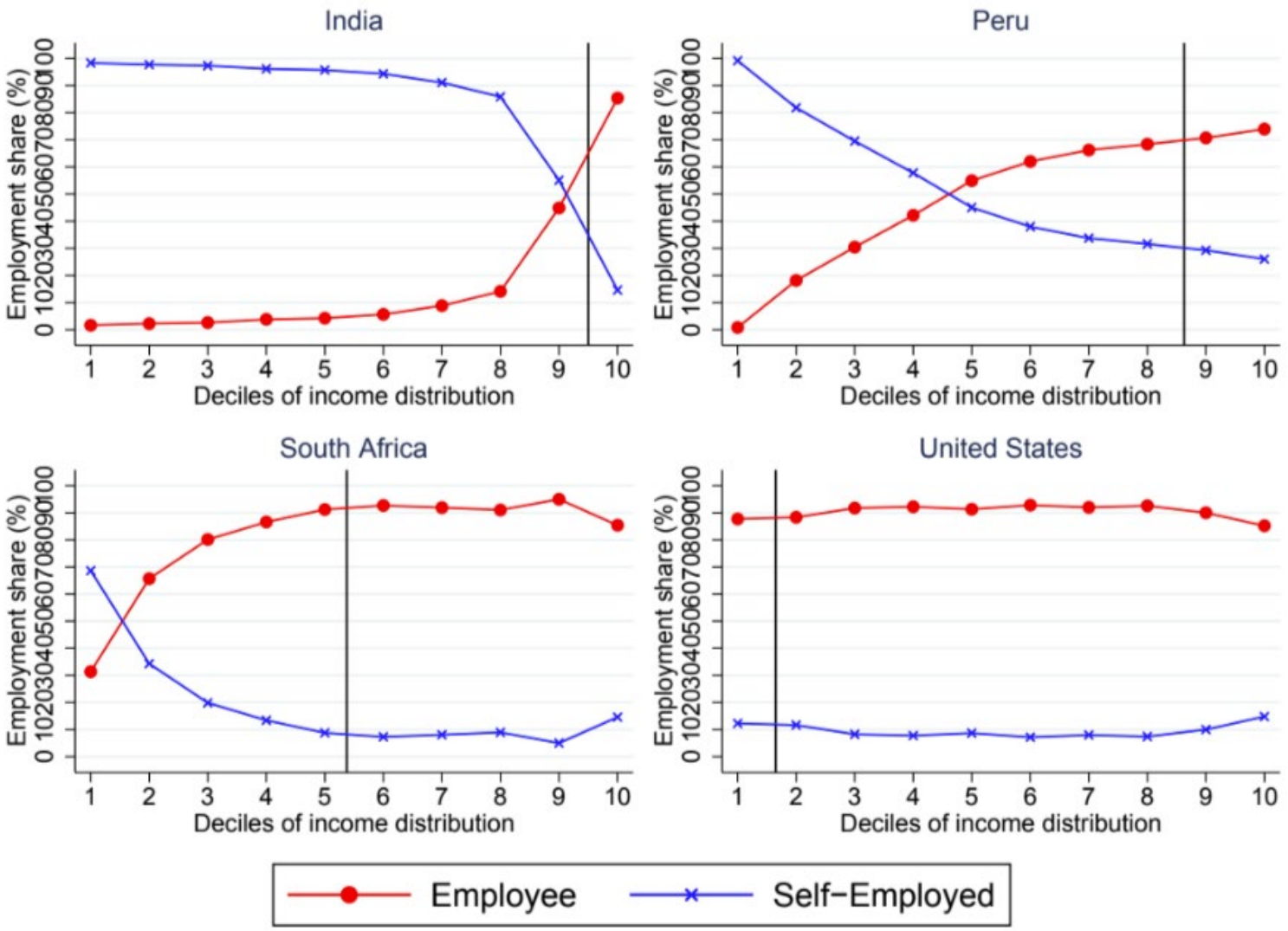
Distributional patterns of informality

- It is possible to leverage harmonized survey data that exists across countries to measure how proxies for informality vary with income both within and across countries

- In the case of employment, we can use household survey data to measure the share of self-employed at various points of the income distribution
  - Available at Luxembourg Income Study (link) and IPUMS (link)

