When Does Introducing a Value-Added Tax Increase Economic Efficiency? Evidence from the Synthetic Control Method

Job Market Paper
November 14, 2015

Bibek Adhikari *

Abstract

The theoretical prediction that a value-added tax (VAT) does not distort firms' production decisions has led to its rapid adoption worldwide, but there is surprisingly little empirical evidence. This paper provides one of the first causal estimates of the efficiency gains (i.e., an increase in GDP per worker) of introducing a VAT in a worldwide sample of countries using the synthetic control method. The synthetic control is a weighted average of countries without a VAT that closely resembles the economic structure and outcomes of the country with a VAT for several years before the adoption of a VAT. In line with previous studies, I find that the VAT has, on average, a positive and economically meaningful impact on economic efficiency. However, this result is driven by richer countries only. There is no significant impact of the VAT on poorer countries. I find similar results when estimating the impact of the VAT on total factor productivity and capital stock per worker, two important channels through which a VAT affects GDP per worker. This paper provides evidence that a success of VAT almost entirely depends on the initial level of income of a country, which, in turn, determines whether a country is able to properly design and enforce a VAT. The findings are robust across a series of placebo studies and sensitivity checks.

Keywords: Value-Added Tax, Economic Efficiency, Synthetic Control Method.
JEL Codes: H20, H25, O40, E6.

*Department of Economics, 206 Tilton Hall, Tulane University, New Orleans, LA 70118 (*email:* badhikar@tulane.edu *web:* www.bibekadhikari.com). I am grateful to James Alm, Stefano Barbieri, Alan Barreca, Romain Duval, John Edwards, Keith Finlay, Sean Higgins, Michael Keen, Jaewoo Lee, Mario Mansour, Ruud de Mooij, Whitney Ruble, Steven Sheffrin, Jay Shimshack, and Dan Teles for helpful comments and discussions. Earlier versions of this paper were presented at the 2014 Annual Meetings of the National Tax Association, 2014 and 2015 Annual Meetings of the Southern Economic Association, Jobs, Growth, and Structural Reforms seminar series at the International Monetary Fund, and economics department seminar series at Tulane University. I am also thankful to the participants of these conferences and seminars for their thoughtful comments. 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 parallel processing of the C++ codes.

1 Introduction

The value-added tax (VAT) is a general tax on consumption that is levied on the difference between the value of a firm's output and its purchased inputs. In its modern form, the VAT was first introduced in Brazil and Denmark in 1967, making it one of the youngest, yet among the most important sources of government revenue. As of 2015, the VAT has been adopted by more than 160 countries and its revenues represent more than seven percent of the world's GDP. These numbers are expected to continue to grow because of increasing reliance on consumption taxes and because of the requirement by many multi-lateral 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 instrument.[1] Under a VAT, tax is charged at all stages of production with some mechanism to provide full credit or refund to firms of the VAT that has been charged on their input costs. Thus, a VAT leaves production undistorted as it does not affect the prices firms ultimately pay for inputs and it does not create cascading - the tax-on-tax as a good passes from one production stage to another. The theoretical justification for an efficient tax system comes from one of the fundamental results in public economics, the production efficiency theorem of Diamond and Mirrlees (1971), which states that tax systems should be designed such that they do not distort firms' production decisions even in second-best environments.

The VAT's efficiency, however, is based on the assumption that it is well-designed and well-enforced, but there are many settings under which a VAT can be inefficient. 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. Exempt firms do not qualify for the refunds on their input costs, which affect the prices they pay for inputs and can introduce production inefficiencies. Second, the production efficiency theorem assumes zero tax evasion, but tax evasion is widespread, especially in developing and transitioning countries (Alm and Embaye, 2013; Schneider, 2005). The presence of tax evasion and the informal economy also distort behavior, which introduces inefficiencies in the VAT system (Emran and Stiglitz, 2005; Piggott and Whalley, 2001). Furthermore, there is some concern that a VAT increases tax evasion and informality 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 (Hines, 2004; Coolidge, 2012). Thus, whether the VAT is an efficient tax instrument is ultimately an empirical question.

This paper claims that the question - Is a VAT an efficient tax instrument - is incomplete.[2]

[1] The other reasons are its high revenue efficiency, its compatibility with outward-oriented economic policy, and its relative simplicity from an administrative perspective (Martinez-Vazquez and Bird, 2011).

[2] Similar claims have been made by many including Bird and Gendron (2011), and Casanegra de Jantscher

Although research has been done to test whether the VAT is as promising in the real world as it is in theory, these studies have ignored the heterogeneous environments under which the VATs are adopted. They include a country's stage of development, tax capacity, tax evasion, and informal economy, as well as the design and enforcement of the VAT. As a result, when it comes to the question of the VAT's efficiency, it is difficult to come to a singular 'yes' or 'no' conclusion when factors determining it are multiple. So perhaps, the real question is - in what conditions should a VAT be adopted for it to be an efficient tax system. This paper is the first to ask if richer countries benefit more from a VAT than poorer countries. In particular, this paper is one of the first to estimate the causal impact of introducing a VAT using a direct measure of economic efficiency (i.e., GDP per worker) and the first paper to investigate two primary channels through which a VAT affects GDP per worker: by increasing total factor productivity and by increasing the stock of capital.[3]

In addition to having limitations, previous study on empirical analysis of the efficiency effects of a VAT is surprisingly sparse as well, largely due to the unavailability of comparable data across countries on economic efficiency. My primary data source is Penn World Table 8.1, which, for the first time, includes comparable cross-country data on output-side real GDP, total factor productivity, stock of physical capital, and stock of human capital (Feenstra, Inklaar and Timmer, 2015).[4] I combine PWT 8.1 with various data sources, such as proprietary data on the manufacturing industry from United Nations Industrial Development Organization (UNIDO), economic and demographic variables from the World Bank's World Development Indicators, and institutional variables from the Polity IV database. Finally, I collect data on the properties of the VAT at time of its introduction from various issues of International VAT Monitor, Tax Notes International, International Monetary Fund's country documents, and tax authority websites.

Even when data availability is not a problem, estimating the causal effect of VAT adoption on economic efficiency using traditional approaches faces several challenges. First, countries choose both whether to adopt a VAT and when to adopt a VAT. Thus, cross-country regressions are subject to numerous biases and are unlikely to reveal the causal 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 VAT adoption and compare countries with a VAT to those countries without a VAT, but with similar propensity to adopt a VAT. However,

(1990), but they do not go beyond descriptive analysis to substantiate these claims.

[3]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.

