Economic Inquiry

DOES A VALUE-ADDED TAX INCREASE ECONOMIC EFFICIENCY?

BIBEK ADHIKARI*

Theory predicts that a value-added tax (VAT) is an efficient tax system, which is one of the primary reasons for its rapid adoption worldwide. However, there is little empirical evidence supporting this prediction, especially for developing countries. I estimate the efficiency gains of introducing a VAT using the synthetic control method. I find that a VAT has, on average, positive and economically meaningful impact on economic efficiency. This result, however, is driven by richer countries only. There is no significant impact of the VAT on poorer countries. The findings are robust across a series of placebo studies and sensitivity checks. (*JEL* H20, H25, O40, E6)

I. INTRODUCTION

The value-added tax (VAT), a broad-based tax on consumption, is one of the youngest, yet among the most important sources of government revenue. In less than 60 years since its introduction, it has replaced various sales and turnover taxes in more than 160 countries and its revenues represent more than 7% of the world's gross domestic product (GDP). These numbers are expected to continue growing due to the increasing reliance on consumption taxes and due to the requirement by many multilateral organizations that their member countries use the VAT.

One of the key rationales for the adoption of the VAT is the belief that it is an efficient tax system. The VAT is collected at each stage of the production chain, but all taxes paid on purchased inputs are fully refunded in a timely manner. In effect, it is levied only on the final consumption and not on the intermediate transaction between firms. It is this feature of the VAT that makes it production efficient. In contrast, if intermediate transactions are taxed, then a profit-maximizing firm will have an adverse incentive to allocate resources to minimize its tax liability rather than allocating resources to their most productive uses. Similarly, taxation of intermediate transaction leads to cascading, a phenomenon where a tax-on-tax occurs as goods pass from one production stage to another. In the presence of cascading, a firm can decrease its tax liability simply by self-producing the inputs, which creates an adverse incentive for firms to vertically integrate to minimize taxable stages.[1] Thus, when a VAT replaces distortionary taxes, a profit-maximizing firm will allocate its resources to produce goods

*This paper is part of my doctoral dissertation at Tulane University and I am very grateful for the support I received from my committee members: Jim Alm, Alan Barreca, and Steven Sheffrin. I want to thank Stefano Barbieri, Sanjukta Basu, Valentina Dimitrova-Grajzl, Romain Duval, John Edwards, Keith Finlay, Tim Harris, Sean Higgins, Michael Keen, Jaewoo Lee, Mario Mansour, Ruud de Mooij, Dimitris Nikolaou, Whitney Ruble, Jay Shimshack, Dan Teles, three anonymous referees, and participants at the National Tax Association conference, Southern Economic Association conference, Jobs, Growth, and Structural Reforms seminar series at the International Monetary Fund, and economics seminar series at Tulane University. I am especially thankful to Steve Sheffrin for providing funds to purchase industrial statistics from the UNIDO, Eric Weding for adding International VAT Monitor to the library collection, and Hideki Fujioka for helping me with some C++ codes. All errors are my own.

Adhikari: Assistant Professor, Department of Economics, Illinois State University, Normal, IL 61790, Phone 309 438 3747, Fax (309)438-5228, E-mail badhik1@ilstu.edu

1. Cascading is unavoidable even in single-stage taxation such as the retail sales tax without a mechanism to fully refund taxes on purchased inputs. For instance, Ring (1989, 1999) and Mikesell (2014) using the U.S. data and Smart and Bird (2009) using the Canadian data estimate that about 35%–45% of the retail sales tax revenue comes from business inputs.

ABBREVIATIONS

DID: Difference-In-Differences
EU: European Union
GDP: Gross Domestic Product
IMF: International Monetary Fund
OPEC: Organization of the Petroleum Exporting Countries
PWT: Penn World Table
RMSPE: Root Mean Square Prediction Error
SCM: Synthetic Control Method
TFP: Total Factor Productivity
VAT: Value-Added Tax
WTO: World Trade Organization

Economic Inquiry
(ISSN 0095-2583)
doi:10.1111/ecin.12847

in the most efficient way, which increases its output per worker, a direct measure of economic efficiency.

The justification for an efficient tax system comes from the production efficiency theorem of Diamond and Mirrlees (1971), which states that the optimal tax system should not distort firms' production decisions even in the presence of other distortions. The production efficiency theorem assumes the absence of tax evasion and informal economy. At the same time, the VAT's efficiency is based on the assumption that it is well designed and well enforced. If any of these assumptions are violated, the VAT becomes a partial tax on production, introducing production inefficiencies. There are many instances under which some of these assumptions are violated. First, almost all countries exempt small firms from the VAT as the costs of administration do not justify the amount of tax collected from these firms. Many developing countries also exempt most of the firms in the agriculture and service sectors due to difficulties in taxing those sectors. Exempt firms do not qualify for tax refunds on their purchased input, which can introduce production inefficiencies. The presence of tax evasion and informal economy that are widespread in developing countries (Alm and Embaye 2013; Schneider 2005) can also distort a firm's behavior (Emran and Stiglitz 2005; Piggott and Whalley 2001; Skinner and Slemrod 1985). Finally, the VAT can push some firms into the informal sector because the cost of complying with a VAT is considered to be higher than that of a sales or a turnover tax, especially for small and medium enterprises (Coolidge 2012; Hines 2004).

Because the design of the VAT, its enforcement, and the broader environment under which it operates are much weaker in poorer countries than in richer countries, a natural question then is whether both richer and poorer countries experience efficiency gains by introducing a VAT? There is, however, surprisingly little evidence evaluating the efficiency gains of the VAT. On top of that, almost all of these studies are exclusively based on high-income countries, largely due to the lack of comparable data across countries on economic efficiency.

This paper uses Penn World Table (PWT) Version 8.1 (Feenstra, Inklaar, and Timmer 2015) data to estimate the impact of introducing a VAT using a direct measure of economic efficiency (i.e., GDP per worker). It also investigates two primary channels through which a VAT affects GDP per worker: by increasing total factor productivity (TFP) and by increasing the stock of physical capital.[2]

The results of this research provide new insight into "VAT and efficiency": the adoption of a VAT does provide significant efficiency gains, but only in richer countries. For instance, 5 years after the reform GDP per worker of the high-income group is 10.9% higher than the control group (all significant at the 1% level). GDP per worker of the upper-middle-income group is also positive and significant at the 5% or better level, but after some time lag. By 5 years after the reform GDP per worker is 25.5% higher compared to the control group. However, I find no significant impact of the VAT on poorer countries. The effects of the VAT on the lower-middle-income countries are slightly positive and the effects on the low-income countries are mostly negative, but these estimates are statistically indistinguishable from zero at the 10% level. The same differential trend in the impact of the VAT is found across income groups for capital stock per worker and TFP, two of the primary channels through which a VAT affects GDP per worker. These results indicate that the theoretical advantages of the VAT do not necessarily translate into practice. The findings are robust across a series of placebo studies and sensitivity checks.

Even when data availability is not a problem, estimating the effect of a VAT adoption on economic efficiency using traditional approaches faces several challenges. Because countries choose whether to adopt a VAT and when to adopt a VAT, cross-country regressions are subject to many biases and are unlikely to reveal the true effect of a VAT on economic efficiency. One way to tackle this problem is to use propensity score matching technique to generate the probability of the VAT adoption and compare countries with a VAT to those countries without a VAT, but with a similar propensity to adopt a VAT. However, King and Nielsen (2019) find that propensity score matching often exacerbates these biases, leading them to suggest replacing propensity score matching with other research designs. On top of that, matching techniques cannot account for the biases caused by unobserved confounders. When the biases by unobservables are of concern, the difference-in-differences (DID) technique can be used. This approach

2. Output-side GDP is a better measure of the productive capacity of an economy than the standard expenditure-side real GDP, which is better at measuring the average living standard of an economy.

allows one to control for both country and year fixed effects so that all time-invariant differences across countries and shocks common to all countries in the same year can be controlled for. However, identifying the effects of a VAT using DID requires the paths of the outcome variable between the VAT adopters and non-VAT adopters to be parallel. As I demonstrate later, this assumption is not satisfied without controlling for country-specific time trends (see Figure 6). However, controlling for country trends using DID potentially biases estimates in the presence of a dynamic treatment effect (Meer and West 2016; Wolfers 2006). Another limitation of the DID is that it assumes the impact of the VAT to be homogenous across reforming countries. Thus, it does not allow us to explore whether the impact of the VAT systematically varies across countries according to their capacity to properly design and implement it.

This paper attempts to address all these challenges. To account for the violation of the parallel trends assumption and the endogeneity caused by time-varying unobserved confounders, I use the synthetic control method (SCM). The SCM is a data-driven research design that constructs counterfactual trajectory of GDP per worker in the absence of a VAT by taking the weighted average of countries without a VAT such that it closely resembles the economic structure and outcomes of the country with a VAT for several years before the adoption of a VAT (Abadie, Diamond, and Hainmueller 2010, 2015; Abadie and Gardeazabal 2003).[3] The SCM calculates the dynamic effects of VAT reforms on a case-by-case basis, which allows for the effects to vary across countries and over time. This enables me to aggregate the dynamic treatment effects into four income groups according to the World Bank's income classification based on the reforming country's level of income at the time of the VAT's introduction. In addition, I also contribute to the improvement of the SCM by developing an intuitive graphical approach of testing whether the average effects are driven by outlier countries by constructing a leave-one-out average effects graph. This test will be useful to researchers applying the SCM to multiple treated units and studying the aggregate impact of the treatment.

3. Abadie, Diamond, and Hainmueller (2010) prove that if a synthetic country can be found such that it matches the pretreatment trajectory of the outcome variable of the treated country, then the size of the bias caused by time-varying unobserved confounders in the difference between the posttreatment outcome variable for the treated and the synthetic control countries goes towards zero as the preintervention period increases.

II. BACKGROUND

A. The Context of VAT Reforms

All tax policies are endogenous to some extent and one of the main reasons to use the SCM is its ability to account for unobserved heterogeneity more flexibly than traditional panel regressions or matching techniques. That said, the context of the VAT adoptions provides a unique and relatively exogenous setting to assess the impact of tax reforms on economic efficiency.

First, VAT take-up decisions are often influenced by external forces that are arguably exogenous to the internal conditions of the reforming countries. Many academic and nonacademic sources highlight the role of multilateral organizations, especially the European Union (EU) and the International Monetary Fund (IMF), in influencing the take-up decision of a VAT. The EU requires that member states adopt a VAT upon entry to the EU. So, any country joining the EU or aspiring to join the EU needs to adopt a VAT. Similarly, the IMF is a strong advocate of the VAT and often puts the adoption of a VAT as one of the conditions for providing a loan, or other assistance. Thus, any country that needs IMF's assistance has a much higher probability of adopting a VAT. For instance, Keen and Lockwood (2010), Ufier (2014), and Cizek, Lei, and Ligthart (2017) find that the probability of a VAT adoption increases by up to 25% in the year following the country's participation in IMF programs. Ebrill, Keen, and Bodin (2001) estimate that more than half of all countries that introduced a VAT during the 1980s and 1990s used IMF advice in doing so.[4]

Second, as Case, Rosen, and Hines (1993) and Besley and Case (1995) demonstrate, the design

4. Riswold (2004) criticizes the role of IMF in sub-Saharan Africa's VAT adoption. James (2015) argues that a significant role is played by other multilateral organizations such as the World Bank (e.g., in the cases of Cameroon, Ghana, Hungary, Philippines, Thailand, Tanzania, and Venezuela), the World Trade Organization (WTO), Inter-American Development Bank, Asian Development Bank, the Organization for Economic Co-operation and Development, as well as aid and development agencies such as the United States Agency for International Development (e.g., in the cases of Serbia, Egypt, El Salvador, Guatemala, and Jamaica), German Aid and Development Agency (e.g., in the cases of Serbia, Croatia, Macedonia, and Pakistan), and the UK Department for International Development (e.g., in the cases of Ghana and Pakistan).

of the tax structure in one jurisdiction influences the design of tax structures in neighboring areas. If countries adopt a VAT due to the influence of their neighbors, then such tax reform will be more exogenous to the economic condition of the adopting country than a reform motivated by their internal economic conditions. The introduction of the VATs indeed followed regional waves. For instance, more than 11 other European countries adopted a VAT within 5 years of France's decision to adopt a VAT. Similarly, 11 other Latin American countries introduced the VAT within a decade of Brazil's decision to adopt a VAT. The copycat behavior is particularly strong in Eastern Europe, where 18 countries adopted a VAT within 5 years of Hungary's VAT adoption. A similar pattern is also found in the developing countries of sub-Saharan Africa and Asia. The copycat behavior in VAT adoption is also demonstrated in more systematic studies by Keen and Lockwood (2010), Ufier (2014), and Cizek, Lei, and Ligthart (2017).