- In the case of consumption, we can use household expenditure diaries to measure how informal consumption varies with household income
  - Data from Bachas et al. (2023) available by request from the World Bank's Microdata Library (link)

Long-run distributional change in employment structure determines PIT design across countries…

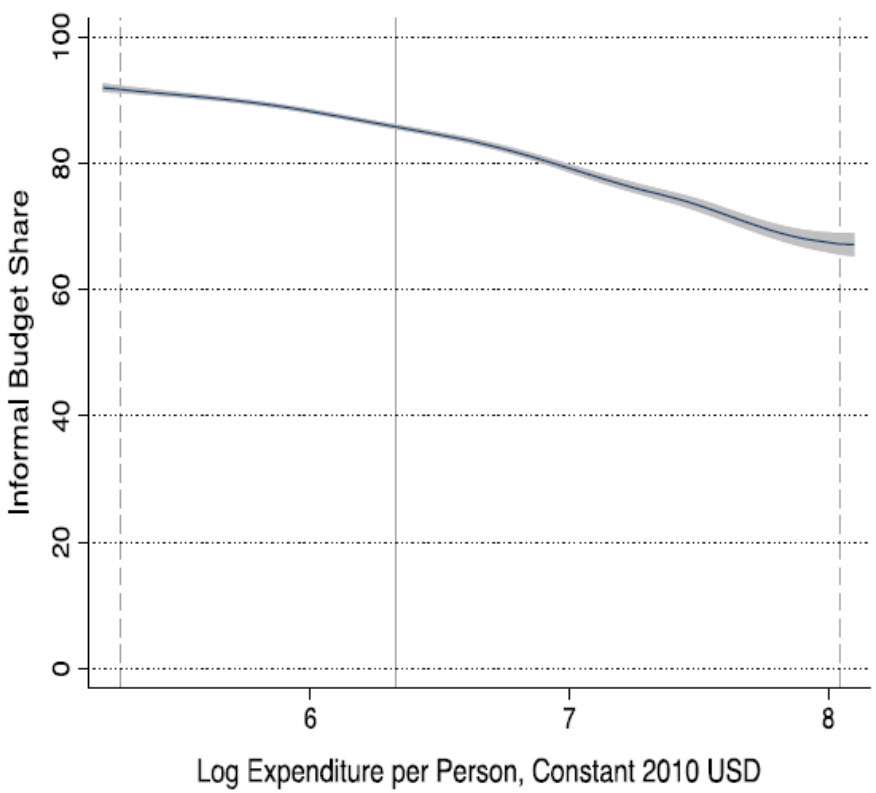


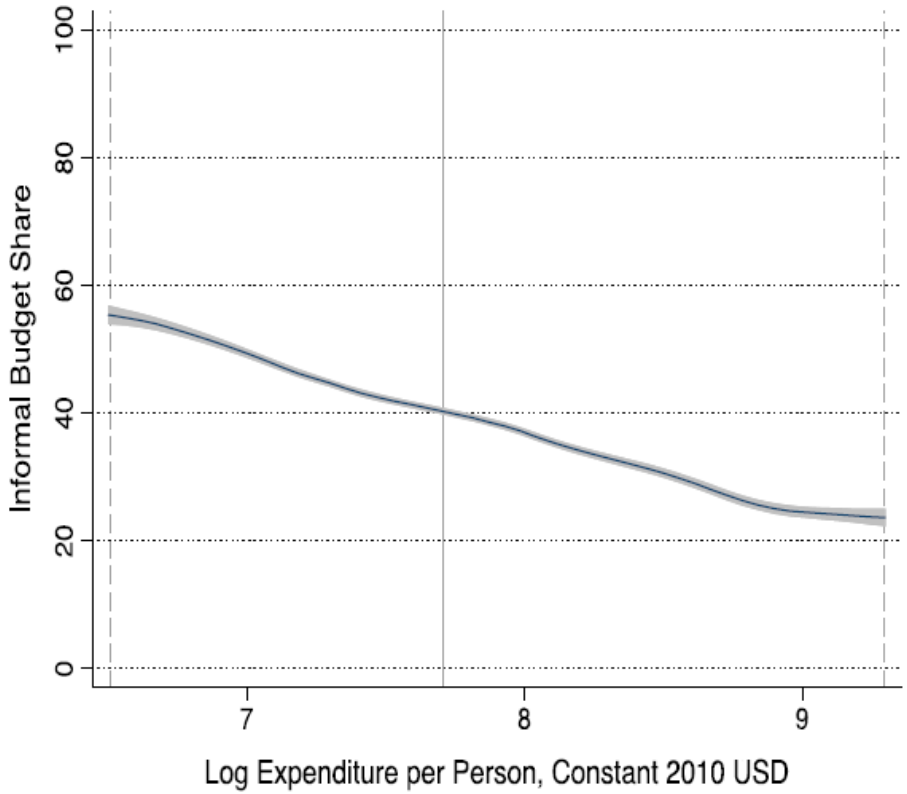
Source: Jensen (2022)

