Online Appendix for

Does the Value-Added Tax Add Value? Lessons Using Administrative Data from a Diverse Set of Countries

Anne Brockmeyer ®Giulia Mascagni ®Vedanth Nair ®Mazhar Waseem ®Miguel Almunia

Journal of Economic Perspectives

Data

Unless otherwise specified, we aggregate monthly and quarterly tax declarations to the annual level. To avoid contamination by "ghost-filers", i.e. nonexistent firms seeking VAT refunds (Mascagni et al., 2020), we exclude firms with zero annual sales. In addition, we made various country-specific decisions in preparing our data, which are laid out in detail in the readme and codes in the data repository. For example, the value-added tax that a firm owes to the government in a specific year is affected by country-specific features, such as the application of interest or fees for late payment, the withholding system, other taxes that are collected alongside the VAT, and un-refunded credits. Table A.1 shows summary statistics for our data.

Overview of Additional Analyses

Figures A.1-A.3 provide additional insights relating to fact 1 on revenue concentration. Figure A.1 shows that the level of the VAT revenue concentration documented in Figure 1 of the main paper and the correlation between VAT revenue concentration and per capita GDP are qualitatively similar when considering the share of VAT contributed by the top 1 percent or top 20 percent of firms. Figure A.2 shows that the concentration of sales is instead flat or even slightly increasing in GDP per capita. Figure A.3 shows that the level of the VAT registration threshold (as share of GDP per capita) is negatively correlated with GDP per capita, and thus unlikely to explain the higher VAT revenue concentration in lower-income countries.

Figures A.4- A.6 provide additional insights relating to fact 2 on effective tax rates. Figure A.4 shows that larger firms have much lower effective tax rates than small firms even in narrowly defined sectors. Figure A.5 shows this for the manufacturing sector and the retail sector individually, and there is again little difference between the two. These results do not

take into account the possibility that firms that operate in sectors that are completely exempt, such as agriculture or financial services, may not file a VAT return. Figure A.6 shows that tax expenditures are a share of total VAT revenue as negatively correlated with GDP per capita.

Figures A.7-A.9 provide additional insights relating to fact 3 on input non-claiming. Figure A.7 provides a breakdown of sales into non-exempt, exempt and reduced or zero-rated sales for each country and firm-size decile, showing that large firms typically have a smaller share of non-exempt sales. Figure A.8 shows that input non-reclaiming is not associated with differences in production technology (and associated differences in the likelihood of making non-taxable purchases) across sectors. Figure A.9 shows that a substantial share of firms in our data have voluntarily registered for the VAT, consistent with the possibility that small firms derive non-tax benefits from remaining in the VAT net.

Figures A.10-A.12 provide additional insights relating to fact 4 on refund delays. Figure A.10 shows that exporters and non-exporters are similarly likely to carry forward credit for more than 12 months. Figure A.11 shows that our measure of the share of firms that are carrying forward credits for more than 12 months is correlated with the World Bank Doing Business measure of refund delays (albeit in a limited sample), which is consistent with our argument that the value-added tax refund system and the credit carryforward system are linked. Figure A.12 shows credits carried forward as a percentage of sales for each country and firm-size percentile. Figure A.13 shows the flat relationship between the statutory VAT rate and GDP per capita.

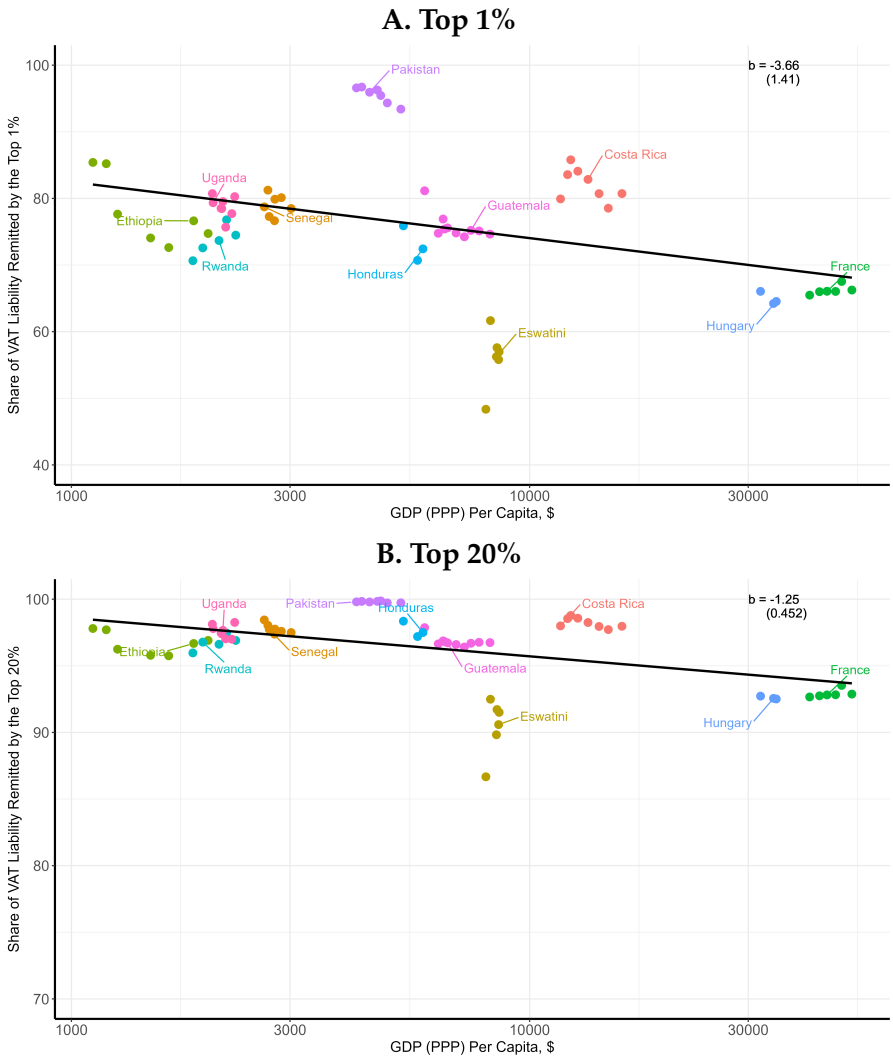
Finally, the last section of this appendix provides details on the counterfactual turnover tax rate calculations and Figure A.14 shows the simulation results.