Third, as documented in Table 1 almost all of the countries adopted a VAT to replace turnover taxes (i.e., sales tax, wholesale tax, or manufacturing tax). In few countries, they also used a VAT introduction to reduce other inefficient indirect taxes such as tariffs, excise taxes, and stamp duties or to slightly reduce direct taxes such as individual income tax rate and corporate income tax rate, but these changes were minor compared to the replacement of turnover taxes. This holds in general as well, and not just in the countries covered in this study (Ufier 2017). For instance, in the sample of 88 countries for which I could collect these data, Japan and Panama are the only two countries that had no general turnover taxes before the VAT adoption. Japan significantly reduced various excise taxes (e.g., automobiles sales tax from 23% to 6%) and Panama significantly reduced various stamp duties when they introduced a VAT.[5] Thus, these VAT reforms provide an ideal setting to evaluate the impact of replacing sales and turnover taxes with a VAT on economic efficiency.

B. *Differences in Institutions and VAT Designs*

One of the key rationales for the adoption of the VAT is the belief that it is an efficient tax system. However, the efficiency property of the VAT crucially depends on whether it is well designed and well enforced as well as the broader environment under which it operates. In this section, I describe the properties of the VATs adopted and compare them with the textbook ideal VAT, highlighting the differences in the design and enforcement of the VAT among developed and developing countries. I also provide the information on the presence of the shadow economy, tax capacity, and overall government quality, to highlight the different economic and political environment under which the VATs operate in developed and developing countries.[6]

An ideal VAT issues a full refund of the VAT paid on purchased inputs in a timely manner. To the extent that there are imperfections in the VAT refund system, the VAT acts as a tax on production. In an extensive survey of VAT refund practices in 36 developing and developed countries, Harrison and Krelove (2005) find that VAT refunds are generally paid within 4 weeks in developed countries, while it takes several months, and sometimes more than a year, in developing countries. They also report that firms in many developing countries do not always receive full refunds they are owed. For instance, in some developing countries, tax officials extract payments from firms in return for speeding up refunds, while in others, they require purchases to be made only through a verified bank account or to be certified by verified accountants before any refunds are made. Finally, PricewaterhouseCoopers and the World Bank (2017) find that claiming a VAT refund triggers audit in many developing countries, which discourages firms from claiming a refund and, thus, acts as a tax on input.

An ideal VAT is broad based and is comprehensive in coverage over the whole production-distribution chain. Table 1 provides a brief description of the characteristics of the VATs at the time of adoption. Most of all the 33 countries that are analyzed in this study introduced the VAT to replace some kind of sales or turnover taxes, while some countries also reduced their income tax, payroll tax, stamp duties, or

5. The few cases where the VAT was introduced in a context with very little existing indirect taxation are the recent adoption of the VAT by several Gulf Cooperation Council countries.

6. These information is not readily available, so I collect them from various sources. I start with Tait (1988) and Cnossen (1998) and supplement them with the data from various issues of International VAT Monitor, Tax Notes International, IMF's country documents, and tax authority websites. Even so, I do not have the information for a few of the countries with a VAT. The information on informal economic activity is obtained from Schneider (2005) and the information on overall government quality is obtained from Marshall and Jaggers (2013) and International Country Risk Guide (2011).

TABLE 1
The Properties of the VAT at the Time of Its Introduction

Country	Year	Coverage	Base	Taxes Replaced
Panel A: High-income countries				
Austria	1973	R	G + S	Cascade wholesale tax
Belgium	1971	R	G + S	Cascade wholesale tax
Canada	1991	R	G + S	Manufacturers sales tax, turnover tax
Denmark	1967	R	G + S	Wholesale tax
France	1968	R	G + S	An earlier and less sophisticated VAT
Greece	1987	R	G + S	Turnover tax, stamp duties, and special import levy
Italy	1973	R	G + S	General and local government sales taxes
Japan	1989	R	G + S	Corporate and personal income tax reduced
Netherlands	1969	R	G + S	Cascade wholesale tax
New Zealand	1986	R	G + S	Wholesale tax
Spain	1986	R	G + S	Cascade production tax and 20 other sales taxes
Sweden	1969	R	G + S	Retail sales tax and capital goods tax
United Kingdom	1973	R	G + S	Multirate wholesale tax
Panel B: Upper-middle-income countries				
Argentina	1975	R	G + ST	Wholesales and provincial cascade turnover tax
Chile	1975	R	G + S	Cascade turnover, manufactures tax, and special luxury tax
Ireland	1972	R	G + S	Wholesale and retail sales
Mauritius	1998	R	G + S	Sales tax abolished, hotel and restaurant tax reduced and eventually abolished
Panama	1977	R	G + S	Stamp duties reduced
Portugal	1986	R	G + S	Single-stage wholesale tax
Panel C: Lower-middle-income countries				
Colombia	1975	R	G + S	Reduced income, property, and capital gains taxes
Costa Rica	1975	R	G + ST	Multistage ring system
Dominican Republic	1983	M	G + ST + CG	Reduced reliance of customs duties
Honduras	1976	R	G + S	Single-stage ring system
Jamaica	1991	R	G + ST	Several indirect taxes including retail, telephone, entertainment, excise
Peru	1973	R	G + S	Cascade production tax and stamp tax
Senegal	1980	R	G + S	Manufacturers VAT
Thailand	1992	R	G + S	Business turnover tax
Uruguay	1968	R	G + S	Manufactures single-stage tax and a cascade turnover tax
Panel D: Low-income countries				
Bangladesh	1991	M	G + ST	Turnover tax, sales tax, excise duties
Guinea	1996	M	G + ST	Turnover taxes on production, services and imports
Kenya	1990	M	G + ST	Manufacturer's sales tax and tax on manufactured imports including capital goods
Nepal	1998	R	G + S	Sales tax, contract tax, hotel tax, and entertainment tax
Pakistan	1990	M	G + ST + CG	Cascading sales

Notes: Column 2 (i.e., year) provides information on the year of VAT adoption. In Column 3 (i.e., coverage), R denotes that the VAT is extended through retail stage, W denotes through wholesale stage, and M denotes through manufacturing stage. In Column 4 (i.e., base), G denotes that goods are in the tax base, S denotes services are in the tax base, ST denotes only selected services are in the tax base, and CG denotes some capital goods are in the tax base. In Column 5 (i.e., taxes replaced), cascade production tax refers to a cascade tax on business turnover restricted to the production stage; cascade wholesale tax extends the turnover tax to include the wholesale stage; cascade retail tax extends the turnover tax to include the retail stage; manufacturers, wholesale, or retail taxes are single-stage taxes. All previous taxes were replaced by the VAT unless it says reduced in the table.

tariffs. The design of the VAT, however, is very heterogeneous, with high-income countries designing the VAT closest to the textbook ideal and the low-income countries designing the VAT furthest from the textbook ideal. To elaborate, all 13 of the countries in the high-income group had a broad-based VAT reform that included both goods and services in the tax base and levied the taxes through the retail stage, while only 83% of the upper-middle-income countries, 67% of the lower-middle-income countries, and 20% of the low-income countries had such broad-based VAT reform.

Table 2 provides a brief description of the economic and political environment across four income groups. I find systematic differences in the environment conducive to effective enforcement of the VAT, with the best environment in high-income countries and the worst environment in low-income countries. For instance, the share of economic activity in the shadow economy is 17.3% in the high-income group, which

TABLE 2
The Economic and Political Environment Under Which the VAT Operates

Country	GDP Per Capita	Polity 2 Score	Shadow Economy	Government Quality	Tax Ratio	Agriculture Share
Austria	0.60	1.00	9.8	0.93	26.02	6.76
Belgium	0.58	1.00	21.9	0.95	23.73	4.02
Canada	0.85	1.00	15.7	1.00	31.03	2.86
Denmark	0.71	1.00	17.7	1.00	29.08	7.83
France	0.62	0.50	15.0	0.92	21.82	8.67
Greece	0.44	1.00	27.5	0.66	17.31	
Italy	0.51	1.00	27.0	0.74	15.46	7.6
Japan	0.70	1.00	11.0	0.91	20.93	2.24
Netherlands	0.61	1.00	13.2	1.00	22.42	
New Zealand	0.55	1.00	12.4	0.99	30.60	7.48
Spain	0.39	1.00	22.5	0.75	15.90	6.18
Sweden	0.67	1.00	18.8	1.00	29.45	
United Kingdom	0.61	1.00	12.5	0.93	26.87	2.81
Average of High-Income Countries	**0.60**	**0.96**	**17.3**	**0.91**	**23.89**	**5.65**
Argentina	0.14	0.60	25.3	0.55		10.23
Chile	0.24	−0.70	19.3	0.61		5.75
Ireland	0.36	1.00	15.8	0.85	26.17	15.81
Mauritius	0.34	1.00	22.7		16.17	9.43
Panama	0.21	−0.70	63.5	0.31		
Portugal	0.30	1.00	23.0	0.75	17.87	14.79
Average of Upper-Middle-Income Countries	**0.27**	**0.46**	**28.3**	**0.61**	**20.07**	**11.02**
Colombia	0.24	0.80	37.3	0.46		24.88
Costa Rica	0.26	1.00	25.7	0.66		
Dominican Republic	0.15	0.60	31.9	0.52	5.82	18.84
Honduras	0.09	−0.10	48.3	0.32		27.33
Jamaica	0.13	1.00	34.8	0.42	23.59	
Peru	0.15	−0.70	58.0	0.37	15.59	16.78
Senegal	0.06	−0.20	43.7	0.45		
Thailand	0.16	−0.10	50.6	0.63	17.70	12.65
Uruguay	0.31	0.80	50.6	0.45		
Average of Lower-Middle-Income Countries	**0.17**	**0.34**	**42.3**	**0.48**	**15.68**	**20.09**
Bangladesh	0.04	−0.50	35.3	0.20	4.68	30.25
Guinea	0.05	−0.10	39.0	0.47	10.28	19.24
Kenya	0.05	−0.70	33.2	0.55	13.28	30.19
Nepal	0.04	0.50	36.7		8.13	41.43
Pakistan	0.07	0.80	35.7	0.37	10.54	26.95
Average of Low-Income Countries	**0.05**	**0.00**	**36.0**	**0.40**	**9.38**	**29.61**

Notes: Column 1 contains GDP per capita normalized using the U.S. values a year before VAT adoption, Column 2 contains value of Polity 2 a year before VAT adoption, which ranges from −1 denoting autocracy to 1 denoting democracy, Column 3 contains average share of economy in informal sector from 1999 to 2007, Column 4 contains average government quality from 1982 to 1997, which ranges from 0 to 1, Column 5 contains tax revenue as a share of GDP a year before VAT adoption, and Column 6 contains agriculture as a share GDP a year before VAT adoption.

increases to 36% in the low-income group. I use three different indicators to measure the country's capacity to administer a VAT and I find systematic differences across income groups in all three indicators. These indicators are overall government quality that is constructed by taking the average of corruption, bureaucratic quality, and rule of law indices and ranges from 0 denoting the lowest quality to 1 detonating the highest; political capacity that is measured using Polity 2 score, which ranges from −1 denoting autocracy to 1 denoting democracy; and tax capacity that is measured using tax as a share of GDP. The overall government quality in the high-income group is 0.91, while it is 0.40 in the low-income group, Polity 2 score decreases from the average of 0.96 in the high-income group to 0.0 in the low-income group, and tax capacity decreases from 23.9% of GDP in the high-income group to 9.4% in the low-income group. The share of agriculture in the overall economy gives yet another indicator of how broadly the VAT can be applied. The agriculture sector is notoriously difficult to tax, so a larger share of agriculture implies a narrower base of VAT. I find that the low-income countries have the largest share of the economy in the agricultural sector (29.6%) and the high-income group has the lowest (5.7%).