(…) and within country over time…

Source: Jensen (2022)

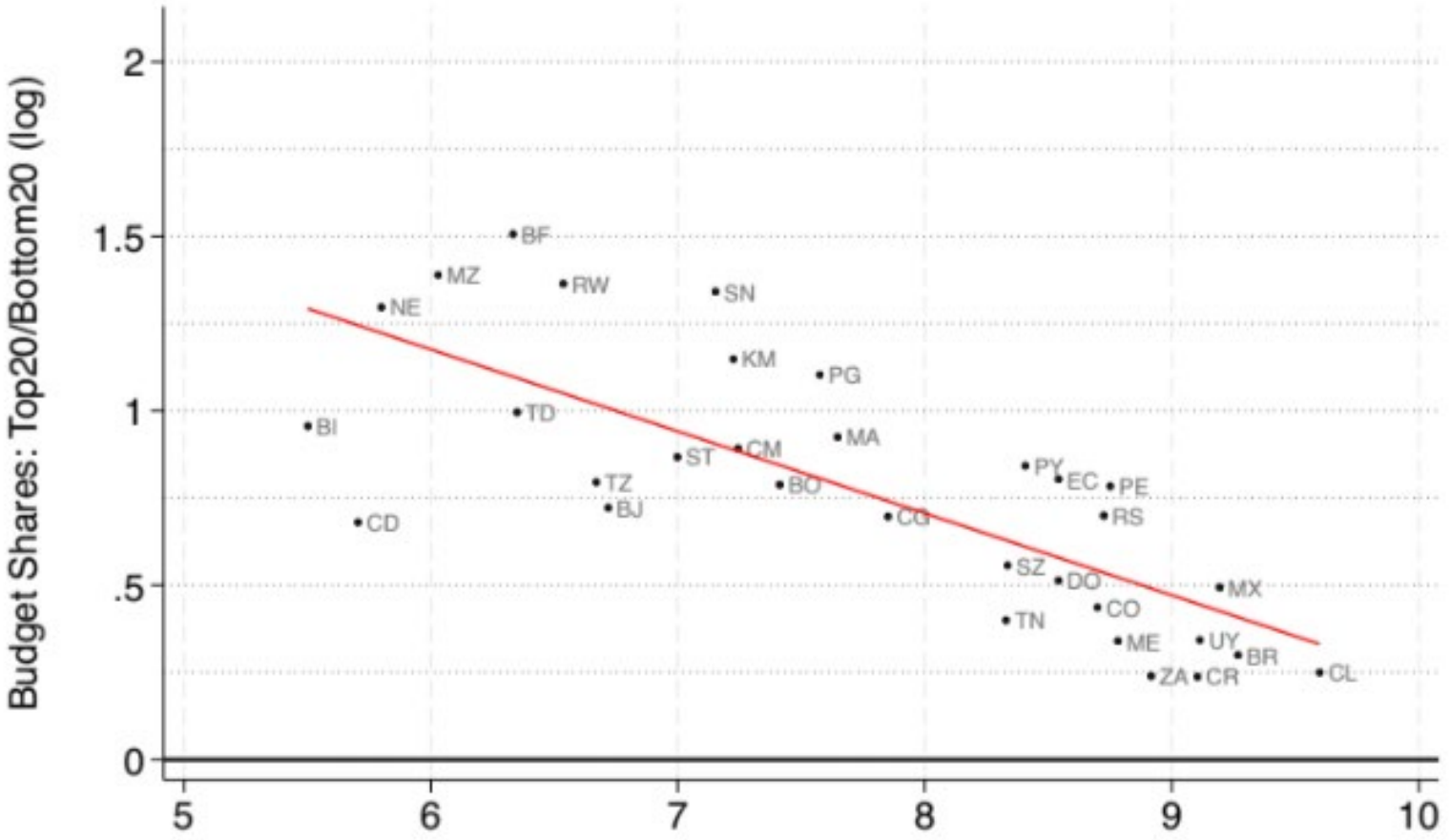
The informality Engel curve is downward sloping in developing countries... [Rwanda on LHS and Mexico on RHS]