TABLE A.1: **SUMMARY STATISTICS**

Country	Year Range	Standard VAT Rate, %	GDP Per Capita (PPP)	Number of Firms	Median Sales, $	Sales at the 99th Percentile, $m	Percentage in Retail	Percentage in Manufacturing
Costa Rica	2007-2014	13	15,905	62,227	43,854	7.4	29.1	11.9
Eswatini	2012-2017	14	8,214	2,701	85,936	11.4	29.4	10.7
Ethiopia	2011-2017	15	1,988	28,607	15,910	3.2	34.7	9.3
France	2015-2020	20	50,501	3,378,161	102,649	10.3	10.3	4.4
Guatemala	2006-2015	12	8,194	148,663	24,362	4.7	14.4	4.7
Honduras	2018-2020	15	5,851	27,623	61,106	15.8	21.6	10.6
Hungary	2018-2020	27	34,541	387,184	65,340	7.6	17.1	8.8
Pakistan	2012-2018	17	4,892	70,627	144,668	22.8	6.4	28.5
Rwanda	2016-2020	18	2,284	16,514	30,402	5.2	27.2	10.2
Senegal	2009-2015	18	2,872	10,776	41,370	16.7	28.7	7.9
Uganda	2011-2018	18	2,128	14,562	76,410	14.5	18	8.2

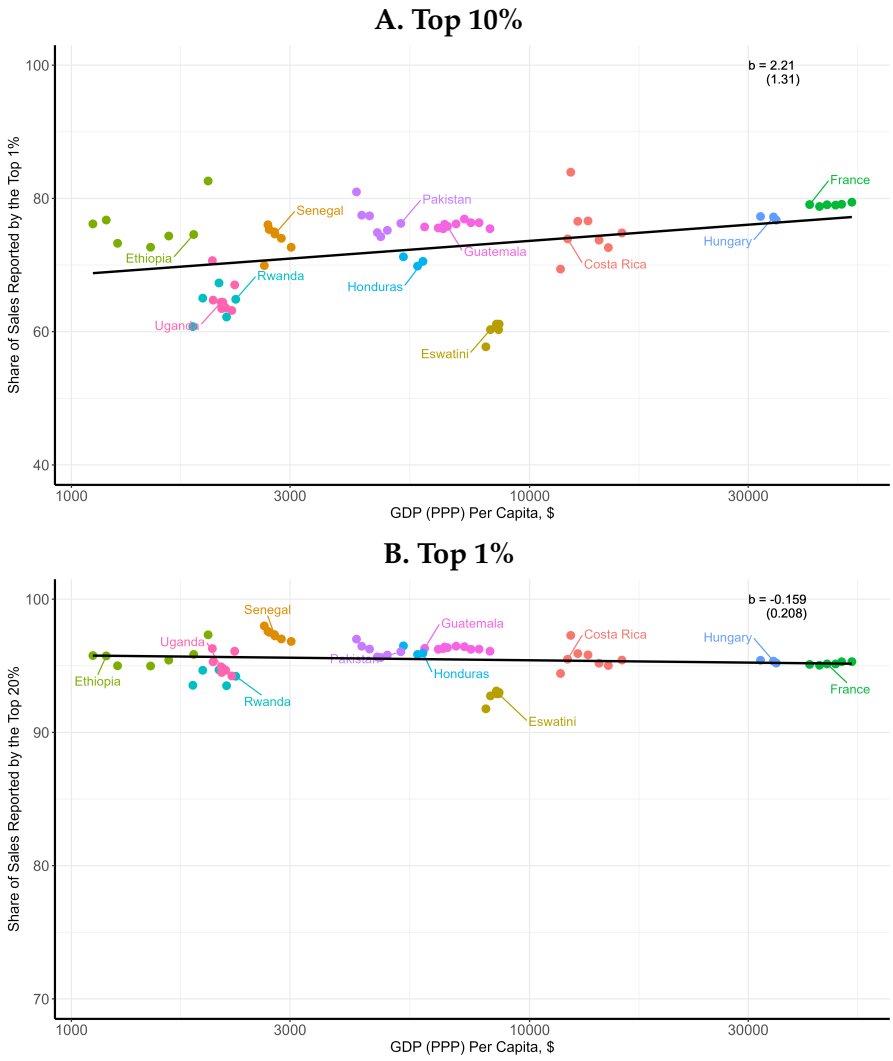
Notes: This table provides summary statistics on the firm-level administrative VAT records used in this paper. Statistics are for the latest year of data available, or 2019 if the latest year of data is 2020. Median sales and sales at the 99th percentile are based on annual sales and are based on the sample of firms that report annual sales greater than zero. Manufacturing is defined as ISIC codes 10-33, and retail is defined as ISIC code 47. Ethiopia and Senegal use different industry labeling conventions; in those cases, manual classification is used to determine whether a firm is in manufacturing, retail, wholesale, services, primary, or other.

FIGURE A.1: **ALTERNATIVE MEASURES OF REVENUE CONCENTRATION**



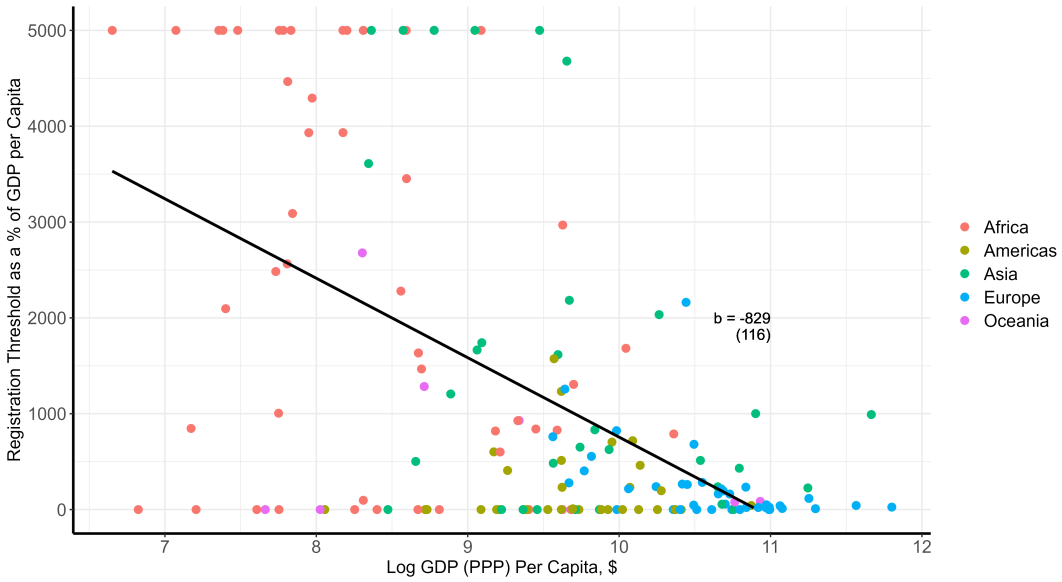
Notes: This figure plots the share of a country's VAT liability that is contributed by the largest 1% (Panel A) and 20% (Panel B) of firms against country log GDP per capita, using the same methodology as Figure 1 in the paper. Each dot represents a country-year observation. Shares are calculated using firm-level data aggregated by calendar year, for firms with positive sales. The VAT liability is defined as max(0,output VAT - input VAT). The largest 1%/20% of firms are those that have the largest tax liabilities. The slope coefficient displayed on the graph is from a simple regression of y on x, with standard errors clustered at the country level. This figure relates to fact 1 in the paper.