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

matching techniques construct the counterfactual by matching on a single index number and, importantly, they cannot account for the biases caused by unobserved confounders. When the biases by unobservables are of concern, difference-in-differences (DID) technique can be used. This approach allows one to control for both country and year fixed effects so that all time-invariant differences across countries (such as geography) and changes common to all countries in the same year can be controlled for. However, identifying the effects of a VAT using DID requires that the paths of the outcome variable between the VAT adopters and non-VAT adopters be parallel. As I demonstrate later, this assumption is not satisfied (see Figure 17). Another important limitation of the DID is that it assumes the impact of the VAT on economic efficiency is homogenous across reforming countries. Thus, it does not allow us to explore whether the impacts of the VAT systematically varies across countries according to their capacity to properly design and implement it. Finally, DID cannot account for the biases caused by time-varying unobservables.

I make progress in addressing all of these challenges. To account for the endogeneity caused by time-varying unobserved confounders, I use the synthetic control method (SCM), 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 (Abadie and Gardeazabal, 2003; Abadie, Diamond and Hainmueller, 2010, 2015). The SCM calculates the dynamic effects of VAT reform on a case-by-case basis, which allows for the effect of the reform to be 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 time of VAT's introduction.[5] In addition, I also contribute to the improvement of the synthetic control method by developing a version of leave-one-out test, which is an intuitive graphical approach of testing whether the average effects are driven by outlier countries. This test will be useful to researchers applying the SCM to multiple treated units and studying the aggregate impact of the treatment.

The results of this research provide a new insight to the topic of 'VAT and efficiency'. I find strong evidence of efficiency gains of VAT adoption, but only in richer countries. For instance, five years after the reform GDP per worker of the high-income group is 10.9 percent higher than the synthetic group, which increases to 11.2 percent ten years after the

[5]The World Bank's country classification starts as early as 1971 for the low-income countries and 1975 for the lower-middle-income countries, but the classification for the upper-middle-income and the high-income countries start from 1987 only, therefore I use the information from the classification of the low-income countries to impute the country classification for years before 1987. For details about the imputation see Appendix 1.

reform (all significant at the 1 percent level). GDP per worker of the upper-middle-income group is also positive and significant at the 5 percent or better level, but after some time lag. By five year after the reform GDP per worker is 25.5 percent higher compared to the synthetic group, which increases to 33.1 percent 10 years after the reform. However, I find no significant impact of VAT on poorer countries. The effect of VAT on the lower-middle-income countries are slightly positive and the effect on the low-income countries are mostly negative, but these estimates are statistically indistinguishable from zero at the 10 percent level. The same differential trend in the impact of VAT is found across income groups for total factor productivity and capital stock per worker, two of the primary channels thorough which VAT affects GDP per worker. These results indicate that the theoretical advantages of VAT do not necessarily translate into practice.

To probe the robustness of these results, I perform a series of additional tests. In some cases VATs were adopted as part of the trade liberalization, World Trade Organization (WTO) membership or in anticipation of joining the European Union (EU). Thus, it is important to test if their impacts are commingled. To that end, I report the average treatment effect calculated by removing the countries that liberalized their economy, joined the WTO or joined the EU in the sample period. My main results remain unchanged. Another concern is that the estimates may be biased due to reverse causation. If VAT reforms are motivated by expectation of future growth prospects, this would bias the estimates obtained from the SCM as long as growth expectations are not captured by the unobservable heterogeneity included in the estimation. To mitigate the concern of reverse causality, I supplement the study with data from the manufacturing industry. The rationale being that the decision to adopt a VAT takes place at the national level and thus they are not likely to be affected by one particular sector of the economy in any significant way. Again, the results are similar to the main results. Likewise, the results are similar when estimated using conventional difference-in-differences methods. I also conduct various other sensitivity checks and placebo studies and the estimates are robust to these tests.

The rest of the paper is organized as follows. Section 2 assesses key advantages and disadvantages of the VAT, provides context of its adoption, describes the features of the VATs that have been adopted, and provides a brief review of the literature. Section 3 lays out the conceptual framework to evaluate the efficiency gains of VAT adoption and discusses the data used for the estimation. Section 4 discusses the research design used to isolate the impact of the VAT from various other confounding factors and discusses the sample selection procedure. Section 5 presents the main results, section 6 presents results for potential mechanisms, and section 7 presents various robustness tests. Section 8 concludes with some policy implications.

2 Background

2.1 Assessing Key Properties of the VAT

There are various properties of a VAT that make it an attractive tool for raising revenue. First, by not taxing intermediate transactions, it does not distort firms' production decisions (unlike sales and turnover taxes). To be more specific, as long as some of the taxed inputs can be substituted by untaxed inputs, any taxes on intermediate transactions will drive a wedge between the buying and selling prices of producers, violating production efficiency. Since a VAT is levied on final consumption and not on intermediate transactions between firms, it does not create a wedge between the prices that producers face in buying and selling from one another, thus, not violating production efficiency (Ebrill, Keen and Bodin, 2001; Keen and Lockwood, 2010; Tait, 1988).

Second, by fully refunding the taxes on inputs, the VAT prevents cascading - a phenomenon where a tax-on-tax occurs as a good passes from one production stage to another (for instance, from manufacturing to wholesale to retail). Cascading creates distortions in real production decisions by providing firms an incentive to vertically integrate to reduce the number of stages that are subject to the tax, implying efficiency loss. An alternative to the VAT that can avoid cascading is taxing only one stage of production, i.e., levy a tax only at manufacturing stage, wholesale stage, or retail stage. However, cascading is unavoidable even in single stage taxation. For instance, many producers buy inputs from manufacturers and wholesalers, or even from big box retailers. Thus, the lack of a mechanism to refund taxes on inputs implies cascading even in a single stage sales tax, with the probability (and magnitude) of cascading greater, the further it is from the retail stage. For instance, Ring (1989, 1999), and Mikesell (2014) estimate that about 40 percent of the retail sales tax revenue in the U.S. comes from business inputs and Smart and Bird (2009) estimate that about 43 percent of the retail sales tax revenue in Canada comes from business inputs.[6]

Third, theory predicts that a consumption tax, which encourages savings and investment, is superior to an income tax in terms of its growth effects (Jones and Manuelli, 1990; Lucas, 1988; Rebelo, 1991; Oakland, 1967*a*) and many empirical studies support this prediction (Alm and El-Ganainy, 2013; Barro, 1991; Kneller, Bleaney and Gemmell, 1999; Lee and Gordon, 2005). As Metcalf (1995) notes, VATs can be considered as a replacement for income taxes in the sense that most countries that adopt a VAT collect a significantly smaller

[6] A VAT is also superior to cascading taxes from a tax incidence perspective since cascading results in a pattern of effective tax rates that depend on arbitrary things such as the number of taxable production stages, making effective taxes opaque and calculation of tax incidence difficult (Ebrill, Keen and Bodin, 2001; Oakland, 1967*b*).

percent of tax revenues from personal and corporate income taxes. Another advantage of the VAT over an income tax is that the VAT does not influence the forms or methods of doing business since tax liability under a VAT remains the same whether the product is made in the corporate, or the non-corporate sector or, whether capital-intensive, or labor-intensive technology is used in production (Cnossen, 1998).