Source: Bachas, Gadenne and Jensen (2023)

(…) which makes a uniform consumption tax progressive



Source: Bachas, Gadenne and Jensen (2023)

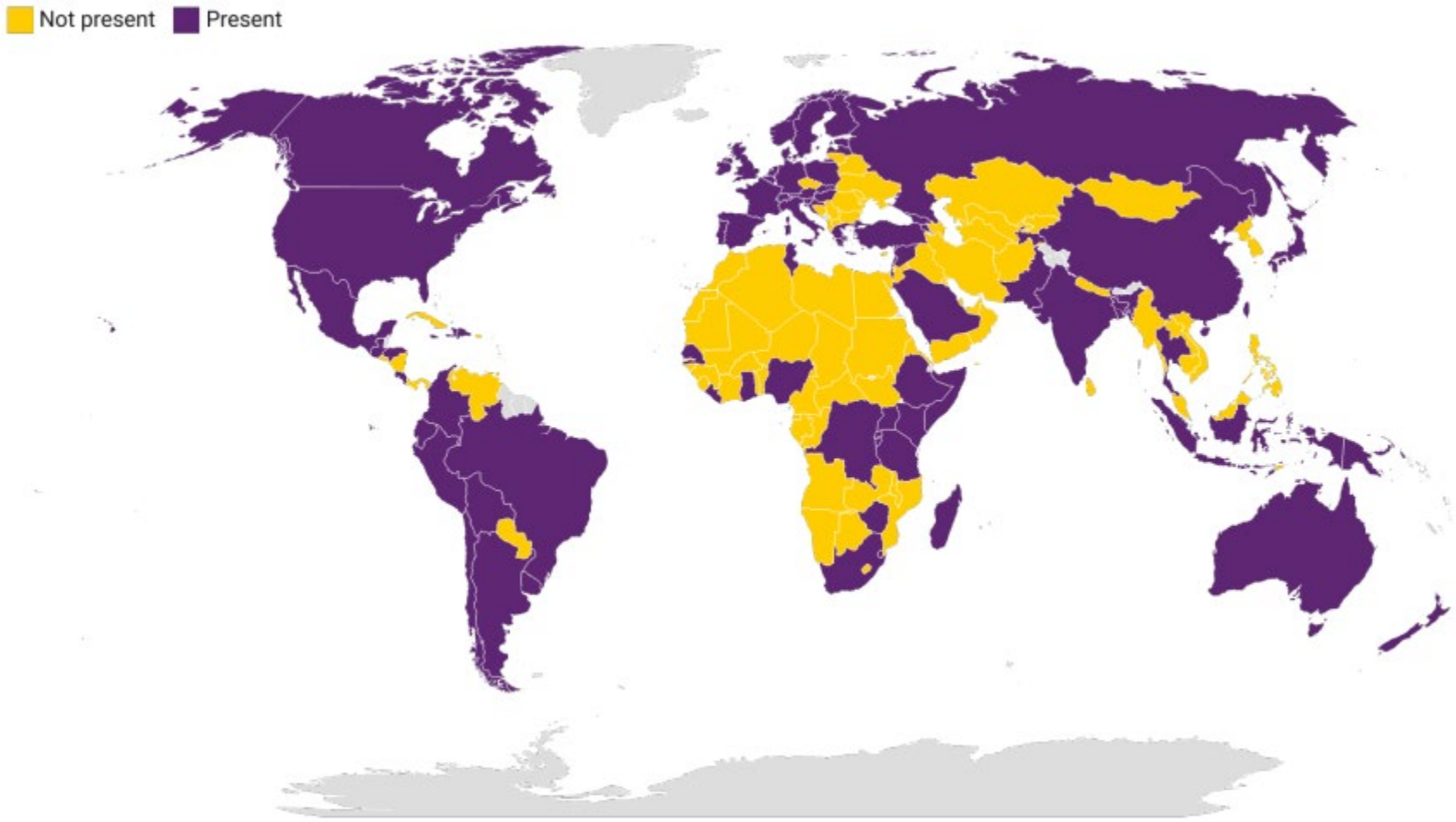
Outline of today's talk

1. Motivation

2. Relevant data for country-level research
   A. Tax collection
   B. Tax burdens: Statutory and Effective
   C. Equity
   D. Informality