This differential trend holds in general as well, and not just in the countries covered in this study. For instance using data for more than 150 countries that have adopted a VAT, I find that average government quality ranges from 0.90 in the high-income group to 0.41 in the low-income group, average Polity 2 score ranges from 0.92 in the high-income group to −0.03 in the low-income group, average shadow economy ranges from 17.7% in the high-income group to 39.4% in the low-income group, average tax capacity ranges from 23.1% in the high-income group to 10.5% in the low-income group, and the share of agriculture ranges from 4.9% in the high-income group to in 34 .4% in the low-income group.

To conclude, the design of the VAT, its enforcement, and the broader environment under which it operates are much weaker in poorer countries than in richer countries. However, the efficiency property of the VAT crucially depends on these factors. Thus, the question remains: do all countries, irrespective of their initial income level, experience efficiency gains after introducing a VAT?

C. Related Literature

The empirical analysis of the economic effects of a VAT is surprisingly sparse. They can be divided into two categories: general equilibrium analysis and reduced-form analysis.[7] All of the general equilibrium studies use data from high-income countries such as the United States (Ballard, Scholz, and Shoven 1987), Canada (Piggott and Whalley 2001), Germany (Boeters et al. 2010), and Norway (Bye, Strom, and Avitsland 2012). They find a positive impact of an "idea l" VAT on economic efficiency and other macroeconomic variables. However, once they incorporate imperfections found in real-world VATs in their model, the efficiency gains achieved by adopting a VAT significantly decreases. For instance, Bye, Strom, and Avitsland (2012) find that including selective services in the VAT base reduces welfare compared to not including services in the VAT base or including all services in the VAT base. Ballard, Scholz, and Shoven (1987) estimate that if the United States introduces a VAT similar to those found in the European countries, it reduces the efficiency gains by 50% to 70% compared to a broad-based and single-rate VAT. Finally, Piggott and Whalley (2001) incorporate the presence of informal economy in their model and find that including more services in the VAT base reduces welfare. These studies highlight the role of a narrow VAT base and the informal economy in undermining the effectiveness of the VAT, even in high-income countries.

The reduced-form studies on the economic effects of a VAT have two limitations. First, they use tax as a share of GDP to measure economic efficiency, which is, at the best, an indirect way of measuring economic efficiency.[8] Second, they either cover high-income countries only (Ferede and Dahlby 2012; Keen and Lockwood 2006; Lee, Kim, and Borcherding 2013; Nellor 1987; Smart and Bird 2009), or pool both developed and developing countries together (Keen and Lockwood 2010; Ufier 2014, 2017). They all find that a VAT has a positive impact on various macroeconomic variables such as tax ratio, investment rate, and economic growth. However, all studies except Keen and Lockwood (2010) and Ahlerup, Baskaran, and Bigsten (2015) ignore the potentially differential impact of a VAT on developed countries versus developing countries. Both studies find the impact of a VAT in developing countries to be insignificant. Similarly, Ufier (2017), in a sample of mostly developing and emerging countries, finds that VAT adoptions did not increase the tax ratio, but they decrease government spending, deficit, and debt. The author also explores the heterogeneity of the impact of the VAT with respect to the probability of VAT adoption, although the author does not study the impact of the initial level of income. This paper differs from Keen and Lockwood (2010) and Ahlerup, Baskaran, and Bigsten (2015) both in terms of the outcomes analyzed and research design. In terms of outcomes, I study the impact of the VAT on a direct measure of economic efficiency as well

7. The main advantage of general equilibrium modeling is that it provides a clear bridge between the theoretical and applied aspects of tax policy analysis. However, they also have important limitations. For instance, they make strong and ad hoc assumptions about the model's functional forms, elasticity type, tax treatment, market structure, technology type, and so on that usually do not hold in the real world.

8. Keen and Lockwood (2010) demonstrate that under some conditions access to a more efficient tax instrument can lead an optimizing (but not necessarily benevolent) government to increase the ratio of tax revenue to GDP. Thus, the tax ratio can be used as an indirect measure of economic efficiency. However, there are various other factors for which no cross-country data exist that can increase the tax ratio even in the absence of any changes in the efficiency of the tax system. Some of such examples are increases in the probability of tax audits, increases in the punishment, and reduction of the informal sector.

as on two important mechanisms: productivity and capital stock, whereas the previous studies analyze the impact on an indirect measure of economic efficiency. In terms of the research design, I use quasi-experimental research design using the SCM which can flexibly control for the effects of time-varying unobserved confounders, unlike fixed-effect panel regressions which can only account for the effects of time-invariant confounders.

III. CONCEPTUAL FRAMEWORK AND DATA

A. Conceptual Framework

Suppose the production function in terms of output per worker is given by

$$y_{it} = A_{it} f(k_{it}, h_{it}), \tag{1}$$

where y_{it} is the output per worker of country i at time t, A is the TFP, k is the stock of physical capital per worker, and h is the average stock of human capital.

I focus on the effect of a VAT introduction on y, A, and k. GDP per worker provides a direct empirical identification of efficiency gains associated with the replacement of sales or turnover taxes with a VAT. An increase in economic efficiency means more of some outputs can be produced from existing capital and labor without decreasing the production of any other outputs. Thus, the adoption of an efficient tax system should lead to higher aggregate output.

The introduction of a VAT can increase output through two channels. First, when a distortionary tax such as a turnover tax is replaced with a relatively nondistortionary tax such as a VAT, a profit-maximizing firm will have an incentive to allocate its resources to their most productive uses rather than more tax-favored uses, which increases A, and thus leads to higher output. Second, as investment goods are not taxed under a VAT, a move from a turnover tax to a VAT will decrease the cost of capital and remove the bias against the use of capital-intensive technology in production, which increases k, and thus leads to higher aggregate output.[9] Finally, it can also increase capital stock and productivity by making vertical integration neutral with respect to taxation and by removing distortions across capital asset types or across firm characteristics such as their sources of financing, their degree of informality, or their size.

B. Data

To study the impact of the VAT on economic efficiency, I create a country-by-year panel for the period between 1950 and 2010 by combining various data sources. My primary data source is the PWT 8.1, which covers 167 countries between 1950 and 2011 (Feenstra, Inklaar, and Timmer 2015). PWT 8.1 has made some major additions to the database, thus, for the first time there now exists a database that is highly comparable across countries and over time, which can be used for comparing economic efficiency, TFP, and capital stock.

I am interested in measuring the relative economic efficiency of the countries, so I use output-side GDP per worker as my primary outcome variable (Inklaar and Timmer 2013).[10] The main explanatory variables are the pretreatment averages of stock of physical capital per worker (k) and the stock of human capital per worker (h).

To anchor the paper to the endogenous growth model, I include pretreatment averages of additional control variables that are meant to capture the impact of institutions (democracy and trade openness), macroeconomic environment (inflation rate), and demography (population growth rate) on the production possibilities (Barro 1991; Mankiw, Romer, and Weil 1992). Data on trade as a share of GDP (trade openness) and population growth are also obtained from the PWT. Data to construct a measure of democracy are obtained from Polity IV (Marshall and Jaggers 2013). Following Persson and Tabellini (2007), I classify a country as democratic if the Polity 2 score in the Polity IV data set is strictly positive. The data on inflation are obtained from the World Bank's World Development Indicators.

The data on the treatment indicator, that is the year the VAT was introduced, is obtained primarily from Ebrill, Keen, and Bodin (2001) and supplemented by International Tax Dialogue (2005) and Ufier (2014).

9. Total factor productivity can also be interpreted as a measure of economic efficiency. However, the interpretation of total factor productivity as one of the determinants of GDP per worker follows directly from the definition of the production function.

10. Feenstra, Inklaar, and Timmer (2015) argue that output-side real GDP per worker is a more appropriate measure of the production possibility of the country compared to the commonly used expenditure-side GDP per worker. Expenditure-side GDP reflects the standard of living of the country, which can differ significantly from the productive capacity of the country due to differences in the terms of trade the country faces.

IV. EMPIRICAL STRATEGY

I employ the SCM developed by Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010, 2015) to isolate the impact of a VAT adoption from other influences. The SCM is a data-driven research design that constructs the counterfactual trajectory of GDP per worker in the absence of a VAT by taking the weighted average of countries without a VAT such that it closely resembles the economic structure and outcomes of the country with a VAT for several years before the adoption of a VAT. The SCM builds on DID estimation but provides a more credible and transparent way of constructing the comparison group. Due to these advantages, applications of the SCM has grown very rapidly and Athey and Imbens (2017) consider the development of the SCM to be "…arguably the most important innovation in the policy evaluation literature in the last 15 years."

Suppose that there are $J+1$ countries, where Country 1 is a treated country and the remaining J countries act as potential controls called the donor pool. Let Y_{it}^{NR} be the outcome variable observed for country i at time t with no reform (NR), and Y_{it}^{R} be the outcome variable for country i at time t with reform (R). The sample period is given by $t = 1, \ldots, T_0, T_{0+1}, \ldots, T$, where T_0 denotes the number of pretreatment periods and T_{0+1} denotes the treatment year. The observed outcome variable can be written as

(2)
$$Y_{it} = \begin{cases} Y_{it}^{\mathrm{NR}}, & \text{in the absence of a VAT reform,} \\ Y_{it}^{\mathrm{R}} \equiv Y_{it}^{\mathrm{NR}} + \tau_{it} D_{it}, & \text{in the presence of a VAT reform,} \end{cases}$$

where $\tau_{it} = Y_{it}^{\mathrm{R}} - Y_{it}^{\mathrm{NR}}$, is the effect of the reform for unit i at time t and $D_{it} = 1$ if $t > T_0$ and $i = 1$, else $D_{it} = 0$.