FIGURE A.2: **CONCENTRATION OF REPORTED SALES**



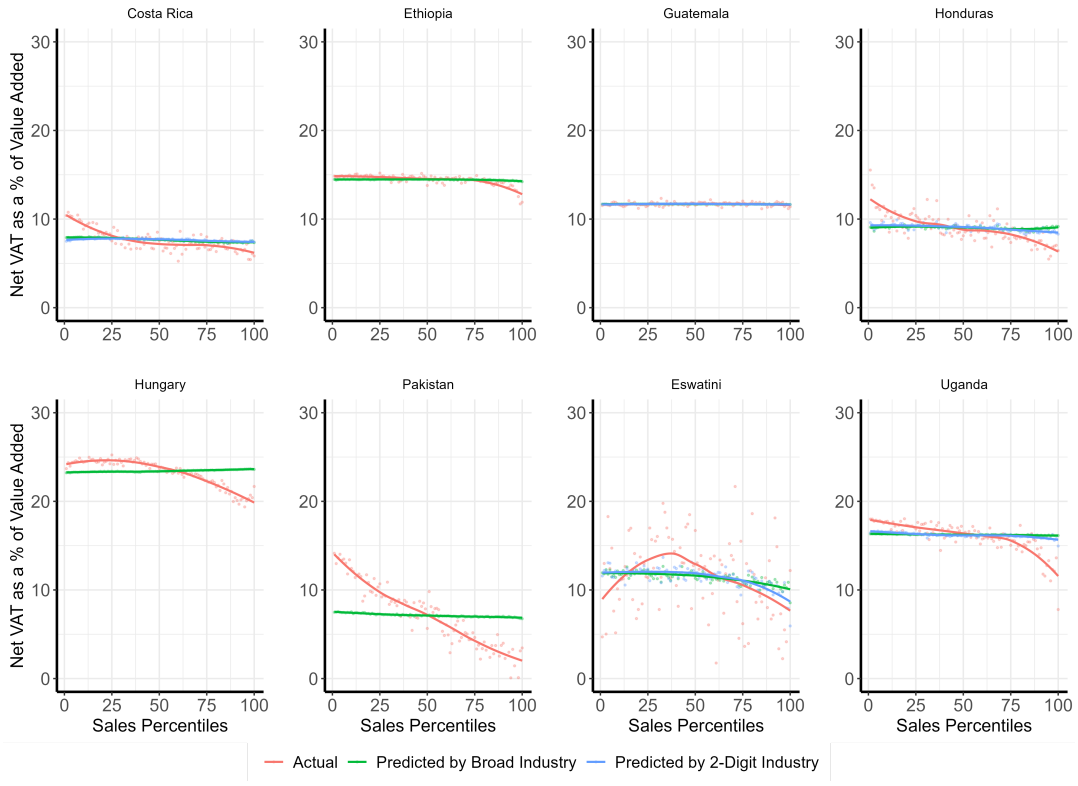
Notes: This figure plots the share of a country's reported sales that is contributed by the largest 10% of firms against country log GDP per capita. Each dot represents a country-year observation. Shares are calculated using firm-level data aggregated by calendar year, for firms with positive sales. The largest 10% of firms are those that have the largest sales. The slope coefficient displayed on the graph is from a simple regression of y on x, with standard errors clustered at the country level. This figure relates to fact 1 in the paper.

FIGURE A.3: **VAT REGISTRATION THRESHOLDS AS A PERCENTAGE OF GDP PER CAPITA**



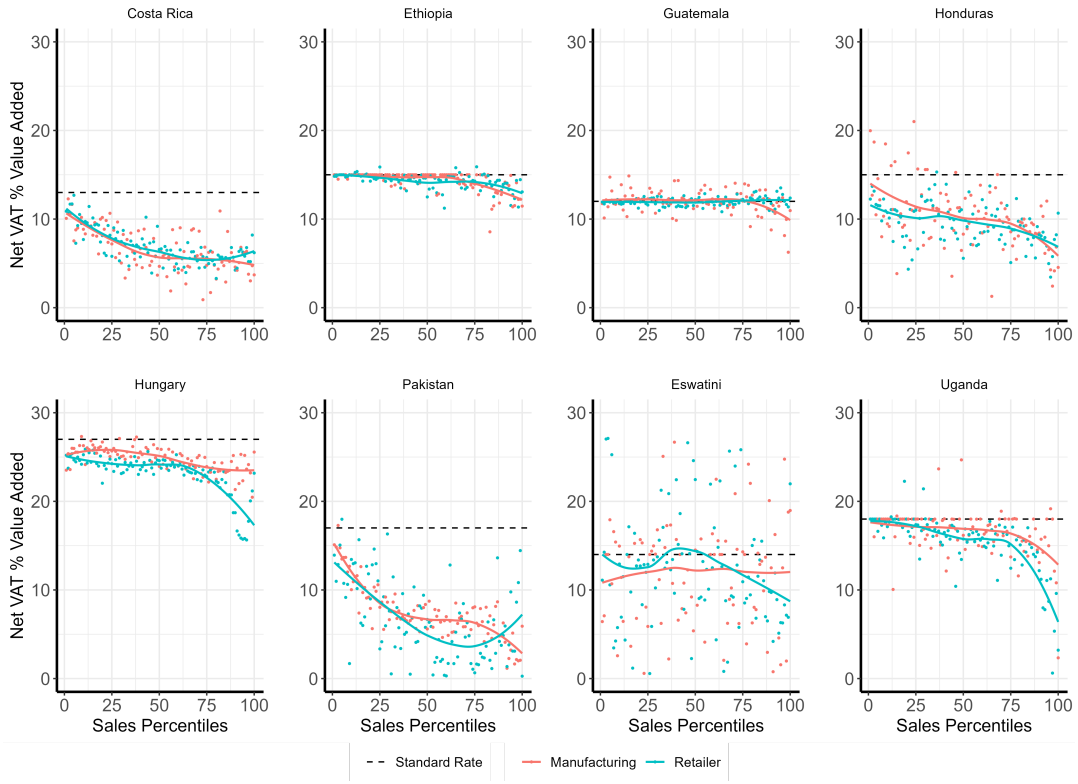
Notes: This figure plots the VAT registration threshold, as a percentage of GDP per capita (USD), against log GDP per capita (PPP). Each dot represents a country, for the latest year of data available. The VAT registration threshold data is from the IMF Tax Policy Assessment Framework `https://www.imf.org/external/np/fad/tpaf/pages/vat.htm`, and is cross-checked with the EY Global GST Guide 2022 `https://www.ey.com/en_gl/tax-guides/worldwide-vat-gst-and-sales-tax-guide` and tax authority websites. Registration thresholds are divided by GDP per capita to get a sense of how binding the registration threshold is for the average firm, under the assumption that average firm size tends to correlate with GDP per capita (consistent cross-country data on average firm size is not available to our knowledge). Registration thresholds in USD are generally uncorrelated with GDP per capita. The slope coefficient displayed on the graph is from a simple regression of y on x. This figure relates to fact 1 in the paper.