3. Relevant data for within-country research

Use of administrative micro-data for research around the world



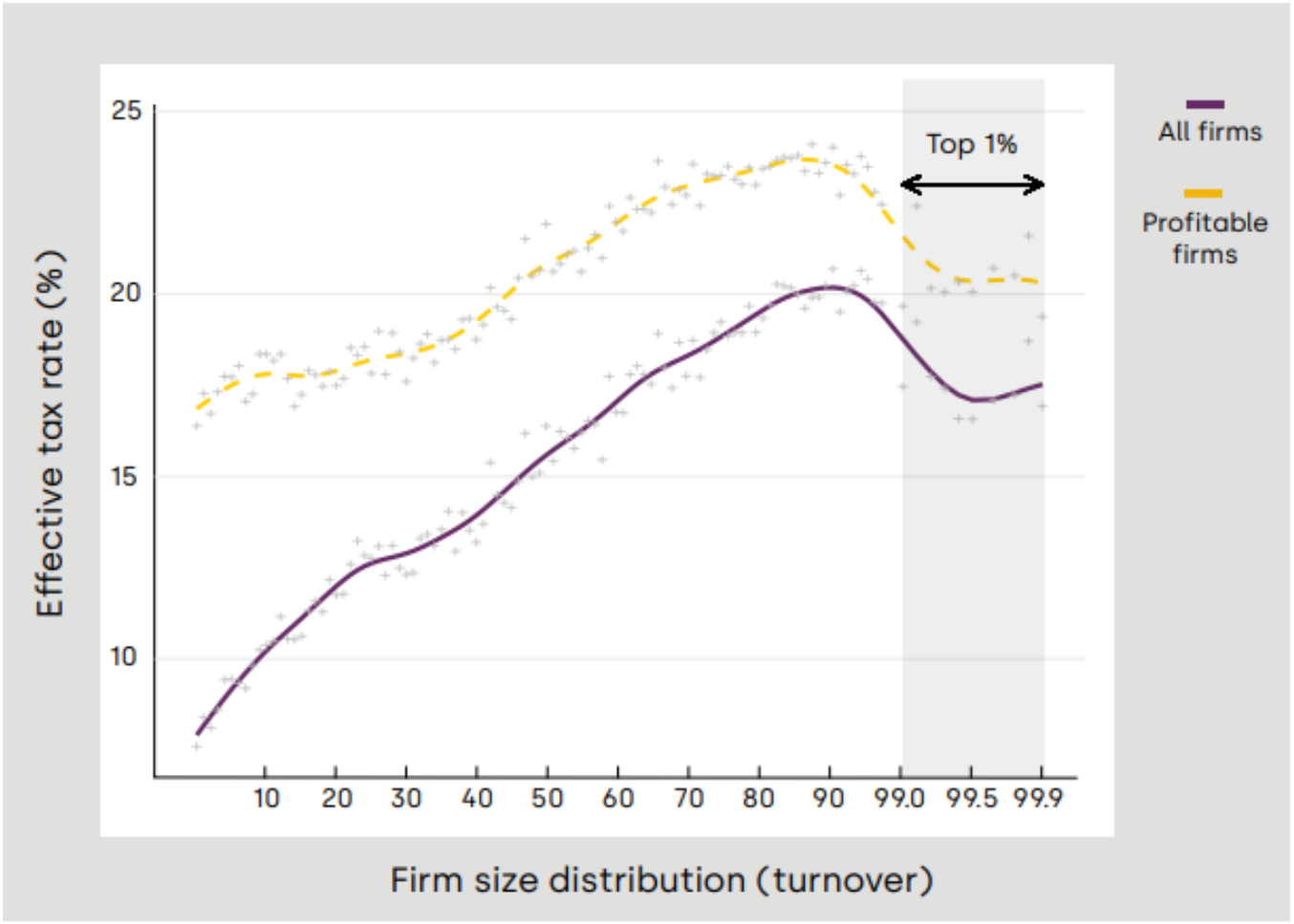
Source: Bachas and Jensen (2023)