Fourth, leakages in revenue are less likely under the VAT than under sales taxes because revenue is collected in all stages of production under the VAT, but only at the final stage under the retail sales tax. The retail stage is also considered to be the weakest link in the chain to collect revenue. Furthermore, the invoice driven VAT system generates a third-party reported paper trail that facilitates enforcement and makes tax evasion more difficult (Ebrill, Keen and Bodin, 2001; Pomeranz, 2015).

The VAT, however, is not without its weaknesses. Exempting some sectors from the VAT or the presence of an informal economy introduces some inefficiency, as they would create incentives for firms to switch to exempt sector or the informal sector to avoid the VAT. Further, the credit and refund mechanism of the VAT creates opportunities for fraud that are unique to a VAT. Some of the examples are false claims for credit or refund, credit claimed for a VAT on purchases that are not creditable, and the creation of fake invoices so that credit can be claimed (Keen and Smith, 2006).

There are two major criticisms of the VAT. The first criticism, ironically, is the perceived efficiency of the VAT. Critics argue that politicians exploit the lower marginal cost of raising revenue through a VAT to raise more revenue than is economically optimal, leading to an inefficiently large level of government spending (Brennan and Buchanan, 1977). For instance, the 2005 presidential panel on tax reform in the United States could not reach a consensus on whether to recommend a VAT in the U.S., primarily because some members were concerned that introducing a VAT would lead to higher total tax collections over time, and would therefore facilitate a larger federal government - in other words, that the VAT would be a "money machine." (The President's Advisory Panel on Federal Tax Reform, 2005).

The second criticism of the VAT is that it is a regressive tax system. Critics argue that any consumption tax will impose a heavier burden on the low-income families than on the high-income families because the fraction of annual income spent on consumption tends to be higher for the low-income than for the high-income families. Also, if a higher reliance on the consumption tax leads to less reliance on the income tax, it means that there will be less redistribution via the income tax. Similarly, the consumption tax exempts capital and capital income from the tax base, which reduces the tax burden on the high-income families where such income tends to dominate, while increasing the burden on the low-income families where income from salary and wages tends to dominate.

2.2 The Context of VAT Reforms

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

First, a 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 non-academic sources highlight the role of multi-lateral 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 help has a much higher probability of adopting a VAT. For instance, Keen and Lockwood (2010), Ufier (2014), and Cizek, Lei and Ligthart (2012) find that the probability of a VAT adoption increases by up to 25 percent in the year following country's participation in IMF program. Ebrill, Keen and Bodin (2001) estimate that more than half of all countries that introduced a VAT during 1980s and 1990s used IMF advice in doing so.[7]

Second, as Case, Rosen and Hines (1993), and Besley and Case (1995) demonstrate, the design 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 again be more exogenous to the economic condition of the adopting country than a reform motivated by their internal economic conditions. Figure 1 documents the spread of the VAT, and it clearly suggests that the VAT was adopted in regional waves. For instance, more than 11 other European countries adopted a VAT between 1967 when it was first introduced and 1973. 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 five years of Hungary's VAT adoption in 1988. A similar pattern is also found in the developing countries of the

[7] 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 multi-lateral organizations such as the World Bank (e.g., in the cases of Cameroon, Ghana, Hungary, Philippines, Thailand, Tanzania, and Venezuela), the World Trade Organization, Inter-American Development Bank, Asian Development Bank, the Organization for Economic Co-operation and Development, as well as aid and development agencies such as 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).

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 (2012).

2.3 Features of the VATs Adopted and the Environment Under which they were Adopted

In this section, I describe the properties of the VATs adopted and compare them with the text-book ideal VAT. I also provide the information on the presence of the informal economic activity and overall government quality.[8] I categorize 33 countries that are analyzed in this study according to their income level in the year the VAT was adopted into four income groups. I do this to highlight differences in the design of the VAT and the economic and political environment among the income groups.

Tables 1 and 2 provide the brief description of the characteristics of the VATs at the time of adoption. Almost of all the 33 countries 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 tariffs. The design of the VAT, however, is very heterogeneous, with high-income countries designing the VAT closest the text-book ideal and the low-income countries designing the VAT furthest from the text-book 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 5 out of 6 upper-middle-income countries, 6 out of 9 lower-middle-income countries, and 1 out of 5 low-income country had such broad-based VAT reform.

Table 3 provides the brief description of economic and political environment across 4 income groups. I find systematic differences in the environment that is conducive to effective enforcement of the VAT, with the best environment in high-income countries and the worst environment in the low-income countries. For instance, the share of economic activity in the informal sector is 16.5 percent in the high-income group, which increases to 28.2 percent in the upper-middle-income group, 42.3 percent in the lower-middle-income group, and 36 percent in the low-income group. To proxy the capacity to well-enforce a VAT, I use the overall government quality ranging from 0 to 1, which is constructed by taking the average of corruption, bureaucratic quality, and rule of law indices. I, again, find systematic differences

[8] These information are not readily available so I collect them from various sources. I start with Tait (1988) and Cnossen (1998) and supplement them by 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 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).

across income groups. To elaborate, overall government quality in the high-income group is 0.93, which decreases to 0.62 in the upper-middle-income group, 0.48 in the lower-middle-income group, and 0.40 in the low-income group. Another proxy for enforcement strength is given by *polity2* score, which ranges from -1 denoting autocracy to 1 denoting democracy and it also exhibits similar differential trends across income groups.

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 110 countries, I find that average government quality ranges from 0.92 in the high-income group to 0.44 in the low-income group, average *polity2* score ranges from 0.91 in high-income group to -0.02 in the low-income group, and average shadow economy ranges from 16.8 percent in the high-income group to 38.8 percent in the low-income group.

To conclude, there are significant differences in the characteristics of the VAT adopted and the environment necessary for its proper implementation and enforcement across countries in different income groups at the introduction of the VAT. Thus, the question still remains: do all countries, irrespective of their initial income level, experience efficiency gains after introducing a VAT?

2.4 Related Literature

There are many applied general equilibrium models evaluating the economic effects of VAT adoptions or of some changes in the structure of the VAT. The main advantage of this methodology is that it provides a clear bridge between theoretical and applied aspect of tax policy analysis. These papers use data from a particular country to calibrate their model. For instance, Boeters et al. (2010) use data from Germany to evaluate the impact of a VAT simplification reform on economic efficiency, Ballard, Scholz and Shoven (1987) use data from U.S. to examine the impact of hypothetical VAT adoption on economic efficiency, Bye, Strom and Avitsland (2012) use data from Norway to estimate efficiency effects of various VAT base-broadening scenarios, and Bovenberg (1987) uses data from Thailand to evaluate the impact of a VAT on economic efficiency. In general, these studies find positive impact of VAT on economic efficiency and other macroeconomic variables. One thing common across all these studies is that they assume there is no informal economy. Piggott and Whalley (2001) use data from Canada where they model the presence of informal economy and assume that a VAT encourages some firms to switch to the informal sector and some individuals to increase home production, both of which are plausibly inefficient compared to the formal sector. Piggott and Whalley (2001) find that VAT base-broadening reduces economic efficiency. Similarly, Bye, Strom and Avitsland (2012) find that including selective services in the VAT base reduces welfare compared to both not including services in the VAT base and including all services in the VAT base. They highlight the role of informal economy and a narrow VAT base in undermining the effectiveness of the VAT, ignoring which can lead to an upward bias of the estimated results.