FIGURE A.4: **EFFECTIVE TAX RATES WITHIN SECTORS**



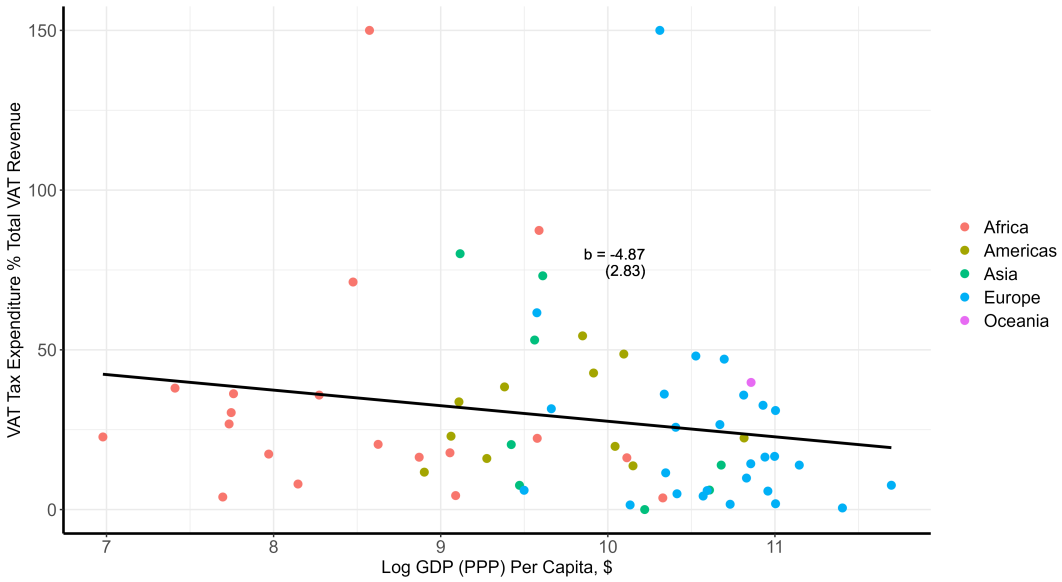
Notes: This figure shows effective tax rates (ETR) by firm size (total sales) percentiles. The methodology for calculating effective tax rates is the same as in Figure 2 of the paper. The ETR is defined as annual net VAT over annual value-added, where net VAT = (output VAT - input VAT), and value added = (total reported sales - total reported purchases). Exporters (those with exports > 30% of annual sales) are excluded. ETRs are winsorized at the standard rate (or at the higher rate of VAT in Honduras). The red line plots actual ETRs. The blue and green lines plot ETRs that are predicted by a regression of ETRs on industry fixed effects. Concretely, we estimate for each country the regression $ETR_{itc} = \alpha_c + \beta_c \text{Industry}_{ic} + \gamma_c \text{Year}_{tc}$, for firm i in year t and country c. We then predict $\hat{ETR}_{itc}$ for each firm size percentile based on the estimated coefficients and the industry composition of firms in the size percentile. In the blue series, 'Industry' is a set of fixed effects for broad industry groupings (primary, manufacturing, wholesale / retail, services, other). In the green series, 'Industry' is a set of fixed effects for 2-digit ISIC codes (this is only available in some countries). The goal of this analysis is to identify how much variation in ETRs across firm size percentiles is attributable to differences in ETRs across sectors. This figure relates to fact 2 in the paper.