For each treated country, I can observe Y_{it}^{R}, however, I need to estimate the counterfactual Y_{it}^{NR}, which is the GDP per worker of the country that adopted VAT had the country not adopted it. In order to estimate the counterfactual, I use a linear factor model of the form:

(3)
$$Y_{it}^{\mathrm{NR}} = \alpha_t + \theta_t Z_i + D_i(R_i = R_j) + D_i(I_i = I_j) + \lambda_t \mu_i + \varepsilon_{it},$$

where α_t is an unknown common factor with constant factor loadings across countries and Z_i is a vector of observed covariates with coefficients θ_t. In my main specification, Z_i are the pretreatment averages of GDP per worker, stock of physical capital per worker, stock of human capital per worker, democracy dummy, trade openness, inflation rate, and population growth rate. To prioritize that SCM chooses the synthetic controls from the same geographic region and the same income group, if available, I also include two indicator variables in the main specification: $D_i(R_i = R_j)$, which is an indicator variable that becomes 1 when the treated country and the control country are from the same geographic region and $D_i(I_i = I_j)$, which is an indicator variable that becomes 1 when the treated country and the control country are from the same income group. $\boldsymbol{\mu}_i$ is a vector of unknown parameters, $\boldsymbol{\lambda}_t$ is a vector of unobserved common factors, and ϵ_{it} are the idiosyncratic error terms with zero means. Note that this specification allows country effects to vary with time ($\boldsymbol{\lambda}_t \boldsymbol{\mu}_i$), unlike in DID that restricts country effects to be time-invariant ($\boldsymbol{\lambda} \boldsymbol{\mu}_i$).

The synthetic control for each treated country is a weighted average of the units in the donor pool, which can be represented by a $(\boldsymbol{J} \times \mathbf{1})$ vector of weights, $\boldsymbol{W} = (w_2, \ldots, w_{J+1})'$ such that $w_j \geq 0$ for $j = 2, \ldots, J+1$ and $w_2 + \cdots + w_{J+1} = 1$, where vector $\boldsymbol{W}$ represents a potential synthetic control. Then, the outcome variable for each potential synthetic control unit is given by $\sum_{j=2}^{J+1} w_j Y_{jt}$. The synthetic control algorithm chooses the optimal country weights, W^*, such that it matches the pretreatment trajectory of the outcome variable of the treated country as closely as possible.

Thus, the dynamic treatment effect for time t (τ_{1t}), where $t \in \{T_{0+1}, \ldots, T\}$ can be estimated by $\hat{\tau}_{1t}$, where

(4)
$$\hat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}.$$

Abadie, Diamond, and Hainmueller (2010) show that the size of the bias in the difference between the posttreatment outcome variable for the treated and the synthetic country is close to zero if the pretreatment fit is strong and the number of preintervention periods is large relative to the scale of the transitory shocks. Thus, we need to make sure that the pretreatment fit between the treated country and the synthetic control is strong. Abadie, Diamond, and Hainmueller (2010) use root mean square prediction error (RMSPE) of the outcome variable to measure fit or lack of fit between the paths of the outcome variable for the treated country and its synthetic counterpart. I use the normalized RMSPE called the pretreatment fit index to assess the overall quality of the pretreatment fit. A pretreatment fit index of X implies that the fit of the path of the outcome variable of treated country and its synthetic control is equal to that created by a $100X$ percent deviation of outcome variable on each pretreatment period (Adhikari et al. 2018; Adhikari and Alm 2016).[11]

Next, I follow Cavallo et al. (2013) and aggregate the country-specific treatment effects into four groups based on income classification of the country during the year of VAT adoption. I normalize the estimates before aggregating them by setting the GDP per worker of each treated country equal to one in the treatment year. Aggregating it across countries of group G, where G contains countries $\{1, \ldots, g, \ldots, G\}$ gives the average dynamic treatment effect for group G at time t, or

$$(5) \quad \mathrm{DTE}_t^{\mathrm{G}} = \left(\left(\sum_{g=1}^{G} \hat{\tau}_{1t}^{g}\right) \Big/ G\right) = \left(\left(\sum_{g=1}^{G} Y_{1t}^{g} - \sum_{j=2}^{J_{g+1}} w_j^{g*} Y_{jt}^{g}\right) \Big/ G\right).$$

A. Inference

To evaluate the significance of the treatment effects, I conduct a series of placebo experiments by iteratively estimating the placebo treatment effect for each country in the donor pool (i.e., untreated countries) by assuming that they implemented a VAT reform in the same year as the country of interest and running the SCM. This iterative procedure provides a distribution of estimated placebo treatment effects for the countries where no intervention took place. If the placebo experiments create placebo treatment effects of magnitude greater than the one estimated for the treated country in more than 10% of the placebo experiments (i.e., if the corresponding p value is greater than 0.1), then I conclude that there is no statistically significant evidence of an effect of a VAT reform in the treated country. However, if the placebo experiments create placebo treatment effects of magnitude greater than the one estimated for the treated country in less than 10% of the placebo experiments (i.e., if the corresponding p value is less or equal to 0.1), then I conclude that there is a statistically significant evidence of an effect of a VAT reform in the treated country. Assuming that I am doing inference about a positive treatment effect, the p value at each posttreatment year $t \in \{T_0 + 2, T_0 + 3, \ldots, T\}$ is given by

$$(6) \quad \Pr(\hat{\tau}_{1t} < \hat{\tau}_{jt}) = \left(\left(\sum_{j=2}^{J+1} 1(\hat{\tau}_{1t} < \hat{\tau}_{jt})\right) \Big/ J\right),$$

where $j = 1$ denotes the treated unit and $j \neq 1$ denotes the placebo units, and $1(\cdot)$ is the indicator function.

As I am interested in drawing inference about the significance of the aggregate effects, I need to account for the fact that the average smooths out some noise. To that end, I follow Cavallo et al. (2013) and compute p values for the average treatment effect for group g at each postintervention year $t \in T_{0+2}, \ldots, T$ in the following way. First, for each country i of the group g, I compute the placebo effect for all J_i placebo units from the placebo pool. Second, at each posttreatment year, I compute every possible placebo average effect by picking a single placebo estimate corresponding to each country i and then taking the average across the G placebos. This results in $N_{\overline{\mathrm{PL}}}$ placebo averages where $N_{\overline{\mathrm{PL}}} = \prod_{j=1}^{G} J_i$.[12] Third, at each posttreatment year, I calculate the p values for the dynamic treatment effect obtained in Equation (5) by using the following equation:

$$(7) \quad \Pr(\bar{\tau}_{gt} < \bar{\tau}_{gt}^{\mathrm{PL}}) = \left(\left(\sum_{1}^{N_{\overline{\mathrm{PL}}}} 1(\bar{\tau}_{gt} < \bar{\tau}_{gt}^{\mathrm{P}L})\right) \Big/ N_{\overline{\mathrm{PL}}}\right),$$

where $1(.)$ is the indicator function, $\bar{\tau}_{gt}$ is the average treatment effect for group g at time t after the treatment, and $\bar{\tau}_{gt}^{\mathrm{PL}}$ is the placebo average treatment effect for group g at time t.

11. The RMSPE is defined as follows: $\sqrt{(1/T_0)\sum_1^{T_0}(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt})^2}$. The pretreatment fit index is given by RMSPE$/((1/T_0)\sum_1^{T_0}(Y_{1t})^2)$ The pretreatment fit index provides an index number that makes assessing the quality of fit intuitive as well as facilitates the comparison of pretreatment fit between the treated country and synthetic controls across different countries, especially when GDP per worker varies quite significantly across the treated countries.

12. Note that this number grows rapidly when there are many countries in the group or when there are many placebo units for each country in the group, and especially when both are true.

B. Sample Restrictions

The first step in using the SCM to evaluate the impact of the VAT reform is to choose suitable countries for analysis from the group of countries that adopted the VAT. As I want to analyze the impact of the VAT up to 10 years after the treatment, I choose countries that adopted a VAT before 2000. Next, the SCM requires at least a few pretreatment observations of GDP per worker to calibrate the synthetic control. If some treated countries do not have pretreatment data on GDP per worker, then those countries are excluded from the study. When data availability is not an issue, I restrict the sample period to 10 pretreatment years to calibrate the synthetic unit and 10 posttreatment period to evaluate the impact of the treatment. It becomes difficult to isolate the impact of a VAT if the country experiences major shocks to its economy around the time of the VAT adoption. Thus, I also exclude countries that were known to have shocks of bigger magnitude (e.g., civil war) or those that have higher propensity to experience idiosyncratic shocks such as, transitional countries, small island countries, and resource-rich countries.[13] Finally, some treated countries do not have pretreatment data on some of the predictors used, when that happens, I exclude that predictor from the specification.[14]

The second step in using the SCM is to select the donor pool for each treated country. For any country to be included in the donor pool, it needs to meet two requirements. First, there cannot be any missing observations for the outcome variable in the sample period. Second, there must be at least one nonmissing observation before the treatment for each of the predictors used in the estimation. Any country not meeting these requirements is dropped from the donor pool. Furthermore, I exclude from the potential donor pool any country that adopted a VAT before or within the sample period (i.e., 20-year window around the VAT adoption) because the synthetic unit is meant to reproduce the path of the outcome variable that would have been observed for the treated unit in the absence of treatment. Including any country in the donor pool that was treated in the sample period implies that the synthetic unit is not reproducing the potential outcome in the absence of treatment. I, however, allow countries that eventually adopt a VAT to be in the control group if they implement a VAT 10 years or later after the treated country in question. I also exclude transitional countries, small island countries, resource-rich countries, and countries that went through the civil war in the sample period from the donor pool.

In the final step, I run the synthetic control algorithm and aggregate their effects by income group. While aggregating the treatment effects, I discard the treated countries for which the SCM was not able to produce a synthetic unit with a good pretreatment fit. Note that discarding from the analysis the countries with poor pretreatment fit is similar to confining the analysis to the common support when using matching estimators (Cavallo et al. 2013). I use the pretreatment fit index to measure the quality of the fit and exclude any country with a pretreatment fit index greater than 0.10. For the same reason, I also exclude placebo countries with pretreatment fit index greater than 0.10 from the placebo experiment. In a few cases, SCM is unable to find any placebo countries with a good pretreatment fit, leaving the placebo pool empty. When that happens, I also exclude those treated countries with the empty placebo pool. These sample selection restrictions meant I had to exclude 28 treated countries from the analysis. The majority of these countries were from the low-income group (18), followed by the lower-middle-income group (6), the high-income group (3), and the upper-middle-income-group (1).[15]

The final sample includes 54 countries, with 33 in the treatment group.[16] In some of the analysis, I divide the countries into four groups according to the World Bank's income classification, which results in 13 high-income countries

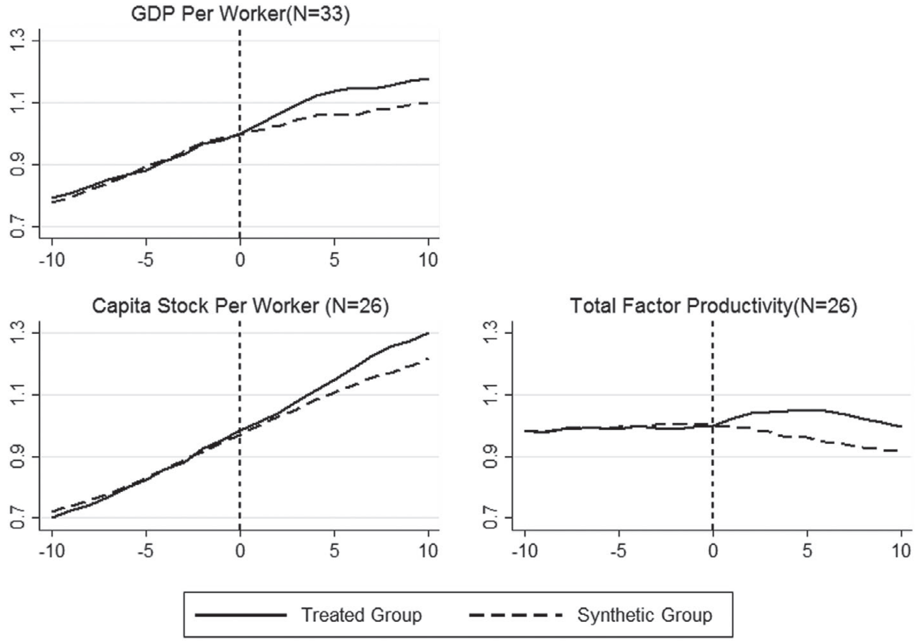
13. Almost all of the transitional countries adopted a VAT within few years of transitioning to a market economy, thus the effects would be commingled with the effects of wider structural reforms. Very few small island countries and resource-rich countries have a VAT or any sales or turnover taxes for that matter. I define resource-rich countries as OPEC members plus countries whose petroleum and natural gas share of GDP (10-year average) is 10% or more than the OPEC average using World Development Indicators data.