The applied general equilibrium models of the VAT reforms 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. This methodology also suffers from the lack of statistical test to confirm the validity of model specifications. Finally, almost all of these are single-country studies focusing on the high-income countries only (Bovenberg, 1987; Shoven and Whalley, 1984; Andre, Cardenete and Romero, 2010).

In contrast, there are few reduced-form studies on the economic effects of the VAT. Nellor (1987) is one of the earliest attempts to empirically study the impact of a VAT on economic efficiency. Nellor (1987) analyzes 11 European countries that introduced a VAT in the 1960s and 1970s using cross-country regression and he finds that introduction of a VAT raised the revenue ratio in those countries. As demonstrated by Keen and Lockwood (2010), under some weak conditions access to a more efficient tax instrument leads an optimizing government to increase the revenue ratio. In two recent studies, Keen and Lockwood (2006) use a panel of Organization for Economic Co-operation and Development (OECD) countries and Keen and Lockwood (2010) use an unbalanced panel of 143 countries over 26 years. Both estimate the impact of the VAT on the revenue ratio using fixed effects regression, and find that a VAT adoption is associated with an increase in economic efficiency. Similarly, Smart and Bird (2009) and Ferede and Dahlby (2012) use a DID research design and province-level data from Canada, and find that investment and economic growth respond positively when a VAT replaces a sales tax. More recently, Ufier (2014) studies the effect of the VAT on various macroeconomic variables such as growth, investment, trade, inflation, and government size using propensity score matching. He finds evidence that the VAT increases investment, and he finds weak evidence that it increases growth.

Most of the studies mentioned above use standard cross-country regression methodology that do not attempt to identify the causal link between a VAT adoption and economic efficiency. Smart and Bird (2009) and Ferede and Dahlby (2012) are the only studies that attempt to establish causality while also controlling for the biases caused by time-invariant unobservables, but they only analyze the impact of province-level VAT in Canada. My study is more similar to Ufier (2014) in that both of us account for the selection of a country into a VAT. However, my paper differs from Ufier (2014) both in terms of the methodology used and the topics analyzed. In terms of the methodology, I use the SCM which can flexibly control for both observed and unobserved confounders unlike propensity score matching which can only account for observed confounders. Second, the SCM constructs the counterfactual by matching the level and the trend of the outcome variable (10 years in this study) unlike propensity score matching that constructs the counterfactual by matching on a single index number. Thus, the SCM produces more credible counterfactual. In terms of the topics analyzed, while both of us analyze the impact of VAT on economic efficiency, none of the previous studies including Ufier (2014) empirically estimate the impact on productivity and capital stock. Moreover, rather than estimating the average impact of the VAT across all countries, I allow for the impact of the VAT to vary across countries. This allows me to provide new insights: that the impact of a VAT depends on the initial level of income of a country, which can lead to very different policy implications.

3 Conceptual Framework and Data

3.1 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 total factor productivity (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 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 the production of some outputs can be increased 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 sales tax or a turnover tax is replaced with a relatively non-distortionary tax such as a VAT, a profit maximizing firm will have an incentive to allocate its resources to their most productive uses, which increases A, and thus leads to higher output. Second, as investment goods are not taxed under a VAT, a move from sales tax or 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.

3.2 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 Penn World Table (PWT) version 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 such as reintroducing measures of capital stock and introducing for the first time measures of human capital, the share of labor income in GDP, and total factor productivity.[9] This implies that 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 stocks. Prior to PWT 8.1, the data on TFP was only available for a small sample of countries, with most of them being the high-income countries because the construction of TFP is very difficult as it requires comparable data on output, labor stock, and capital stock. Even so, TFP data prior to PWT 8.1 came from various different databases, making comparisons of countries from different income groups problematic. Furthermore, PWT 8.1 uses an improved methodology to compute these series, making PWT 8.1 the best database for the study.

The outcome of interest is output-side real GDP per worker, which is calculated by dividing output-side real GDP by number of persons engaged, both of which are from the PWT 8.1. 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.[10] However, I am interested in measuring relative economic efficiency and not the relative living standard of the countries. Hence, output-side GDP per worker is an appropriate measure (Inklaar and Timmer, 2013).

The main explanatory variables are stock of physical capital per worker (k), and stock of human capital per worker (h) and they are obtained from the PWT.[11] To anchor the paper to

[9]PWT 5.6 had data on physical capital stocks up to 1992 and was based on older basic data using a single depreciation rate.

[10]For instance, small open countries such as Singapore and Luxembourg enjoy a much higher standard of living than what their productive capacities allow, due to favorable terms of trade.

[11]For the construction of capital stock, the PWT uses the perpetual inventory method. Unlike previous databases that assume constant depreciation rate for all assets, PWT divides assets into six categories (structures, transport equipment, computers, communication equipment, software, and other machinery and assets) and allows the depreciation rate to vary across each category. This adjusts for differences in the composition of investment across countries and over time, thus producing tighter estimates of capital stock. The PWT calculates the average stock of human capital as a function of the average years of schooling for the population aged 15 and older and the rates of return of schooling where the return to schooling is given

the endogenous growth model, I include 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, which contains unbalanced panel data on the types and qualities of government of 167 countries spanning 1800 to 2012 (Marshall and Jaggers, 2013). Following Persson and Tabellini (2007), I classify a country as democratic if the *polity2* 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 VAT was introduced, is obtained primarily from Ebrill, Keen and Bodin (2001) and supplemented by International Tax Dialogue (2005) and Ufier (2014).