FIGURE A.5: **EFFECTIVE TAX RATES IN THE MANUFACTURING AND RETAIL SECTOR**



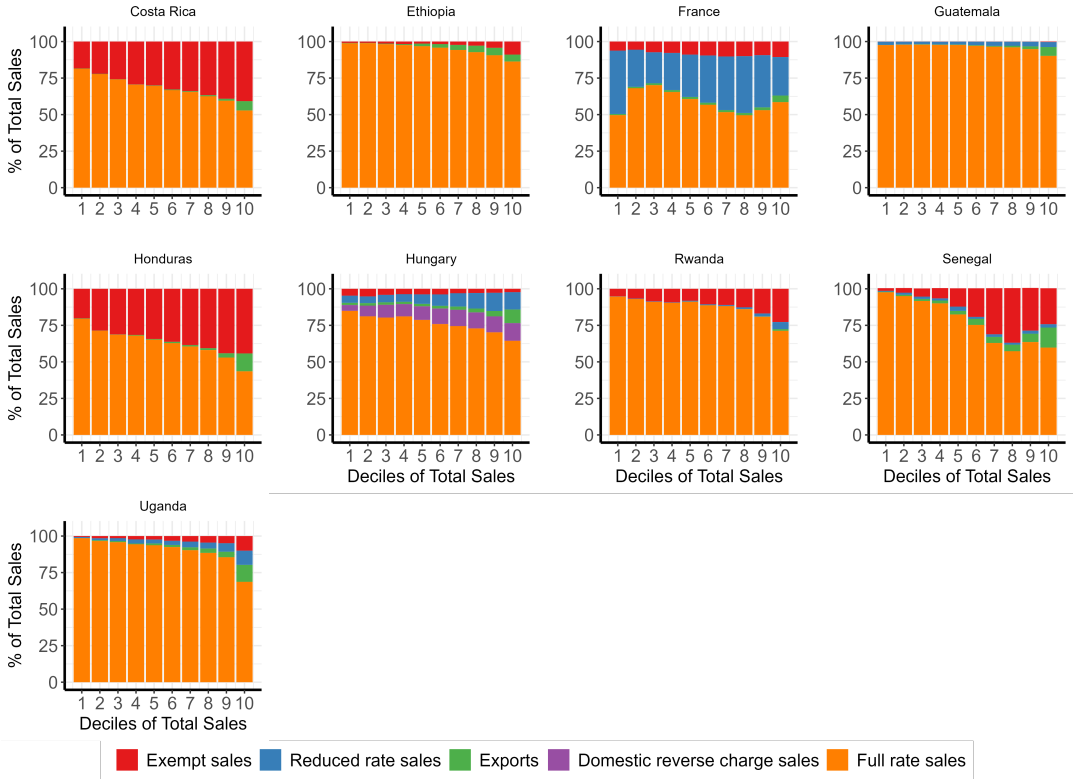
Notes: This figure shows effective tax rates (ETR) by firm size (total sales) percentiles, for firms in the manufacturing sector and the retail sector specifically. Manufacturing is defined as ISIC codes 10-33, and retail is defined as ISIC code 47. Ethiopia and Senegal use different industry labeling conventions; in those cases, manual classification is used to determine whether a firm is in manufacturing, retail, wholesale, services, primary, or other. The methodology for calculating effective tax rates is the same as in Figure 2 of the paper. The ETR is defined as annual net VAT over annual value-added, where net VAT = (output VAT - input VAT), and value added = (total reported sales - total reported purchases). Exporters (those with exports > 30% of annual sales) are excluded. ETRs are winsorized at the standard rate (or at the higher rate of VAT in Honduras). The red line plots actual ETRs. The blue and green lines plot ETRs that are predicted by a regression of ETRs on industry fixed effects. Concretely, we estimate for each country the regression $ETR_{itc} = \alpha_c + \beta_c \text{Industry}_{ic} + \gamma_c \text{Year}_{tc}$, for firm i in year t and country c. We then predict $\hat{ETR}_{itc}$ for each firm size percentile based on the estimated coefficients and the industry composition of firms in the size percentile. In the blue series, 'Industry' is a set of fixed effects for broad industry groupings (primary, manufacturing, wholesale / retail, services, other). In the green series, 'Industry' is a set of fixed effects for 2-digit ISIC codes (this is only available in some countries). The goal of this analysis is to identify how much variation in ETRs across firm size percentiles is attributable to differences in ETRs across sectors. This figure relates to fact 2 in the paper.

FIGURE A.6: **VAT TAX EXPENDITURES AS A PERCENTAGE OF VAT REVENUES**



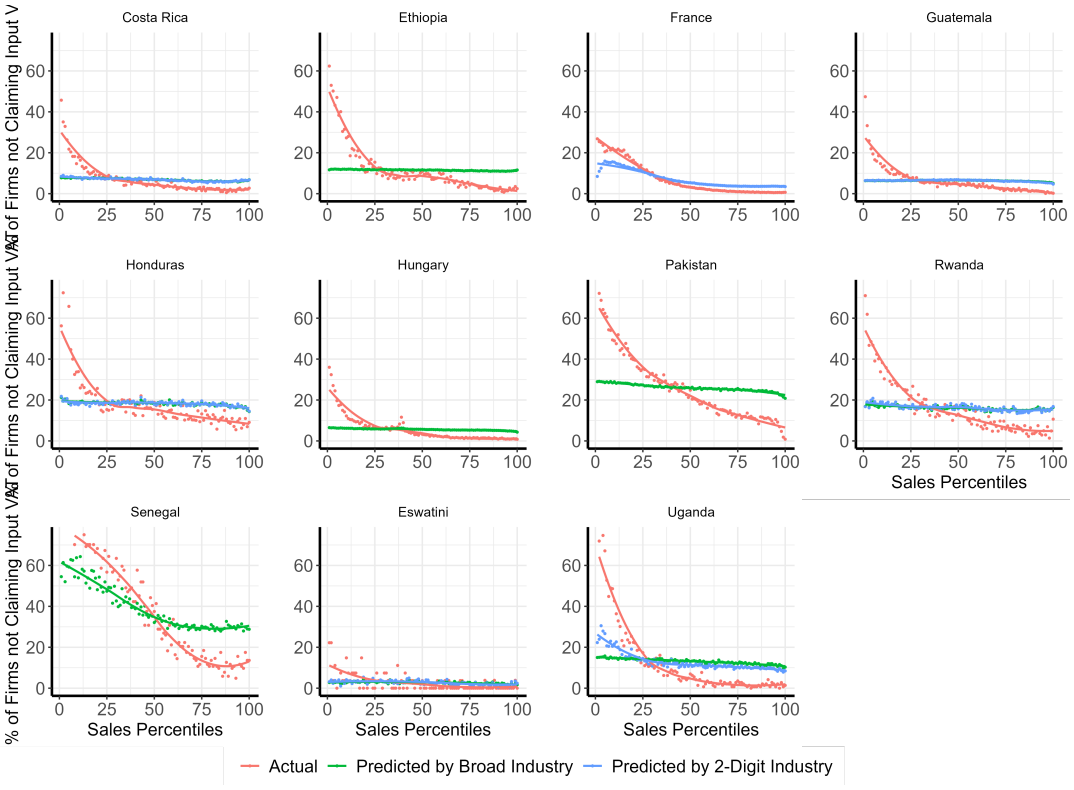
Notes: This figure plots VAT tax expenditures share of a country's VAT revenue against country log GDP per capita. The Tax expenditure data comes from the Global Tax Expenditure Database (Redonda, von Haldenwang, Christian and Aliu, 2022). VAT tax expenditures are the revenue foregone from policies that lead to deviations of VAT from the 'benchmark tax system'. The definition of the benchmark tax system varies across countries, but for VAT, the largest source of tax expenditures are reduced/zero rates for domestic sales, and exemptions. Note that tax expenditures do not include behavioral responses. VAT revenue data is from UNU-WIDER Government Revenue Dataset (UNU-WIDER, 2022). The slope coefficient displayed on the graph is from a simple regression of y on x, with standard errors clustered at the country level. This figure relates to fact 2 in the paper.