The value of administrative data for public finance research

- Microeconomic datasets that governments continuously create to administer tax and transfer programs

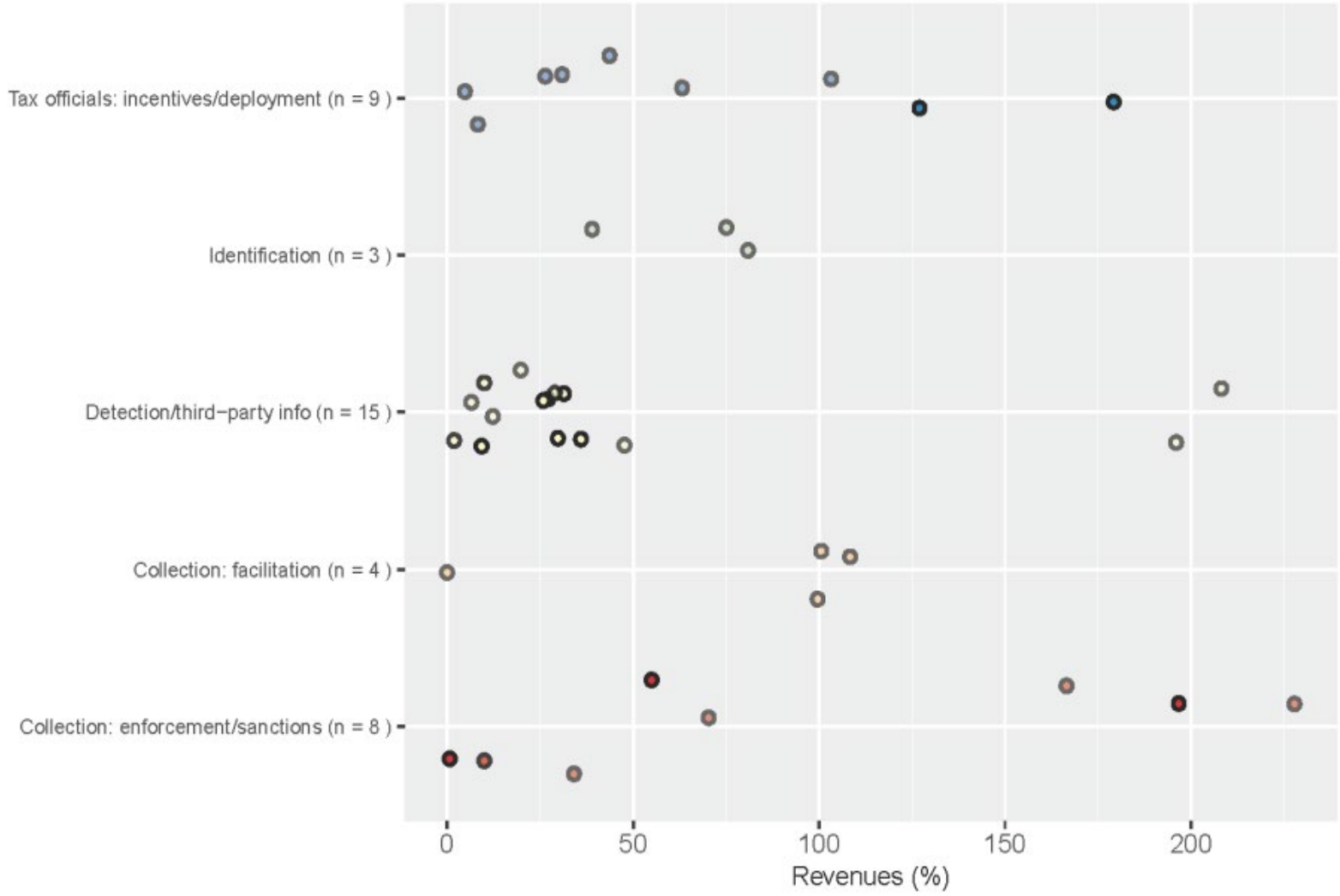
- Administrative data are characterized by three features which make it especially valuable for research:
  1. Administrative data often contains rich information on firms and citizens at the micro-level
  2. Administrative data are collected at high frequency (yearly, or, in the case of firm data, quarterly or monthly), and permit tracking over time
  3. Administrative data has comprehensive coverage of the formal sector: much larger sample size than survey data