4 Empirical Strategy

I employ the synthetic control method (SCM) to isolate the impact of a VAT adoption from other influences. The SCM is a data-driven way of finding the most appropriate counterfactual in generalized difference-in-differences (DID) estimation. The DID estimation consists of identifying a specific treatment, and then comparing the difference in outcomes before and after the treatment for the treated country to the difference in outcomes before and after the treatment for the untreated countries. The primary motivation to use the SCM is the belief that the effect of a particular intervention can be empirically assessed only by comparison with the appropriate counterfactual. The SCM was developed by Abadie and Gardeazabal (2003) and expanded by Abadie, Diamond and Hainmueller (2010, 2015). Abadie and Gardeazabal (2003) use the approach to quantify the impact of political violence on economic growth in the Basque region of Spain, Abadie, Diamond and Hainmueller (2010) use it to estimate the impact of a large anti-tobacco initiative on the per capita sales of cigarettes in California, and Abadie, Diamond and Hainmueller (2010) use it to assess the economic effects of the 1990 German reunification in West Germany. The use of the SCM has been growing, and it is now being applied to study a very diverse set of topics. [12]

by a piece-wise linear function to account for the fact that early years of education have a higher return (i.e., higher wages) than the later years. For further details see Inklaar and Timmer (2013) and Feenstra, Inklaar and Timmer (2015).

[12] These applications include: the effects of trade liberalization on economic growth (Nannicini and Billmeier, 2011; Billmeier and Nannicini, 2013), the impact of the terrorist attacks on electoral outcome (Montalvo, 2011), the effects of relaxing restrictions on home equity lending on retail spending by households (Abdallah and Lastrapes, 2012), the impact of natural disasters on economic growth (Cavallo et al.,

There are numerous advantages of using the SCM in my setting. First, the SCM allows for the effect of the reform to be heterogeneous across both countries and over time, as opposed to the DID estimator where only average effects can be analyzed. The average effects can mask large differences across countries. This is especially important for my purposes because the broader environment under which the VAT is adopted, such as the stage of development of a country, design of the VAT, and its implementation vary significantly across countries, thus, the assumption that the impacts across countries are homogenous might not be valid. Indeed, I find very heterogeneous impacts of the reforms.

Second, the SCM can substantially reduce any potential endogeneity problem caused by the omitted variables. Abadie, Diamond and Hainmueller (2010) prove that if a synthetic country can be found such that it matches the pre-treatment 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 post-treatment outcome variable for the treated and the synthetic control countries goes towards zero as the pre-intervention period increases. The intuitive explanation is that only countries that are alike in both observed and unobserved predictors of the outcome variable as well as in the effect of those predictors on the outcome variable should produce similar trajectories of the outcome variable over extended periods of time.

In comparative case studies, the researcher is allowed to choose the comparison group that resembles the counterfactual. However, Abadie, Diamond and Hainmueller (2010) argue that this introduces substantive ambiguity about how comparison groups are chosen, since researchers select comparison groups based on subjective measures of affinity between the treated country and the untreated countries. A third advantage is that the SCM resolves this problem by using a data-driven method to create a comparison unit that closely resembles the fundamentals of the treated unit in the pre-treatment period, using a weighted average of all other control units. The weights are chosen so that the pre-treatment outcomes and the covariates of the synthetic control unit match on average the outcomes and covariates of the treated unit. By taking the weighted average of control countries, the SCM makes explicit the contribution of each comparison unit to the counterfactual of interest. This provides better transparency. Contrast this with a regression technique like conventional DID where the counterfactual is constructed using the regression-weighted average of all control countries. However, these weights are implicitly calculated and are not reported in

2013), the effect of civil conflict on economic growth (Dorsett, 2013), the impact of nutrition policies on dietary behavior and childhood obesity (Bauhoff, 2013), the effect of immigration laws on demographic composition (Bohn, Lofstrom and Raphael, 2014), the impact of decrease in police enforcement on traffic fatalities and injuries (DeAngelo and Hansen, 2014), and the impact of major natural resource discoveries on economic growth (Smith, 2015).

practice, which makes them opaque. Moreover, since the weights in the SCM are restricted to be positive and sum to one, it safeguards against extrapolation bias that plagues traditional regression models. Similarly, the weighted average of a few similar countries can provide a more credible counterfactual than the regression-weighted average of all control countries.

4.1 Synthetic Control Method

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 pre-treatment periods and T_{0+1} denotes the treatment year. The observed outcome variable can be written as:

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

where $\tau_{it} = Y_{it}^{R} - Y_{it}^{NR}$ 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 the 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:

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

where α_t is an unknown common factor with constant factor loadings across countries, Z_i is a vector of observed covariates with coefficients θ_t, μ_i is a vector of unknown parameters, λ_t is a vector of unobserved common factors, and ϵ_{it} are the idiosyncratic error terms with zero means. $D_i(R_i = R_j)$ is a vector of indicator functions that become 1 when the treated country and the control country belongs to the same geographic region, and $D_i(I_i = I_j)$ is a vector of indicator functions that become 1 when the treated country and the control country belongs to the same income group. Note that this specification allows country effects to vary with time ($\lambda_t \mu_i$), unlike in difference-in-differences that restricts country effects to be time invariant ($\lambda \mu_i$).

Let us define a synthetic control unit as a weighted average of the units in the donor pool. That is, a synthetic control can be represented by a $(J \times 1)$ vector of weights, $W = (w_2, \ldots, w_{J+1})'$ such that $w_j \geq 0$ for $j = 2, \ldots, J+1$ and $w_2 + \cdots + w_{J+1} = 1$, where vector

W represents a potential synthetic control. Then the outcome variable for each potential synthetic control unit is given by $\sum_{j=2}^{J+1} w_j Y_{jt}$. Abadie, Diamond and Hainmueller (2010) prove that the size of the bias in the difference between the post-treatment outcome variable for the treated and the synthetic country is close to zero if the pre-treatment fit is strong and the number of pre-intervention periods is large relative to the scale of the transitory shocks. Thus, the unbiased dynamic treatment effect for time t (τ_{1t}), where $t \in \{T_{0+1}, \dots, T\}$ can be estimated by $\hat{\tau}_{1t}$ where:

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

To find the vector of country weights, w^*, let X_1 be $(k \times 1)$ vector containing the values of the pre-treatment variables of the treated country that I aim to match as closely as possible. Let X_0 be the $k \times J$ matrix where each column of the matrix is a vector of the same pre-treatment variables for each potential donor country. Then, the differences between the pre-treatment characteristics of the treated unit and its synthetic control is given by $\|X_1 - X_0 W\|$. The synthetic control algorithm chooses W^* to minimize the distance, $\|X_1 - X_0 W\| V = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$, where V is a diagonal matrix that assigns weights to the covariates in accordance to their predictive power on the outcome variable. The optimal V assigns weights to linear combinations of the variables to minimize the root mean square prediction error (RMSPE) for the pre-treatment period.[13]

I use a cross-validation method to choose the matrix of predictors weight V. If all available pre-treatment data is used to both choose matrices V and W and to minimize the RMSPE then the algorithm can make a more accurate prediction on the data it has seen (i.e., pre-treatment data), but it may perform poorly on the data it has not seen (i.e., post-treatment data). This phenomenon is called over-fitting in the statistical literature. It is crucial to insure that the algorithm performs well in the post-treatment data, since I am interested in constructing the counterfactual trajectory of GDP per capita such that it can accurately reflect how the trajectory of GDP per capita would evolve in the absence of the treatment. To avoid over-fitting, I use a cross-validation technique suggested by Abadie, Diamond and Hainmueller (2015) to choose the matrix of weights V that minimizes out-of-sample prediction errors. This is achieved by splitting the pre-treatment data into two halves and using the first half of the data for training and the second half for validation.[14]