FIGURE A.7: **BREAKDOWN OF SALES BY CATEGORY**



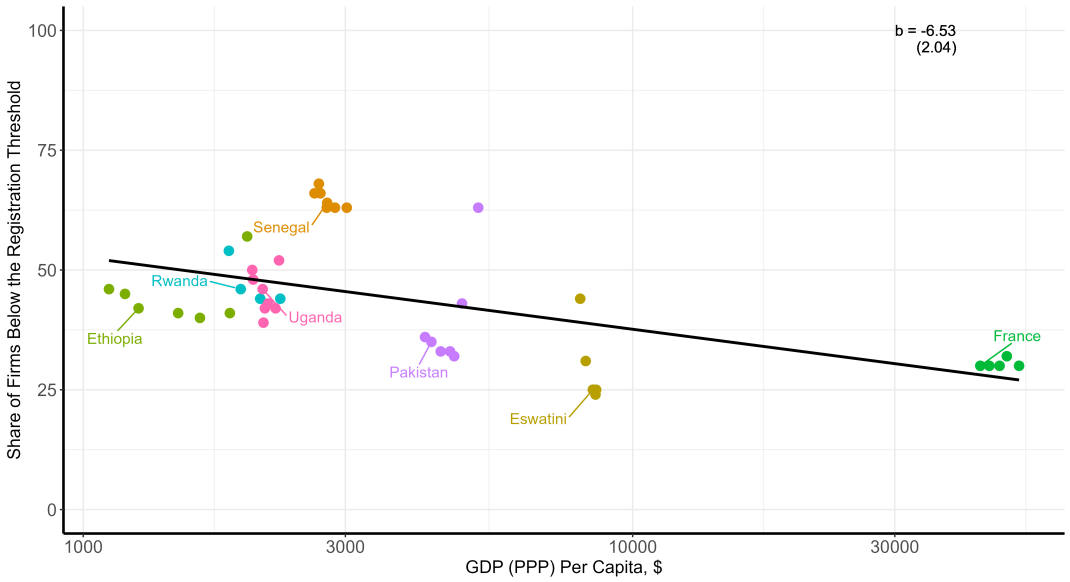
Notes: This figure plots the share of (domestic) sales that are at the full rate of VAT, at a reduced or zero rate of VAT, and are exempt, for firms which have positive domestic sales. Firms are categorized into deciles based on annual sales. The statistic for each decile is the average share of full rate sales, reduced/zero rate sales, and exempt sales, for firms within that decile.

FIGURE A.8: **INPUT VAT NON-CLAIMING WITHIN SECTORS**



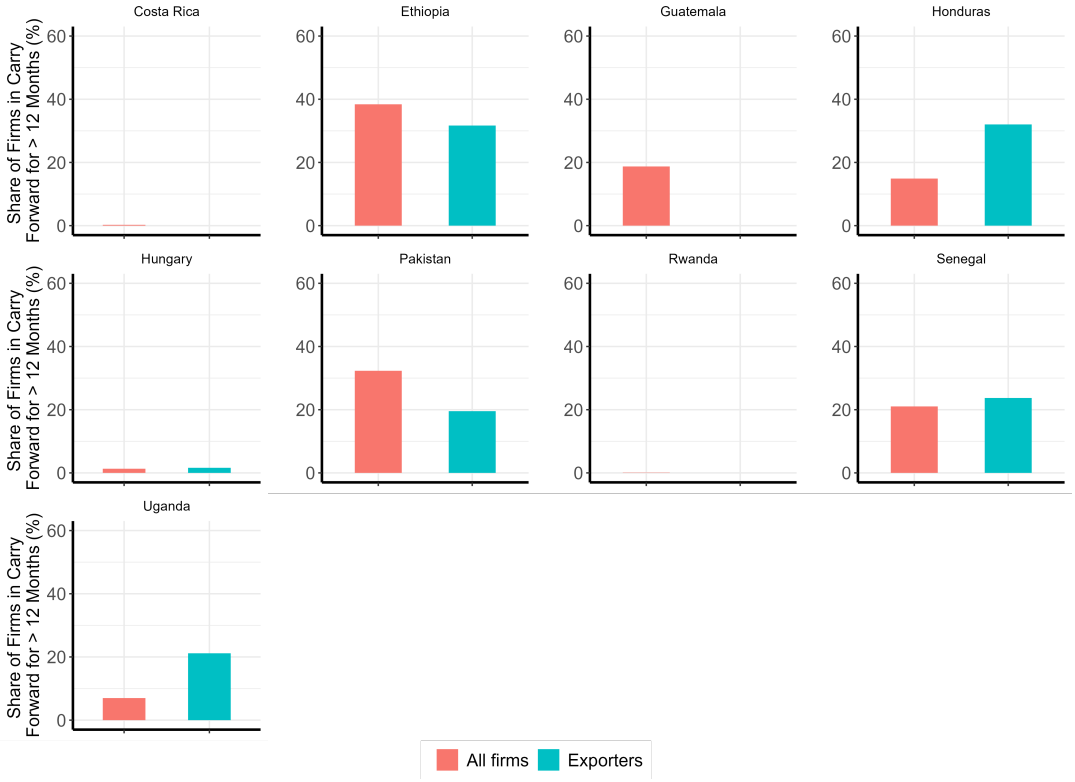
Notes: This figure shows the share of firms not claiming input VAT by firm size (total sales) percentiles. The figure uses the same methodology as Figure 3 in the paper. A firm has zero input VAT if they have declared zero input VAT over the entire year, conditional on reporting positive sales. The red line plots the actual share of input VAT non-claiming. The blue and green lines plot input VAT non-claiming predicted by a regression of non-claiming on industry fixed effects. Concretely, we estimate for each country the regression $ZeroIV_{itc} = \alpha_c + \beta_c \text{Industry}_{ic} + \gamma_c \text{Year}_{tc}$, where $ZeroIV_{itc}$ is a dummy that takes value 1 if firm i in year t and country c did not claim input VAT. We then predict $\hat{ZeroIV}_{itc}$ for each firm size percentile based on the estimated coefficients and the industry composition of firms in the size percentile. In the blue series, 'Industry' is a set of fixed effects for broad industry groupings (primary, manufacturing, wholesale / retail, services, other). In the green series, 'Industry' is a set of fixed effects for 2-digit ISIC codes (this is only available in some countries). The goal of this analysis is to identify how much variation in input VAT non-claiming across firm size percentiles is attributable to differences across sectors. This figure relates to fact 3 in the paper.