14. Because the PWT has a very long and wide coverage of national income data, these are, in practice, very mild restrictions.

15. Of note, I find significant and consistent results for the high-income group only and I was able to include almost all of the high-income countries because I was able to find a good fit synthetic control for them. Thus, the exclusion of countries with bad pretreatment fit does not affect my most consistent and significant results.

16. For each of the countries analyzed, the list of countries in the donor pool and their respective weights in the construction of the synthetic unit is presented at the end of the paper. The weights and balance for each of the predictor variables are presented in Section V of Appendix S1.

FIGURE 1
The Aggregate Impact of VAT on GDP per Worker, Capital Stock per Worker, and TFP Across All Countries.



Notes: The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome of its synthetic control. *N* denotes the total number of countries. The left axis consists of labels for the outcome variable. To avoid the size effect caused due to varying levels of outcome variable across countries, the estimates are normalized before aggregating them by setting the outcome variable of each treated country equal to 1 in the treatment year

(H), 6 upper-middle-income countries (UM), 9 lower-middle-income countries (LM), and 5 low-income countries (L).[17]

V. MAIN RESULTS

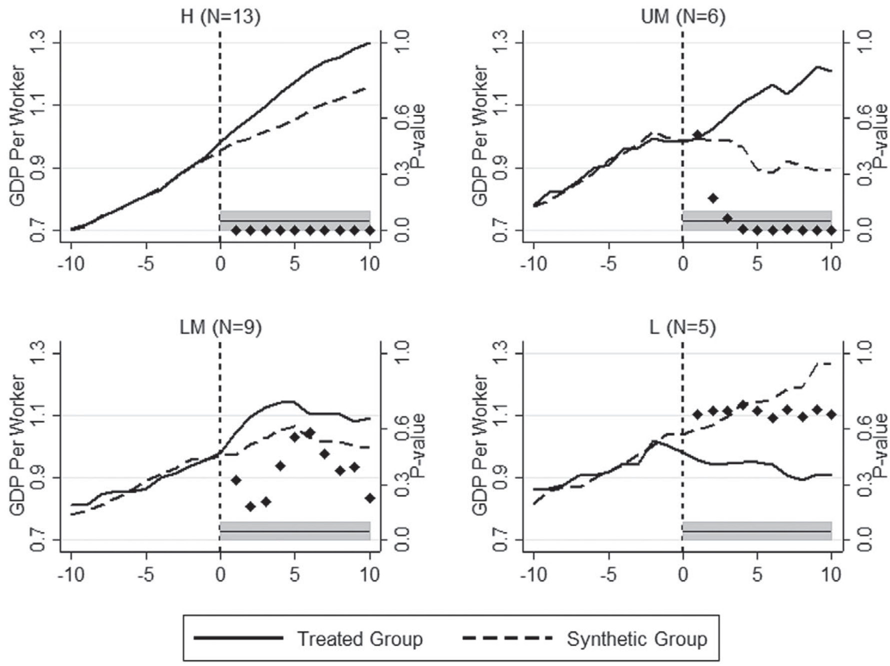
This section presents the results of the empirical analysis. I first present evidence that the VAT increases economic efficiency in the aggregate sample of 33 treated countries. I then analyze the heterogeneity in the effects across countries, showing that the VAT is efficiency improving in richer countries, but not in the poorer countries.[18]

Figure 1 presents the aggregate impact of the VAT reforms on GDP per worker for all treated countries. The aggregate effects are positive and economically meaningful. Five years after the reform, the average GDP per worker of all treated countries is 7.6% higher than the synthetic group. The result is consistent with the theory that suggests that a VAT adoption will have a step effect that could build up over time given the lags of the economy. The result is also consistent with

17. The use of SCM allows us to analyze the impacts of treatment on the much smaller sample than the traditional regression analysis. For instance, the original SCM papers used SCM as a case study approach and several recent papers have used the SCM in settings similar to mine with multiple countries. Cavallo et al. (2013) aggregate 44 individual country cases into 3 groups: largest disasters (4 countries), large disasters (18 countries), and medium disasters (22 countries), Smith (2015) aggregates 17 individual country cases into two groups: non-OECD (12 countries) and OECD (5 countries), and Acemoglu et al. (2016) aggregate SCM estimates of multiple treated firms. Finally, many other studies, such as Campos, Coricelli, and Moretti (2019), Nannicini and Billmeier (2011), and Billmeier and Nannicini (2013) use SCM to analyze multiple country cases even though they do not aggregate the impacts.

18. There is some heterogeneity in the impact of the VAT even within the income group, nevertheless, the distribution of treatment effects clearly indicates that higher the income at the time of the VAT adoption, higher the chance that the VAT had a positive impact. The full distribution of treatment effects is presented Figures 15–19 in Appendix S1.

FIGURE 2
The Impact of VAT on GDP per Worker Across Different Income Groups.



Notes: The four income groups are high-income (H), upper-middle-income (UM), lower-middle-income (LM), and low-income (L) countries. *N* denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome of its synthetic control. The left axis consists of labels for the outcome variable, and the right axis consists of labels for p values (the probability that the gap in outcomes would occur by chance for each year after the treatment), which are shown with the scatter plot, where the area below.10 is shaded in gray and a horizontal line is drawn to indicate.05. To avoid the size effect caused due to varying levels of outcome variable across countries, the estimates are normalized before aggregating them by setting the outcome variable of each treated country equal to 1 in the treatment year

Ufier (2014), which finds that VAT adoption is associated with GDP per capita growth rate ranging from 0.38% to 2.17%. However, are these positive results representative of all countries?

Figure 2 presents the dynamic treatment effects for countries in each of the income groups, which allows me to test the differential effects of a VAT by the development stage at which it was adopted. I find that VAT adoptions have positive and significant impacts on the high-income and the upper-middle-income countries, but mixed and insignificant impacts on the lower-middle-income and the low-income countries. For the high-income countries, 5 years after the reform GDP per worker is 10.9% higher than the synthetic group, and it increases to 11.2% higher 10 years after the reform, all significant at the 1% level.[19] Similarly, I find that the upper-middle-income countries have a positive and significant impact on GDP per worker, although after a time lag. The estimates are significant at the 10% level from Year 3 onwards and at 5% level from Year 4 onwards. I find that 5 years after the VAT adoption, GDP per worker is 25.5% higher in the upper-middle-income group than their synthetic controls, which increases to 33.1% higher 10 years after the reform. For the lower-middle-income group, I find average effects ranging from 6.7% to 9.3% higher than the synthetic control, but the estimates are indistinguishable from zero at the 10%

19. Because the literature usually uses GDP per capita growth rate to measure economic growth, the estimates in the paper are not directly comparable to the GDP per capita growth rate estimates. Thus, I also use GDP per capita estimates as a robustness test, which are presented in Section H of Appendix S1. The GDP per capita estimates are smaller than the GDP per worker estimates. For instance, 5 years after the VAT adoption, GDP per capita is around 5% to 6% higher in the treated high-income group than their synthetic counterparts.

level. In the case of the low-income countries, I find negative but statistically insignificant impact of the VAT.

VI. POTENTIAL MECHANISMS

As demonstrated in Section 3, the introduction of a VAT can increase GDP per worker through two primary channels: by increasing capital stock and by increasing TFP. Thus, in this section, I explore how the VAT affects these variables using the SCM. To make the estimates as comparable as possible, I only run the analysis on the countries that were included in the main results.[20]

Figure 1 presents the average impact of the VAT reforms for all treated countries. As in GDP per worker, the overall effects are positive and economically meaningful for both capital stock per worker and TFP. By the end of the sample period, capital stock per worker is 6.1% higher and TFP is 8.9% higher in the treated group compared to the control group. However, are these results representative of both high-income and low-income countries?

A. *Capital Stock*

The results for the dynamic treatment effects of the VAT reforms are presented in Figure 3. The same differential trend in the impact of the VAT is found across income groups for capital stock per worker. For the high-income countries, I find a positive and significant impact of the reform at the 5% level for all years except Years 5–7 where the estimates are significant at the 10% level. Five years after the reform, capital stock per worker is 5% higher in the treated group compared to the synthetic group, which increases to 7.1% by the end of the sample. However, the average effects are positive but not significant for the middle-income countries and significantly negative for the low-income countries. By the end of the sample, the low-income countries with a VAT have 10% lower capital stock per worker compared to the synthetic control.

B. *Total Factor Productivity*

Figure 4 presents the results for the average impact across income groups. I find the same differential trend in the impact of the VAT across income groups for TFP. All dynamic treatment effects for the high-income countries are positive and significant at the 1% level. Five years after the reform, TFP is 10.1% higher than the synthetic group, which increases to 11.8% by the end of the sample. The estimates for the upper-middle-income countries are also positive and significant at the 10% or better level, but after a time lag. Note that the trend in TFP of the treated group starts declining a few years before the reform, but the trend starts improving immediately after the reform and by Year 4 the treated group has higher productivity than the synthetic group. The lower-middle-income group experiences an immediate jump in productivity after the reform, which starts declining after a few years, but always remains higher than the control group. Five years after the reform, TFP of the treated group is 9.9% higher compared to the synthetic group and at its highest, the TFP of the treated group is 11.6% higher than the synthetic group. Moreover, seven of the dynamic effects are significant at the 5% level and the remaining three dynamic effects are significant at the 10% level. However, I find no significant impact on the low-income group. The TFP of the low-income group was on a decreasing trend before the adoption of the VAT and it continued its decreasing trend after the VAT adoption.[21]

VII. ROBUSTNESS TESTS

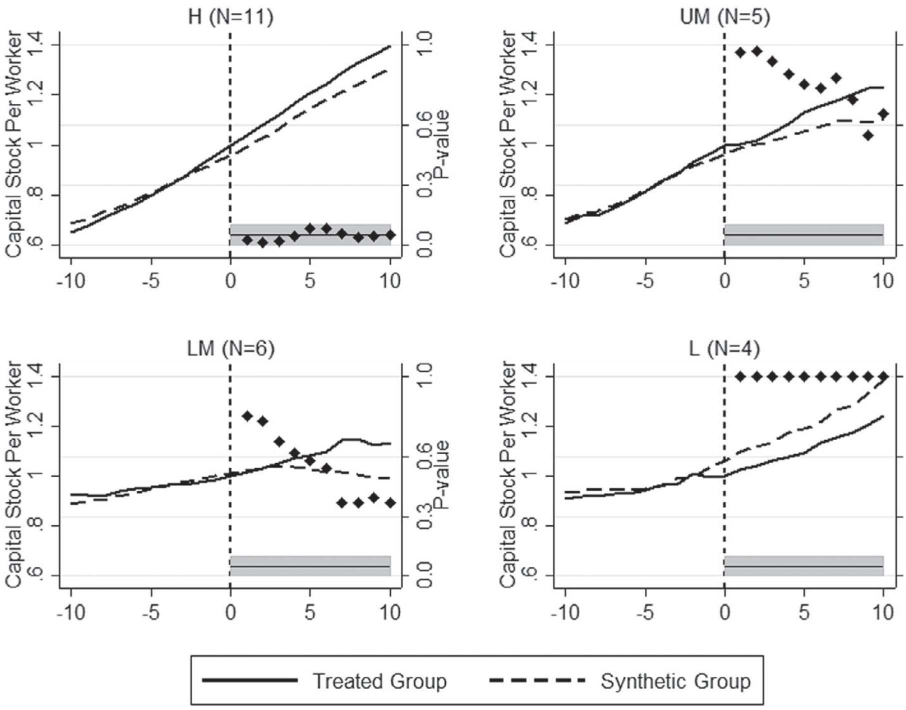
To probe the robustness of these results, I perform a series of additional tests.