[13] The RMSPE measures lack of fit between the trajectory of the outcome variable of the country of interest and its synthetic control. The RMSPE is defined as: $\sqrt{\frac{1}{T_0} \sum_1^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}) \right)^2}$

[14] To elaborate, I use the average values of the predictors from the training period to select the predictor weights V such that the resulting synthetic unit minimizes RMSPE over the validation period. Next, I use

To assess whether the comparison country created using the SCM is a good counterfactual, some measure of how well it resembles the treated country before the treatment is needed. Abadie, Diamond and Hainmueller (2010) use RMSPE of the outcome variable to measure fit or lack of fit between the paths of the outcome variable for treated country and its synthetic counterpart. Following Adhikari and Alm (2015), I use the pre-treatment fit index to assess the overall quality of the pre-treatment fit. A pre-treatment 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 pre-treatment period.[15]

Next, I aggregate the country-specific treatment effects into four groups based on income classification of the country during the year of VAT adoption.[16] 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:

$$\text{DTE}_t^G = \frac{\sum_{g=1}^{G} \hat{\tau}_{1t}^g}{G} = \frac{\sum_{g=1}^{G} Y_{1t}^g - \sum_{j=2}^{J_{g+1}} w_j^{g*} Y_{jt}^g}{G} \tag{5}$$

4.2 Inference

I use placebo experiments to evaluate the significance of the treatment effects. The essence of the placebo experiments is to test whether the estimated impact of the VAT adoption could be driven entirely by chance. Specifically, 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 VAT reform in the same year as the country of interest and running the synthetic control method. This iterative procedure provides a distribution of estimated placebo treatment effects for the countries where no intervention took place. If the placebo experiments create enough placebo treatment effects of magnitude greater than the one estimated for the treated country, then I conclude that there is no statistically significant evidence of an effect of a VAT reform in the treated country. If the placebo experiments show that the treatment effect estimated for the treated country

V from the previous step and the average values of the predictors from the validation period to estimate a synthetic unit, W, that minimizes the RMSPE for the entire pre-treatment period.

[15] The pre-treatment fit index is given by: $\frac{\text{RMSPE}}{\sqrt{\frac{1}{T_0}\sum_1^{T_0}(Y_{1t})^2}}$. One advantage of using the pre-treatment fit index rather than RMSPE is that it normalizes RMSPE, which makes it possible to compare the pre-treatment fit between the synthetic control units across different countries, such as when GDP per worker varies quite significantly across the sample countries. A second advantage is that this approach provides an index number that makes assessing the quality of fit very intuitive.

[16] Following Cavallo et al. (2013), I normalize the estimates before aggregating them by setting the GDP per worker of each treated country equal to one in the treatment year.

is unusually large relative to placebo treatment effects for countries that did not implement a VAT within the sample period, then I conclude that there is statistically significant evidence of an impact of a VAT reform in the treated country. Assuming that I am doing inference about a positive treatment effect, the p-value at each post-treatment year $t \in \{T_0 + 2, T_0 + 3 \ldots, T\}$ is given by:

$$Pr(\hat{\tau}_{1t} < \hat{\tau}_{jt}) = \frac{\sum_{j=2}^{J+1} 1(\hat{\tau}_{1t} < \hat{\tau}_{jt})}{J}, \tag{6}$$

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

Since 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 post-intervention year $t \in T_{0+2}, \ldots, T$ according to the following steps. First, for each country i of group g, I compute the placebo effect for all J_i placebo units from the placebo pool. Second, at each post-treatment 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 $N_{\overline{PL}}$ placebo averages where $N_{\overline{PL}} = \prod_{j=1}^{G} J_i$.[17] Third, at each post-treatment year, I calculate the p-values for the dynamic treatment effect obtained in Equation 5 by using the following equation.

$$Pr(\overline{\tau}_{gt} < \overline{\tau}_{gt}^{PL}) = \frac{\sum_{1}^{N_{\overline{PL}}} 1(\overline{\tau}_{gt} < \overline{\tau}_{gt}^{PL})}{N_{\overline{PL}}}, \tag{7}$$

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

4.3 Sample Selection

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. Since 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 pre-treatment observations of GDP per worker to calibrate the synthetic control. If some treated countries do not have pre-treatment data on GDP per worker, then they are excluded from the study. When data availability is not an issue, I restrict the sample period to 10 pre-treatment years to calibrate

[17]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.

the synthetic unit and 10 post-treatment 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.[18] In my estimation, I include the following set of predictors while estimating the synthetic unit: capital stock per worker, average human capital, a democracy dummy, openness, inflation, and population growth. However, if the country of interest does not have a single observation before the treatment for the relevant predictor, I drop that predictor.

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 non-missing observation before the treatment for each of the covariates used in the estimation. Any country not meeting these requirements is dropped from the donor pool. Further, I exclude from the potential donor pool any country that adopted a VAT before or within the sample period 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 time period implies that the synthetic unit is not reproducing the potential outcome in the absence of treatment. I also exclude transitional countries, small island countries, resource-rich countries, and countries that went through civil war in the sample period from the donor pool. The next step is to run the synthetic control algorithm.

In the third and final step, I select treated countries for which the SCM was able to produce a synthetic unit with a good pre-treatment fit to aggregate their effects by income group. When the pre-treatment fit is bad, we lose the confidence that the treatment effect is due to the reform, because it could be due to the inability of the synthetic unit to mimic the path of the outcome variable of the treated unit. Note that discarding from the analysis the unmatched treatments is similar to confining the analysis to the common support when using matching estimators. I use pre-treatment fit index to decide the quality of the fit, and exclude any country with a pre-treatment fit index greater than 0.10. Similarly, placebo

[18]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 percent or more than the OPEC average. Data are obtained from the World Development Indicators.

countries that do not have a good pre-treatment fit cannot provide information to measure the relative rarity of estimating a large treatment effect. Therefore, I only include placebo countries with pre-treatment fit index less than or equal to 0.10 in the placebo experiment. In some cases, SCM is unable to find any placebo countries with a good pre-treatment fit, leaving the placebo pool empty. When that happens, I also exclude those treated countries with the empty placebo pool.

5 Main Results

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

Figure 2 presents the average causal impact of VAT reforms on GDP per worker for all treated countries. The overall effects are positive and economically meaningful. Five year after the reform, the average GDP per worker of all treated countries is 7.6 percent percent higher than the synthetic group. Overall, post-reform GDP per worker is 6.1 percent higher in the treated group compared to the control group. However, are these results representative of all countries?