FIGURE A.9: **VOLUNTARY REGISTRATION**



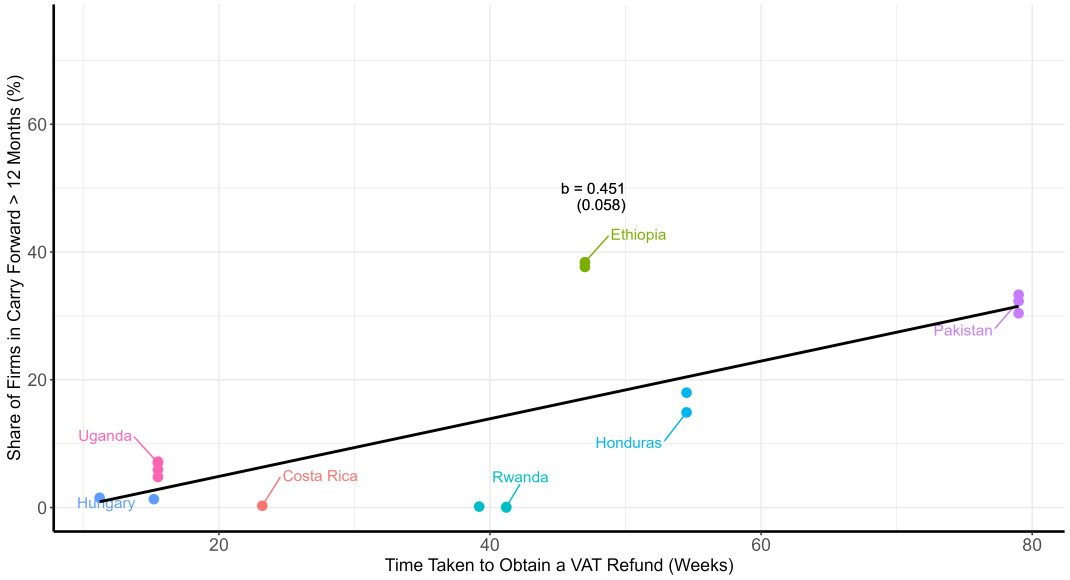
Notes: This figure shows the share of VAT-filing firms that have annual sales below the VAT registration threshold. Each dot represents a country-year pair. Firms with zero annual sales are excluded, due to the issue of nil/ghost-filing. This figure may not exactly equal true 'voluntary registration', as tax administrations may require firms with sales below the threshold to register for the VAT if a certain proportion of their sales are to VAT firms, or if they operate in certain sectors. A firm may also have had sales above the registration threshold in previous years. Note that Costa Rica, Honduras and Guatemala do not have a VAT registration threshold. This figure relates to fact 3 in the paper.

FIGURE A.10: **SHARE OF FIRMS THAT ARE CARRYING FORWARD CREDITS FOR > 12 MONTHS, EXPORTERS AND NON-EXPORTERS**



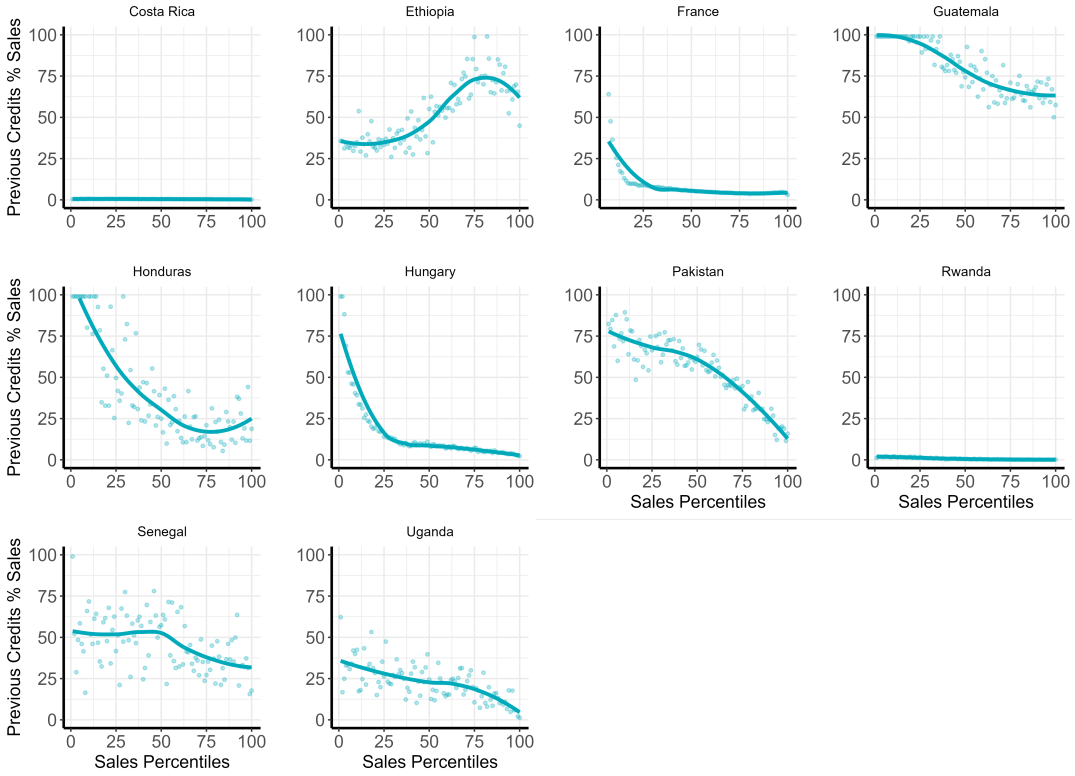
Notes: This figure examines whether exporters are more likely to be exposed to liquidity constraints from slow refunds, compared to non-exporting firms. This is measured using the share of firms receiving credits from a previous period, for at least 12 consecutive VAT filings. The definitions are the same as in Figure 4 in the paper. Exporters are defined as firms where annual exports are >30% of total annual sales. Rwanda and Costa Rica have a marginal number of firms in credit carry forward positions. Results are not available for eSwatini or France due to data limitations. This figure relates to fact 4 in the paper.