20. Following the selection criteria of the main results, some country cases were not included if the pretreatment fit was worse than the threshold of 0.10 or if the donor pool was empty. In some cases, data on capital stock or total factor productivity were not available. Thus, I had to exclude two high-income countries, one upper-middle-income country, three lower-income countries, and one low-income country from the capital stock analysis. Similarly, I had to exclude one upper-middle-income country, one lower-middle-income country, and four low-income countries from the total factor productivity analysis. The covariates used in the estimation of capital stock per worker and total factor productivity are pretreatment averages of the mechanism variable in question, growth rate of real GDP per capita, stock of human capital, openness, inflation, democracy dummy, an indicator variable that become 1 when the treated country and the control country are from the same geographic region, and an indicator variable that become 1 when the treated country and the control country are from the same income group.

21. Only Kenya from the baseline low-income group has data on TFP, so I could not run the analysis on other four low-income countries included in the baseline. When I extended the sample to all low-income countries, nine such countries had data; however, I could not find synthetic control groups for any of them within the baseline pretreatment fit threshold. Thus, in Figure 26 of Appendix S1, I present raw time series lines of TFP for all nine low-income countries and their average. The time series lines suggest no positive impact of the VAT. The average, as well as most of the individual total factor productivity, continued its decreasing trend after the VAT reform.

FIGURE 3

The Impact of VAT on Capital Stock per Worker Across Different Income Groups.



Notes: The four income groups are high-income (H), upper-middle-income (UM), lower-middle-income (LM), and low-income (L) countries. *N* denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome of its synthetic control. The left axis consists of labels for the outcome variable, and the right axis consists of labels for *p* values (the probability that the gap in outcomes would occur by chance for each year after the treatment), which are shown with the scatter plot, where the area below .10 is shaded in gray and a horizontal line is drawn to indicate .05. To avoid the size effect caused due to varying levels of outcome variable across countries, the estimates are normalized before aggregating them by setting the outcome variable of each treated country equal to 1 in the treatment year

A. Leave-One-Out Averages

I develop a graphical test to inspect whether the average effect is driven by a single outlier country or not by plotting the leave-one-out averages of the difference in the outcome variable between the treated and synthetic groups. The leave-one-out average is constructed by iteratively removing one treated country and its synthetic counterpart and taking the average of the remaining $N-1$ countries. If the average effect is not driven by any particular country then the leave-one-out averages should track the overall average very closely.[22] Figure 5 presents leave-one-out averages for GDP per worker. The leave-one-out averages track the average result pretty closely across all four income groups. Although in a few cases removing one country from the average slightly changes the size of the effect, it never changes the sign. Thus, I can conclude that the main results are not driven by the inclusion of any treated country in particular.[23]

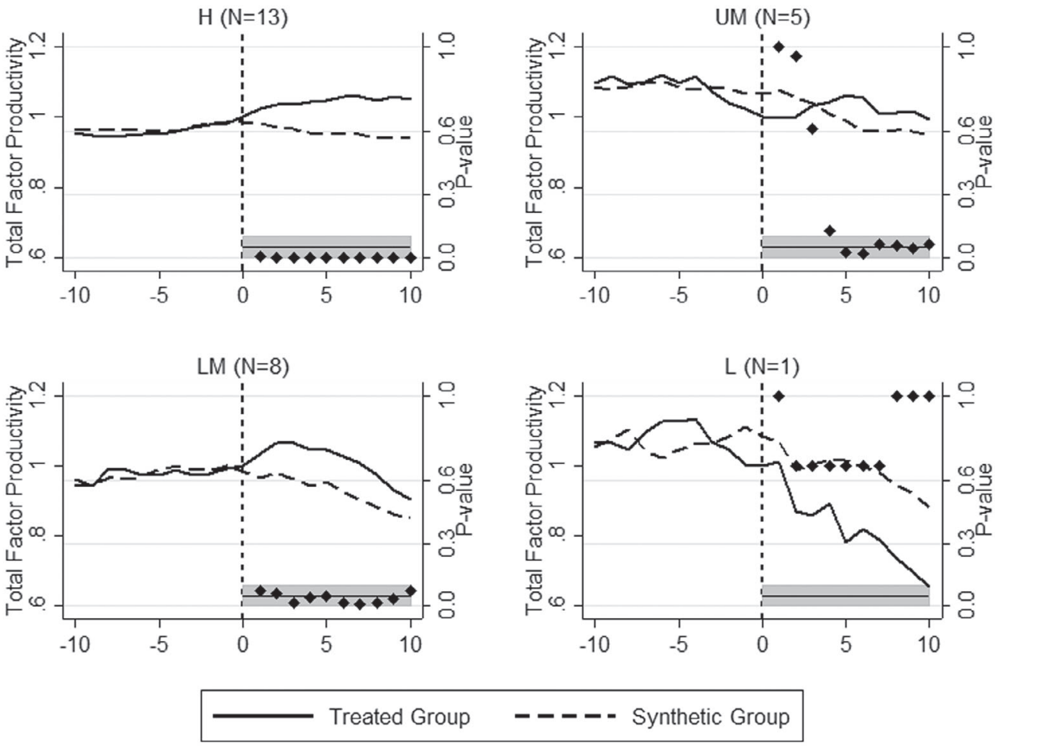
B. B. Controlling for Contemporaneous Events

In some cases, the VAT was adopted as part of the trade liberalization or the EU membership. Thus, the baseline estimates might confound their impact. To test whether the results

22. This test can be generalized to leave-*k*-out test. To do so, first rank the countries according to the size of the effect. Next, calculate the leave-*k*-out averages by iteratively removing *k*-treated countries and their synthetic counterparts and then take the average of the remaining $N-k$ countries.

23. In Figure 24 and 25 of Appendix S1, I present leave-one-out averages for capital stock and total factor productivity, respectively. I do not find that any one country is influencing the results.

FIGURE 4
The Impact of VAT on TFP Across Different Income Groups.



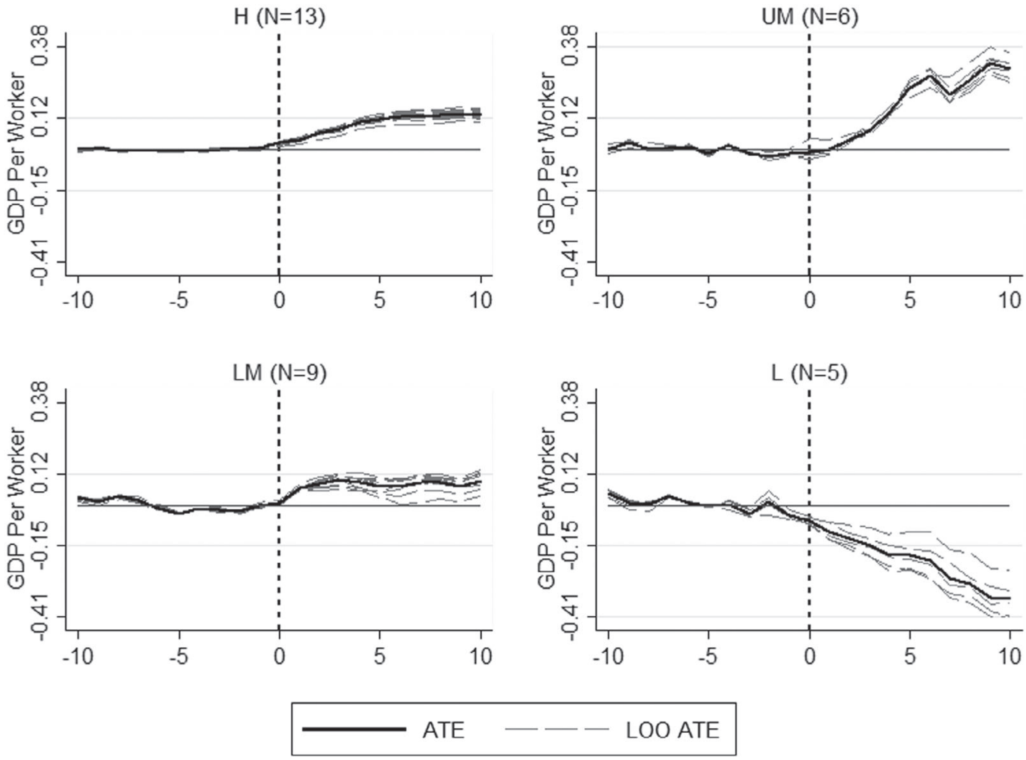
Notes: The four income groups are high-income (H), upper-middle-income (UM), lower-middle-income (LM), and low-income (L) countries. *N* denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome of its synthetic control. The left axis consists of labels for the outcome variable, and the right axis consists of labels for *p* values (the probability that the gap in outcomes would occur by chance for each year after the treatment), which are shown with the scatter plot, where the area below .10 is shaded in gray and a horizontal line is drawn to indicate .05. To avoid the size effect caused due to varying levels of outcome variable across countries, the estimates are normalized before aggregating them by setting the outcome variable of each treated country equal to 1 in the treatment year

are confounded, I calculate the average treatment effect by removing the countries that liberalized their economy or joined the EU in the entire sample window (i.e., 20 years around the VAT adoption). I use two of the most widely used measures of economic liberalization: WTO membership and a binary indicator originally constructed by Sachs and Warner (1995) and updated by Wacziarg and Welch (2008).

The results removing these six countries that joined the EU within 10 years of the VAT adoption is very similar to the baseline results. For the high-income countries, the impact of the VAT controlling for the EU membership ranges from 6.6% to 14.6% higher compared to the synthetic group, all significant at the 1% level. For the upper-middle-income countries, the impact is positive and significant, but after a 2-year time lag. The statistically significant estimates range from 4.2% to 29.5% higher compared to the synthetic group.

The aggregate effects controlling for the WTO membership (14 countries in total) are also very similar to the baseline result. For the high-income countries, I find the impact on GDP per worker is positive and statistically significant at the 1% level, although the size of the effects are slightly smaller (ranging from 1.9% to 7.4%). Similarly, the upper-middle-income countries also have a positive and significant impact, after a time lag. Five years after the reform, GDP per worker is 22.4% higher, which increases to 37.1% by the end of the sample period. The lower-middle-income countries

FIGURE 5
The Average Effect Versus Leave-One-Out Average Effects of VAT on GDP per Worker Across Different Income Groups.



Notes: The four income groups are high-income (H), upper-middle-income (UM), lower-middle-income (LM), and low-income (L) countries. N denotes the number of countries from each income group. The solid black line denotes the difference in the outcome of the group of treated countries and their synthetic controls, and the dashed black line denotes the difference in the outcome of $N-1$ treated countries and their synthetic controls. The left axis consists of labels for the difference in the normalized outcome variable

experience economically meaningful, but statistically insignificant improvement. The results controlling for Wacziarg and Welch (2008) economic liberalization are also similar to the baseline result.