To test whether I find differential effects of the VAT adoption by the development stage at which it was adopted, I divide the countries into four groups according to the World Bank's income classification: high-income countries (H), upper-middle-income countries (UM), lower-middle-income countries (LM), and low-income countries (L). This results in 13 high-income countries, 6 upper-middle-income countries, 9 lower-middle-income countries, and 5 low-income countries.[19]

As a first step, I graph country-specific average treatment effects obtained from running the SCM in Figure 3. The horizontal axis represents GDP per worker during the year of VAT adoption normalized by the GDP per Worker of the U.S. The vertical axis represents the change in GDP per worker after a VAT reform of the reforming country compared to

[19] I exclude 6 treated countries from the analysis because the SCM algorithm could not find any placebo countries with pre-treatment fit less than or equal to 0.10 for these countries. Similarly, I exclude 22 treated countries from the analysis because the SCM algorithm produced the synthetic control with pre-treatment fit greater than 0.10. 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). When I increase the pre-treatment fit threshold to include countries with the fit index less than or equal to 0.20, the number of countries in the low-income group increases to 11, the number of countries in the lower-middle-income group increases to 10, and the number of countries in the upper-middle-income group increases to 6. Even then, the impact of the VAT on economic efficiency displays same differential trends across income groups.

change in GDP of its synthetic control. Figure 3 indicates that the impacts of VAT reforms are very heterogeneous. Importantly, the average effects are increasing with the initial level of income, with almost none of the low-income countries having a positive effect, about half of the lower-middle-income countries having positive effects and most of the upper-middle-income and the high-income countries having positive effect. To further explore this pattern and to provide the significance of the estimates, I calculate dynamic treatment effects and their p-values next.

Figure 4 presents the dynamic treatment effects for each of the income group. 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 and the low-income countries. For the high-income countries, 5 years after the reform GDP per worker is 10.9 percent higher than the synthetic group, and it increases to 11.2 percent higher 10 years after the reform, all significant at the 1 percent level. Similarly, I find that the upper-middle-income countries have positive and significant impact on GDP per worker, although after a time lag. The estimates are significant at the 10 percent level from year 3 onwards and at 5 percent level from year 4 onwards. Five years after the reform GDP per worker is 25.5 percent higher than the synthetic group, and it increases to 33.1 percent higher 10 years after the reform. For the lower-middle-income group, I find average effects ranging from 6.7 percent to 9.3 percent higher than the synthetic control, but indistinguishable from zero at the 10 percent level. In the case of the low-income countries, I find negative, but statistically insignificant, impact of VAT adoption.

6 Exploring Potential Mechanisms

In this section, I explore two mechanisms through which a VAT affects economic performance: capital stock per worker and total factor productivity. I use the synthetic control method to estimate the causal effect of the VAT reforms on these variables following the same procedure used when estimating the effect on GDP per worker. To make the estimates as comparable as possible, I only run the analysis on the countries that were included in the main results.[20] The covariates used in the estimation of capital stock per worker and total factor productivity are growth rate of real GDP per capita, average stock of human capital, openness, inflation,

[20]Following the selection criteria of the main results, some country cases were not included if the pre-treatment 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 was not available. Thus, I had to exclude 2 countries from the high-income group, 1 country from the upper-middle-income group (), 3 countries from the low-income group, and 1 country from the low-income group from the capital stock analysis. Similarly, I had to exclude 1 upper-middle-income country, 1 lower-middle-income country and all 5 low-income countries from the total factor productivity analysis.

and democracy dummy.

6.1 Capital Stock

The results for country-specific effects of the VAT reforms on capital stock per worker are presented in Figure 5. Average treatment effects are at the vertical axis and GDP per capita normalized by the GDP per capita of U.S are on the horizontal axis. The impacts on capital stock are heterogeneous, with none of the low-income countries having a positive impact, but most of the middle-income and high-income countries having a positive impact. The results for dynamic treatment effects are presented in Figure 6. For the high-income countries, I find positive and significant impact of the reform at the 5 percent level for all years except year 5, 6, and 7 where the estimates are significant at the 10 percent level. Five years after the reform, capital stock per worker is 5 percent higher in the treated group compared to the synthetic group, which increases to 7.1 percent 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 percent lower capital stock per worker compared to the synthetic control. To summarize, I find the same differential effect across income groups on capital per stock worker as in GDP per worker.

6.2 Total Factor Productivity

Figure 7 graphs country-specific effects of VAT reforms on total factor productivity at the vertical axis and GDP per capita normalized by the GDP per capita of U.S on the horizontal axis. The impact on total factor productivity is heterogeneous, but mostly positive and it is increasing with the initial level of income. Figure 8 presents the results for the average impact across income-groups. Due to insufficient data, I could not analyze the impact on the low-income countries. [21] For the rest of the three income groups, I find that VAT increases total factor productivity. All dynamic treatment effects for the high-income countries are positive and significant at the 1 percent level. Five years after the reform, total factor productivity is 10.1 percent higher than the synthetic group, which increases to 11.8 percent by the end of the sample. The estimates for the upper-middle-income countries are also positive and significant at the 10 percent 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 see an

[21] Only Kenya had data on TFP and Kenya's TFP continued its decreasing trend after the VAT reform.

immediate jump in the 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 percent higher compared to the synthetic group and at its highest, the TFP of the treated group is 11.6 percent higher than the synthetic group. Moreover, 7 of the dynamic effects are significant at the 5 percent level and the remaining 3 dynamic effects are significant at the 10 percent level.

7 Robustness Tests

I conduct a series of robustness checks. First, I test whether one particular treated country is driving the average effect. Second, I test the impact of the VAT using more disaggregated data from manufacturing industry. Third, I test whether the impact of the VAT reform is commingled with other contemporaneous events like the EU membership, the WTO membership, or economic liberalization. Fourth, I estimate the impact of the VAT by restricting the donor pool to the same geographic region or the same income group. Fifth, I estimate the impact of the VAT adoption using conventional DID method to test whether my results are the artifact of using the SCM. 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. Seventh, 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. Eighth, I use tax as a share of GDP to control for overall tax burden while estimating the treatment effects using the SCM. Ninth, I implement leave-one-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. Tenth, I test for the anticipation effects by assuming that the VAT was adopted two years before the actual implementation year. Eleventh and finally, I use in-time placebo experiments by assuming that the VAT was adopted five years before the actual implementation year.

I find the baseline estimates to be largely robust and the first four robustness tests are discussed in detail below. The rest of them are discussed briefly below and in detail in the Appendix.