Administrative data has informational value to precisely describe firms and citizens



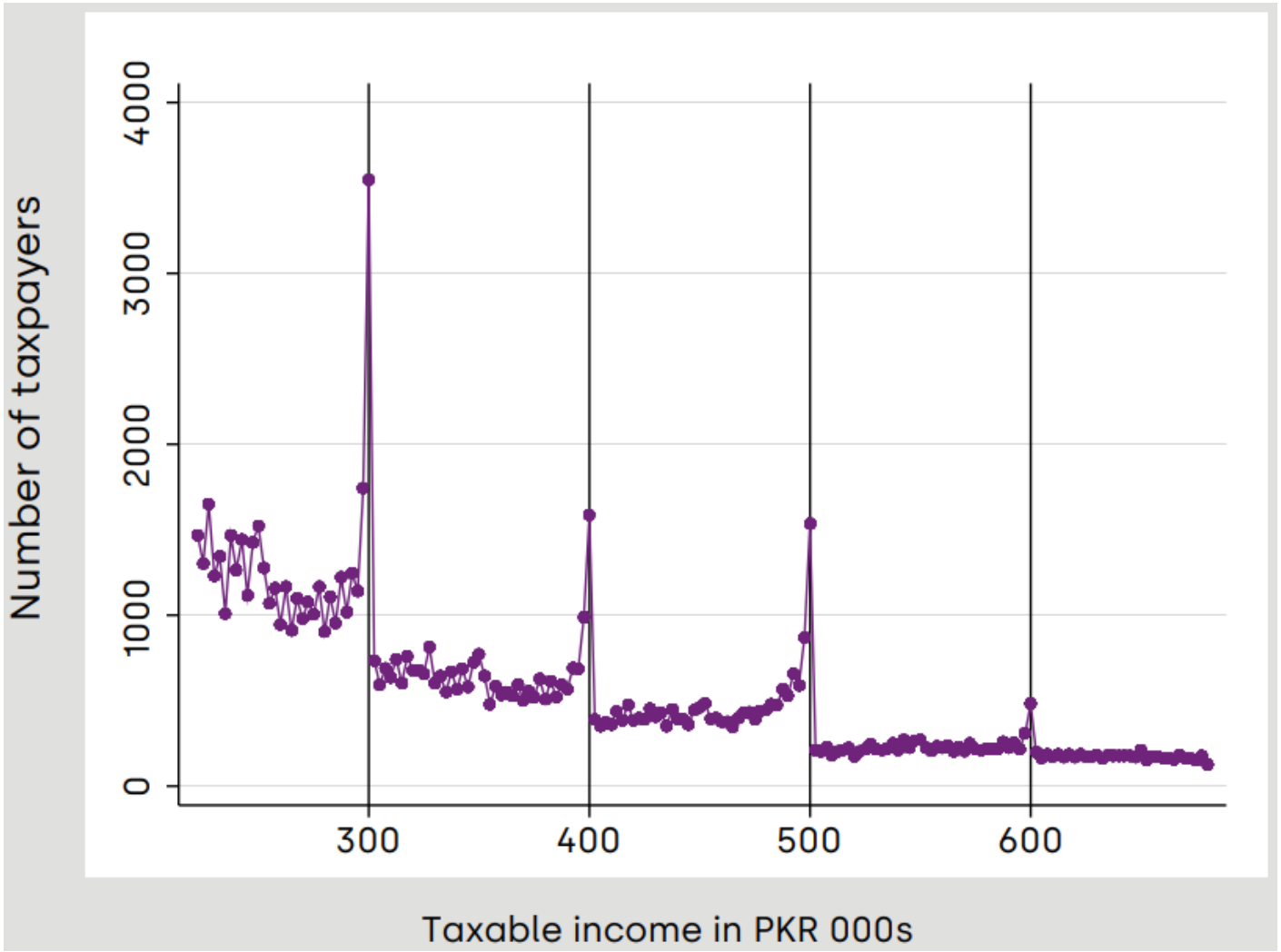
Source: Bachas and Jensen (2023)

Administrative data allows for precise estimation of policy impacts, in combination with experimental methods (…)



Source: Okunogbe and Tourek (2023)

(…) and non-experimental methods



Source: Kleven and Waseem (2013)

Q&A

Thank you!