Finally, this trend holds even when I control for the EU membership, WTO membership, and economic liberalization at the same time. Thus, I conclude that the results are driven by the VAT adoption itself and not by other contemporaneous events.[24]

C. *Expanded Set of Countries*

In the main analysis, I imposed various sample selection restrictions and pretreatment fit restrictions that are standard in the SCM literature to ensure that the synthetic controls provide a credible estimate of the counterfactual outcomes. However, due to such restrictions, I was only able to include 33 treated countries. In this robustness test, I test if my results are sensitive to those restrictions as well as how my results change when I expand the number of treated countries analyzed to 61.[25]

24. All the figures of these results are presented in Figures 22 and 23 of Appendix S1.

25. I relaxed most of my sample restrictions, keeping only the restrictions that if a country is flagged as an outlier in my primary data (PTW), then I do not include it in the analysis and that countries need to have at least 5 years of pre-VAT data for finding a synthetic control group with matching pretrends. I also decreased the post-VAT analysis window to 5 years. I was able to almost double the number of treated countries analyzed from 33 countries to 61 countries. The number of countries analyzed increased from 13 to 19 in the high-income group, 6 to 13 in the upper-middle-income group, 9 to 16

The results for the high-income group are only slightly smaller than the original analysis. For instance, 5 years after the reform, GDP per worker of the high-income group is 8.3% higher (compared to 10.9% in the original analysis) than their synthetic group. All estimates are significant at 5% or better level as well. The results for upper-middle-income countries become negative, while the lower-middle-income countries become positive, although none of the positive or negative estimates are statistically significant at the 10% level. The impact on the low-income group remains negative but the estimates are more precisely estimated.[26]

D. Conventional DID

The literature on the economic impact of tax reforms, such as a VAT, often uses the DID research design. However, the DID design suffers from two important shortcomings in my setting due to the violation of parallel trends before the adoption of the VAT and the lack of power to analyze small subsamples of countries. Thus, in my main analysis, I use the SCM, which addresses both of these issues. In this section, I estimate the impact of VAT reforms using the conventional DID method to test whether the estimates are dependent on the use of the SCM rather than the DID method. I follow the event study approach by adding leads and lags of the treatment, omitting the year before VAT adoption from the regression. This design also allows me to test whether the parallel trends assumption is violated or not and to assess the dynamic nature of the VAT's impact on economic efficiency. If coefficients on the leads are significantly different from zero, then this might indicate the failure of the DID approach to create a comparison group with counterfactual trends parallel to the treatment group. In such a case, the use of DID produces biased results. The coefficients on lags describe the transition, capturing the average effect of tax reform in years following adoption relative to the effect before the adoption.

The results using conventional DID are presented in Figure 6. The figure in the left panel includes country fixed effects, year fixed effects, and predictors used in the baseline SCM estimation, while the figure in the right panel also includes country-specific linear time trends. The figure in the top row is obtained by running a regression on the sample of 54 treated and control countries used in the baseline SCM estimates. The figure in the bottom row includes additional treated countries that were excluded from the baseline SCM estimates due to their poor-pretreatment fit and their respective control countries, resulting in an extended sample of 76 countries. The result from the extended sample provides another test for the sensitivity of baseline SCM estimates to the selection of treated countries based on the pretreatment fit criteria. In all specifications, the standard errors are clustered at the country level.

In both of the figures in the left panel, most of the leads are significant, indicating that just using country and year fixed effects violates the parallel trends assumptions. I can control for the differential trends by using country-specific linear trends in the estimation. The results controlling for differential country trends are presented in the right panel. Both figures in the right panel indicate that the VAT adoption has a positive and significant impact on economic efficiency. These effects increase with time for the first few years and then start decreasing. To conclude, the DID estimates provide further assurance that my main results are not the artifact of the SCM methodology or the selection of treated countries based on good pretreatment fit.

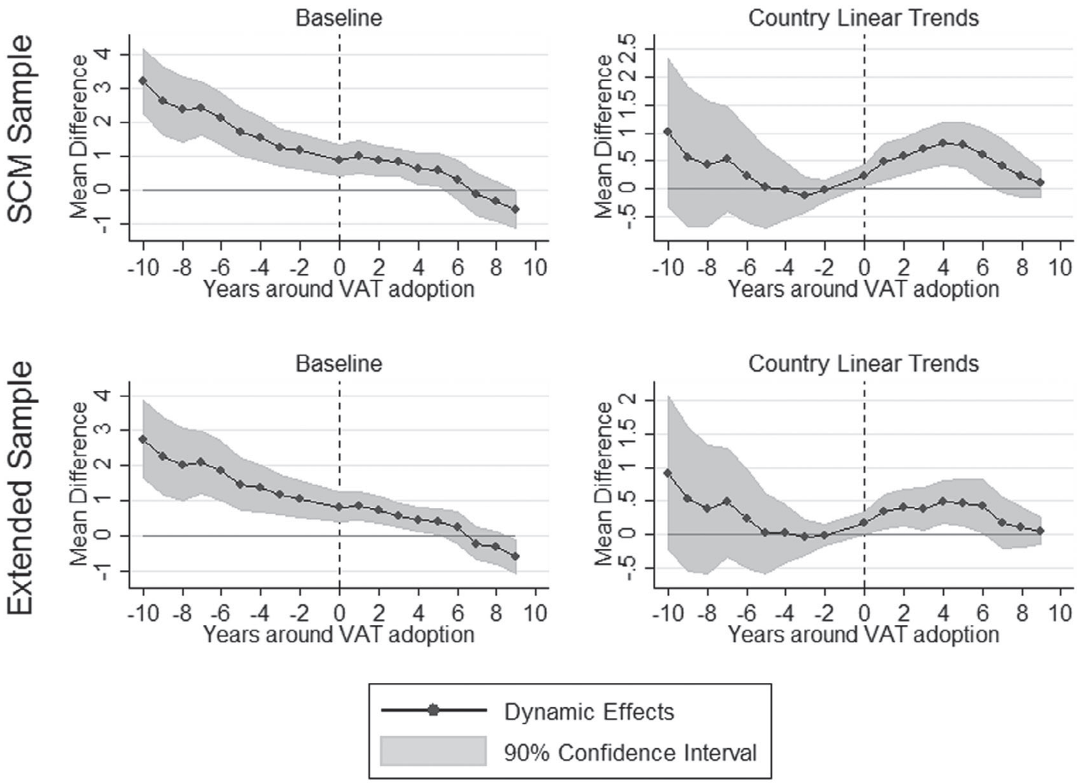
E. Additional Robustness Tests

I also conduct several additional robustness tests in Appendix S1. First, I examine the impact of the VAT using more disaggregated data from the manufacturing industry. Second, I estimate the impact of the VAT by restricting the donor pool to the same geographic region or the same income group. Third, I construct synthetic controls by matching on pretreatment averages of various additional variables suggested in the tax and growth literature such as government consumption as a share of GDP, rents from natural resources (oil, gas, etc.) as a share of GDP, agriculture as a share of GDP, life expectancy, land area, and population. Fourth, I test if there is a systematic variation in the impact of the VAT among early adopters and late adopters. Fifth, I control for the potential impact of the IMF loan programs by adding the participation

in the lower-middle-income, and 5 to 13 in the low-income group. The list of all 61 treated countries analyzed is presented in Section IV of Appendix S1.

26. In Figures 28 and 29 of Appendix S1, I present the results for capital stock and total factor productivity, respectively. They are also similar to the baseline results for capital stock and total factor productivity.

FIGURE 6
The Impact of VAT on GDP per Worker Using Conventional Difference-in-Differences Approach.



Notes: In the top row 54 countries used in baseline SCM are included, while the bottom row adds an additional 22 countries that were excluded in the baseline due to poor pretreatment fit. All specification includes country and year fixed effects, and standard errors are clustered at the country level. The figures in the right panel also include country-specific linear time trends. The connected line plots the estimates for leads to the left of the vertical line and the estimates for lags to the right. The variable year = −1 is omitted

in the loans program as dummy variables when constructing the synthetic controls. Sixth, I test whether the results are sensitive to the World Bank's income classification system. Seventh, I test whether the results are sensitive to the use of cross-validation technique used while constructing the synthetic controls. Eighth, I use various alternative measures of economic efficiency such as GDP per capita, expenditure-side GDP, and GDP data from PWT 8.0 rather than PWT 8.1. Ninth, I use tax as a share of GDP to control for overall tax burden while estimating the treatment effects using the SCM. Tenth, I implement leave-one-control-unit-out tests suggested in Abadie, Diamond, and Hainmueller (2015) to check the sensitivity of the baseline estimates to the inclusion of specific donor country in the construction of synthetic units. Eleventh, I test for the anticipation effects by assuming that the VAT was adopted 2 years before the actual implementation year. Twelfth, I use in-time placebo experiments by assuming that the VAT was adopted 5 years before the actual implementation year.

The results from these additional robustness tests are discussed in detail in Section II of Appendix S1. Although in a few of the robustness tests the effect sizes are slightly different, they never change the sign. Thus, I conclude that the baseline estimates are largely robust.

VIII. CONCLUSIONS

This paper analyzes the impact of replacing sales and turnover taxes with a VAT on economic efficiency using newly available PWT 8.1 data and the SCM recently developed in Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010, 2015). The procedure involves identifying the effects by comparing the trajectory of posttreatment GDP per worker with a carefully constructed counterfactual trajectory of GDP per worker. I find that a VAT adoption has an economically meaningful and statistically significant impact on the trajectory of real GDP per worker. However, the positive impact is conditional on the level of development of the country in question. That is, there is a strong correlation between the positive impact of a VAT adoption and the initial level of development, with the high-income countries benefiting the most from the reform and the low-income countries benefiting the least. I also analyze two potential mechanisms (capital accumulation and TFP) through which a VAT can affect GDP per worker and find that both the channels are important. I find the same pattern of differential effect of a VAT adoption along the initial income level for these channels.

This paper has two main policy implications. First, these results indicate that the theoretical advantages of a VAT do not necessarily translate into practice. In particular, the impact of the VAT on economic efficiency depends on the level of development of the country. The level of the development is highly correlated with factors such as tax capacity, tax evasion, and informal economy, which can severely undermine the effectiveness of a VAT. Second, a VAT is often considered a “silver bullet” that can both replace the distortionary taxes and raise much-needed revenue for public spending, especially in the developing countries. My study suggests that the results are more nuanced, highlighting the need to modernize tax administration along with VAT adoption, to benefit from the efficiency properties of a VAT. This policy implication is consistent with the voluminous literature on taxation and development arguing that the real tax system is not that which is passed as legislation, but that which is administered (Bird and Gendron 2011; Casanegra de Jantscher 1990; Gordon and Li 2009). Thus, it is critical to ensure that the implementation and administration of a VAT receive as much attention as the adoption of a VAT does, especially because my results indicate sizable gains in economic efficiency from adopting a well-designed and well-enforced VAT.

A. Treated Countries and Donor Countries

The list given below displays each of the treated countries (in boldface) and their respective synthetic controls. The relative contribution of each donor country with positive weights is shown in parenthesis.

Austria: Finland (0.413), Greece (0.35), and Switzerland (0.237).

Argentina: Ghana (0.579), Mali (0.108), Niger (0.138), and Portugal (0.175).

Bangladesh: Laos (0.783), Sierra Leone (0.192), and the United States (0.025).

Belgium: Finland (0.716), Jamaica (0.002), and Switzerland (0.282).

Canada: Botswana (0.179), Lesotho (0.002), and the United States (0.819).

Chile: Australia (0.07), Jamaica (0.853), Japan (0.021), New Zealand (0.032), and Philippines (0.025).

Colombia: Japan (0.287), Philippines (0.458), and South Africa (0.254).

Costa Rica: Canada (0.108), Jamaica (0.685), and Niger (0.206).