FIGURE A.11: **VAT REFUND DELAYS AGAINST SHARE OF FIRMS CARRYING FORWARD CREDITS FOR > 12 MONTHS**



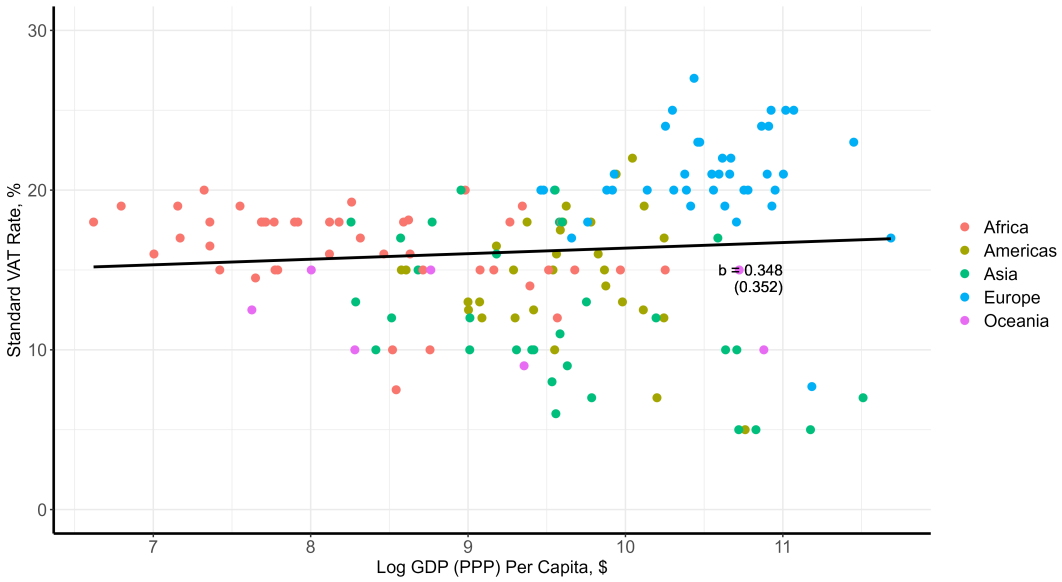
Notes: This figure shows the correlation between the time taken to obtain a VAT refund (in weeks), from the World Bank Doing Business survey, and the share of firms carrying forward credits for > 12 months, from the microdata. The definitions are the same as in Figure 4. Each dot represents a country-year pair. Note that for Costa Rica, the share of firms carrying forward credits is for 2015, but the time taken to obtain a VAT refund is for 2016 (as it is the earliest year available). The time taken to obtain a VAT refund is not available for Guatemala. Results are not available for eSwatini or France due to data limitations. This figure relates to fact 4 in the paper.

FIGURE A.12: **CREDITS CARRIED FORWARD AS A PERCENTAGE OF SALES**



Notes: This figure shows the accumulated amount of credits carried forward over annual sales, by firm size percentiles, for the latest year of data available in each country. This is a replica of Figure 4, Panel C, in the paper. Note that unlike in other graphs, the percentiles of sales for the x-axis here are constructed by taking percentiles of average sales across the entire sample, due to concerns about measurement error in sales affecting both the x and y variable. Credits carried forward as a % of sales is winsorized at 100%.

FIGURE A.13: **STANDARD VAT RATES ACROSS COUNTRIES**



Notes: This figure shows the correlation between standard VAT rates and (log) GDP per capita (PPP).
Source: KPMG indirect tax database (https://kpmg.com/it/it/home/services/tax/tax-tools-and-resources/tax-rates-online/indirect-tax-rates-table.html), manually removing countries that do no have a VAT.

Counterfactual Turnover Tax Rate Calculations

This section describes the details of the counterfactual turnover tax (TOT) rate calculations discussed in the paper and shown in Figure A.14.

The revenue-neutral turnover tax (TOT) rate (domestic), τ_{tc}^{totd}, in year t and country c is calculated by solving the equation: $\sum_i \text{Net VAT}_{itc} = \tau_{tc}^{totd} \sum_i \text{Domestic Taxable Sales}_{itc}$. The left-hand side is VAT revenue (aggregate VAT liabilities). The right-hand side is the TOT rate, multiplied by the tax base for the TOT (domestic sales).

The revenue-neutral TOT rate (domestic and imports) requires information on the VAT collected on imports. We use the share of total VAT revenue that is collected at the import stage, γ_c^{IS}, which is typically available from tax administration annual reports. Notice the lack of t subscript, as this number is typically only available in one or two years of the sample period, so we assume that the share of import VAT in total VAT revenue is constant within countries over time. The import share of VAT varies from 40% to 70% in our sample.

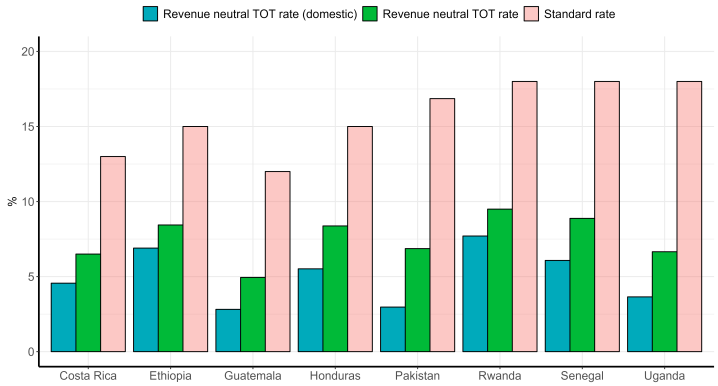
The revenue-neutral TOT rate, τ_{tc}^{tot}, is calculated by solving the equation
$\frac{1}{1-\gamma_c^{IS}} \sum_i \text{Net VAT}_{itc} = \tau_{tc}^{tot}[\sum_i \text{Domestic Taxable Sales} + \frac{\gamma_c^{IS}}{\tau_{tc}^{VAT}(1-\gamma_c^{IS})} \sum_i \text{Net VAT}_{itc}]$,
where τ_{tc}^{VAT} is the standard VAT rate. The term on the left-hand side is approximated total (domestic and import) VAT revenue. The term $\frac{\gamma_c^{IS}}{\tau_{tc}^{VAT}(1-\gamma_c^{IS})} \sum_i \text{Net VAT}_{itc}$ is estimated revenue from VAT on imports, divided by the standard rate of VAT, to proxy total imports as TOT base.

In Figure A.14, we display the average of τ_{tc}^{totd} and τ_{tc}^{tot} over the sample period t for each country.

FIGURE A.14: **COMPARING THE VAT TO ALTERNATIVE TAXES**

A. Counterfactual Revenue from a Retail Sales Tax, Pakistan

B. Counterfactual Revenue-Neutral Turnover Tax Rate



Notes: This figure shows results for counterfactual simulations comparing the VAT to alternative taxes. Panel A compares VAT revenue to the counterfactual revenue collected by a retail sales tax imposed on business-to-consumer (B2C) sales at the standard VAT rate. B2C sales are identified using transaction-level data. Findings are currently only available for Pakistan, and forthcoming for Uganda and Rwanda. Panel B displays the turnover tax (TOT) rates that would be required to collect the same amount of revenue as the VAT. For the revenue-neutral TOT (domestic), the counterfactual involves replacing domestic VAT with TOT, but leaving VAT collected on imports untouched. For the revenue-neutral TOT, the counterfactual involves replacing VAT, collected both domestically and on imports, with a TOT.

References

Redonda, Agustin, von Haldenwang, Christian, and Flurim Aliu. 2022. "Global Tax Expenditures Database (GTED)."

UNU-WIDER. 2022. "UNU-WIDER Government Revenue Dataset – Version 2022."