Denmark: Finland (0.558) and Switzerland (0.442).

Dominican Republic: Switzerland (.0212) and Tanzania (0.788).

France: Spain (0.705), Switzerland (0.08), Turkey (0.002), and the United States (0.213).

Greece: Ghana (0.153), Tanzania (0.345), and the United States (0.502).

Guinea: Laos (0.971) and the United States (0.029).

Honduras: Cote d’Ivoire (0.718), Jamaica (0.009), Jordan (0.21), and Paraguay (0.063).

Ireland: Canada (0.054), Jamaica (0.261), New Zealand (0.106), and Portugal (0.58).

Italy: Finland (0.898), Morocco (0.002), and Switzerland (0.099).

Jamaica: Lesotho (0.878) and the United States (0.122).

Japan: Jordan (0.631) and the United States (0.369).

Kenya: Central African Republic (0.625), Gambia (0.324), Jordan (0.028), and the United States (0.023).

Mauritius: Swaziland (0.84) and the United States (0.16).

Nepal: Laos (1).

Netherlands: Australia (0.259), Canada (0.12), Finland (0.51), Greece (0.077), and Switzerland (0.034).

New Zealand: Australia (0.582) and Ghana (0.418).

Pakistan: Botswana (0.144), Central African Republic (0.618), Gambia (0.178), and the United States (0.06).

Panama: Philippines (0.588), South Africa (0.37), and the United States (0.041).

Peru: Cote d'Ivoire (0.152), Ghana (0.111), Spain (0.112), Tanzania (0.529), and the United States (0.096).

Portugal: Ghana (0.252), Jordan (0.425), Tanzania (0.109), and the United States (0.214).

Senegal: Cote d'Ivoire (0.037), Ghana (0.019), Jamaica (0.041), Jordan (0.036), and Tanzania (.867).

Spain: Ghana (0.119), Jordan (0.398), and the United States (0.483).

Sweden: Australia (0.107), Canada (0.098), Finland (0.157), Japan (0.189), New Zealand (0.085), Switzerland (0.199), Turkey (0.128), and the United States (0.038).

Thailand: Lesotho (0.583), Swaziland (0.401), and the United States (0.017).

United Kingdom: Canada (0.163), Finland (0.728), Portugal (0.029), and Switzerland (0.079).

Uruguay: Jamaica (0.84) and the United States (0.16).

REFERENCES

Abadie, A., and J. Gardeazabal. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*, 93(1), 2003, 113–32.

Abadie, A., A. Diamond, and J. Hainmueller. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490), 2010, 493–505.

———. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*, 59(2), 2015, 495–510.

Acemoglu, D., S. Johnson, A. Kermani, J. Kwak, and T. Mitton. "The Value of Connections in Turbulent Times: Evidence from the United States." *Journal of Financial Economics*, 121(2), 2016, 368–91.

Adhikari, B., and J. Alm. "Evaluating the Economic Effects of Flat Tax Reforms Using Synthetic Control Methods." *Southern Economic Journal*, 83(2), 2016, 437–63.

Adhikari, B., R. Duval, B. Hu, and P. Loungani. "Can Reform Waves Turn the Tide? Some Case Studies Using the Synthetic Control Method." *Open Economies Review*, 29(4), 2018, 879–910.

Ahlerup, P., T. Baskaran, and A. Bigsten. "Tax Innovations and Public Revenues in Sub-Saharan Africa." *The Journal of Development Studies*, 51(6), 2015, 689–706.

Alm, J., and A. Embaye. "Using Dynamic Panel Methods to Estimate Shadow Economies Around the World, 1984–2006." *Public Finance Review*, 41(5), 2013, 510–43.

Athey, S., and G. W. Imbens. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives*, 31(2), 2017, 3–32.

Ballard, C. L., J. K. Scholz, and J. B. Shoven. "The Value-Added Tax: A General Equilibrium Look at its Efficiency and Incidence," in *The Effects of Taxation on Capital Accumulation*. University of Chicago Press, 1987, 445–80.

Barro, R. J. "Economic Growth in a Cross Section of Countries." *The Quarterly Journal of Economics*, 106(2), 1991, 407–43.

Besley, T., and A. Case. "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *The Quarterly Journal of Economics*, 110(3), 1995, 769–98.

Billmeier, A., and T. Nannicini. "Assessing Economic Liberalization Episodes: A Synthetic Control Approach." *The Review of Economics and Statistics*, (4), 2013, 95.

Bird, R., and P.-P. Gendron. *The VAT in Developing and Transitional Countries*. Cambridge Books, Cambridge University Press, 2011.

Boeters, S., C. Bohringer, T. Buttner, and M. Kraus. "Economic Effects of VAT Reforms in Germany." *Applied Economics*, 42(16–18), 2010, 2165–82.

Bye, B., B. Strom, and T. Avitsland. "Welfare Effects of VAT Reforms: A General Equilibrium Analysis." *International Tax and Public Finance*, 19(3), 2012, 368–92.

Campos, N. F., F. Coricelli, and L. Moretti. "Institutional Integration and Economic Growth in Europe." *Journal of Monetary Economics*, 103, 2019, 88–104.

Casanegra de Jantscher, M. "Value-Added Taxation in Developing Countries," in *Chapter Administering the VAT*, edited by M. Gillis, C. Shoup and G. P. Sicat. Washington, DC: World Bank, 1990, 171–90.

Case, A. C., H. S. Rosen, and J. J. Hines. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics*, 52(3), 1993, 285–307.

Cavallo, E., S. Galiani, I. Noy, and J. Pantano. "Catastrophic Natural Disasters and Economic Growth." *The Review of Economics and Statistics*, (5), 2013, 95.

Cizek, P., J. Lei, and J. E. Ligthart. "Do Neighbours Influence Value-Added-Tax Introduction? A Spatial Duration Analysis." *Oxford Bulletin of Economics and Statistics*, 79(1), 2017, 25–54.

Cnossen, S. "Global Trends and Issues in Value Added Taxation." *International Tax and Public Finance*, 5(3), 1998, 399–428.

Coolidge, J. "Findings of Tax Compliance Cost Surveys in Developing Countries." *eJournal of Tax Research*, 10(2), 2012, 250–87.

Diamond, P. A., and J. A. Mirrlees. "Optimal Taxation and Public Production: I—Production Efficiency." *American Economic Review*, 61(1), 1971, 8–27.

Ebrill, L., M. Keen, and J.-P. Bodin. *The Modern VAT*. Washington, DC: International Monetary Fund, 2001.

Emran, M. S., and J. E. Stiglitz. "On Selective Indirect Tax Reform in Developing Countries." *Journal of Public Economics*, 89(4), 2005, 599–623.

Feenstra, R. C., R. Inklaar, and M. P. Timmer. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10), 2015, 3150–82.

Ferede, E., and B. Dahlby. "The Impact of Tax Cuts on Economic Growth: Evidence from the Canadian Provinces." *National Tax Journal*, 65(3), 2012, 563–94.

Gordon, R., and W. Li. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." *Journal of Public Economics*, 93(7–8), 2009, 855–66.

Harrison, G., and R. Krelove. "VAT Refunds: A Review of Country Experience." IMF Working Paper, WP/05/218, 2005.

Hines, J. R. "Might Fundamental Tax Reform Increase Criminal Activity?" *Economica*, 71(283), 2004, 483–92.

Inklaar, R., and M. P. Timmer. *Capital, Labor and TFP in PWT8.0*. Groningen Growth and Development Centre, University of Groningen, 2013.

International Country Risk Guide. *IRIS-3: International Country Risk Guide, 1982–1997*. The Political Risk Services Group, 2011.
International Tax Dialogue. *The Value Added Tax: Experiences and Issues*. International Tax Dialogue, 2005.
James, K. *The Rise of the Value-Added Tax*. Cambridge University Press, 2015.
Keen, M., and B. Lockwood. "Is the VAT a Money Machine?" *National Tax Journal*, 59(4), 2006, 905–28.
———. "The Value Added Tax: Its Causes and Consequences." *Journal of Development Economics*, 92(2), 2010, 138–51.
King, G., and R. Nielsen. "Why Propensity Scores Should Not be Used for Matching." *Political Analysis*, 2019, 1–20.
Lee, D., D. Kim, and T. E. Borcherding. "Tax Structure and Government Spending: Does the Value-Added Tax Increase the Size of Government?" *National Tax Journal*, 66(3), 2013, 541–70.
Mankiw, N. G., D. Romer, and D. N. Weil. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics*, 107(2), 1992, 407–37.
Marshall, M. G., and K. Jaggers. *Polity IV Project: Politial Regime Characteristics and Transitions*, 2013, 1800–2012.
Meer, J., and J. West. "Effects of the Minimum Wage on Employment Dynamics." *Journal of Human Resources*, 51(2), 2016, 500–22.
Mikesell, J. L. "Misconceptions about Value-Added and Retail Sales Taxes: Are they Barriers to Sensible Tax Policy?" *Public Budgeting & Finance*, 34(2), 2014, 1–23.
Nannicini, T., and A. Billmeier. "Economies in Transition: How Important Is Trade Openness for Growth?" *Oxford Bulletin of Economics and Statistics*, 73(3), 2011, 287–314.
Nellor, D. "The Effect of the Value-Added Tax on the Tax Ratio." IMF Working Paper, 1987, 1–28.
Persson, T., and G. Tabellini. in *The Growth Effect of Democracy: Is it Heterogenous and how Can It Be Estimated?* edited by National Bureau of Economic Research Working Paper. 2007, 13150.
Piggott, J., and J. Whalley. "VAT Base Broadening, Self Supply, and the Informal Sector." *American Economic Review*, 91(4), 2001, 1084–94.
PricewaterhouseCoopers and the World Bank. 2017. "*Paying Taxes*."
Ring, R. J. Jr. "The Proportion of Consumers' and Producers' Goods in the General Sales Tax." *National Tax Journal*, 42(2), 1989, 167–79.
———. "Consumers' Share and Producers' Share of the General Sales Tax." *National Tax Journal*, 52(1), 1999, 79–90.
Riswold, S. "VAT in Sub-Saharan Africa - A Critique of IMF VAT Policy." *International VAT Monitor*, 15(2), 2004, 427–45.
Sachs, J. D., and A. Warner. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity*, 26(1), 1995, 1, 25th A–118.
Schneider, F. "Shadow Economies Around the World: What Do we Really Know?" *European Journal of Political Economy*, 21(3), 2005, 598–642.
Skinner, J., and J. Slemrod. "An Economic Perspective on Tax Evasion." *National Tax Journal*, 38(3), 1985, 345–53.
Smart, M., and R. M. Bird. "The Impact on Investment of Replacing a Retail Sales Tax with a Value-Added Tax: Evidence from the Canadian Experience." *National Tax Journal*, 62(4), 2009, 591–609.
Smith, B. "The Resource Curse Exorcised: Evidence From a Panel of Countries." *Journal of Development Economics*, 116(C), 2015, 57–73.
Tait, A. A. *Value Added Tax: International Practice and Problems*, Vol. 24. International Monetary Fund, 1988.
Ufier, A. "Quasi-Experimental Analysis on the Effects of Adoption of a Value Added Tax." *Economic Inquiry*, 52(4), 2014, 1364–79.
———. "The Effect of VATs on Government Balance Sheets." *International Tax and Public Finance*, 24(6), 2017, 1141–73.
Wacziarg, R., and K. H. Welch. "Trade Liberalization and Growth: New Evidence." *World Bank Economic Review*, 22(2), 2008, 187–231.
Wolfers, J. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review*, 96(5), 2006, 1802–20.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1 Does Introducing a Value Added Tax Increase Economic Efficiency?