7.1 Leave-One-Out Averages

I develop a graphical test to inspect whether the average effect is driven by any one treated country or not by drawing 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 9 presents the result for GDP per worker. The leave-one-out averages track the average result pretty closely across all four income groups. Thus, I can conclude that the main results are not driven by the inclusion of any treated country in particular. Similarly, Figure 10 and Figure 11 present the results for capital stock and total factor productivity respectively. I again do not find that any one country is influencing the results. Although in a few cases removing one country from the average changes the size of the effect, it never changes the sign. Thus, I conclude that a few large outliers are not driving the central finding.

7.2 Manufacturing Industry

There are two concerns with using aggregate data to analyze the effect of the VAT reforms that the use of manufacturing sector level data can help address. First, if the VAT reforms are motivated by expectation of future growth prospects, this would bias the estimates obtained from the SCM as long as growth expectations are not captured by the unobservable heterogeneity included in the estimation. Thus, I cannot rule out the possibility of the bias caused by reverse causality. The use of manufacturing-level data helps mitigate the concern of reverse causality because VAT adoptions take place at the national level and thus are not likely to be affected by the underlying trends in one of the numerous sectors in any significant way.

Second, how a VAT affects different sectors of the economy is different across countries. In some countries, services are taxed broadly while in others they are only taxed selectively. In most cases, the VAT is extended through retail stage, but in others it is only extended through manufacturing stage. Similarly, sectors such as agriculture, financial sector, education sector, and health services sector are all treated differently in each country. Thus, the aggregate impact of the VAT masks this heterogeneity. However, the treatment of manufacturing sector is quite homogenous across countries. For instance, in almost all countries, the VAT was introduced to replace sales and turnover taxes that directly affected manufacturing industry and in all cases the VAT includes the manufacturing industry in the tax base. Thus, by focusing exclusively on the manufacturing industry I can isolate the impact of a more uniform VAT.

[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.

I use proprietary data on industrial statistics (INDSTAT2) from United Nations Industrial Development Organization (UNIDO), which provide industrial statistics for a wide range of countries at manufacturing level going back to early 1960s. It covers value-added, employment, and wage, among other things, allowing me to calculate labor productivity (value-added per employee) and average wage (wage per employee). Average wage provides an additional robustness test on labor productivity because if labor is more productive then it increases the marginal product of labor and as a result, the market-clearing wage. The data are reported in current U.S. dollars. I follow the approach suggested by Levchenko, Rancire and Thoenig (2009) and convert them into constant international dollars using price deflators from the PWT 8.1. To my knowledge, this is the only source of data that goes back to 1960s, allowing me to analyze countries from Western Europe and Latin America that adopted VAT in the 1960s and 1970s. It is also the only source of data for the middle-income and the low-income countries. However, only a few middle-income and low-income countries have long time series, but to make the results comparable to the main results, I need 21 years of non-missing data. Accordingly, the empirical analysis is based on 8 high-income countries (Austria, Belgium, Denmark, France, Ireland, Italy, Netherlands, Sweden, and United Kingdom) for labor productivity and 6 high-income countries (Austria, Belgium, Denmark, Italy, Netherlands, and United Kingdom) for average wage that meet the data requirements as well as the requirements of the SCM.

The results for labor productivity, average wage, and their leave-one-out averages are presented in Figure 12. Both labor productivity and average wage increase more in the reforming countries than their respective synthetic groups after the VAT reform. The dynamic treatment effects range from 2.7 percent higher 1 year after the reform to 10.3 percent higher 4 years after the reform, after which it starts decreasing and becomes statistically insignificant at the 10 percent level, but remains positive throughout the sample period. The leave-one-out averages suggests that the results are not driven by the inclusion of any particular treated country in the group. The results are slightly stronger and more precisely estimated for average wage. The estimates range from 3.7 percent higher to 13.6 percent higher in the treated group compared to the synthetic group. All estimates from year 2 onwards are significant at the 10 percent level with estimates from year 4 onwards being significant at the 5 percent or better level. Again, the leave-one-out test rules out the possibility that the results are driven by an outlier country.

7.3 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 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., 10 years before or 10 years after VAT adoption). Furthermore, I use two of the most widely used measures of economic liberalization: WTO membership and a binary indicator by constructed by Sachs and Warner (1995) and updated by Wacziarg and Welch (2008).

The aggregate effects controlling for the EU membership are presented in the top row of Figure 13. Six countries joined the EU within 10 years of VAT adoption and they belonged to either the high-income or the upper-middle-income group.[23] The results are unchanged by the exclusion of these countries. For the high-income countries, the impact of the VAT controlling for the EU membership ranges from 6.6 percent to 14.6 percent higher compared to the synthetic group, all significant at the 1 percent level. For the upper-middle-income countries, the impact is positive and significant, but after two year time lag. The estimate from year 3 is significant at the 10 percent level and the estimates from year 4 onwards are significant at the 1 percent level. The statistically significant estimates range from 4.2 percent to 29.5 percent higher compared to the synthetic group.

The aggregate effects controlling for the WTO membership (14 countries in total) are presented in the bottom row of Figure 13 and the aggregate effects controlling for Wacziarg and Welch (2008) economic liberalization (13 countries in total) are presented in the top row of Figure 14.[24] For the high-income countries, I again find the impact on GDP per worker is positive and statistically significant at the 1 percent level, although the size of the effects are slightly smaller (ranging from 1.9 percent to 7.4 percent). Similarly, the upper-middle-income countries also have positive and significant impact, after a time lag. 7 out of 10 estimates are significant at the 10 percent level and 6 of those are significant at the 1 percent level. Five years after the reform, GDP per worker is 22.4 percent higher, which increases to 37.1 percent by the end of the sample period. The lower-middle-income countries experience economically meaningful, but statistically insignificant improvement.[25]

The results controlling for Wacziarg and Welch (2008) economic liberalization are pre-

[23] Denmark, Greece, Ireland, Portugal, Spain, and United Kingdom joined the EU within the sample period and thus were removed from the analysis.

[24] The WTO members were Bangladesh, Canada, Greece, Guinea, Jamaica, Japan, Kenya, Mauritius, Nepal, New Zealand, Pakistan, Portugal, Spain, and Thailand. The countries that underwent economic liberalization were Bangladesh, Chile, Denmark, the Dominican Republic, France, Guinea, Ireland, Jamaica, Kenya, Nepal, Netherlands, New Zealand, and Sweden.

[25] Since all four low-income countries joined the WTO within 10 years of VAT adoption, I could not analyze the impact on the low-income countries.

sented in the top two rows of Figure 14. I again find that VAT adoption has positive and significant impact on the high-income countries, positive and significant impact after a time lag on the upper-middle-income countries, slightly positive, but insignificant impact on the lower-middle-income countries, and mixed, but statistically insignificant impact on the low-income countries. Moreover, this trend holds even when I control for the EU membership, WTO membership, and economic liberalization at the same time as illustrated in the bottom row of Figure 14.

To conclude, in every specification, I find an economically and statistically significant increase in economic efficiency on the high-income countries and the upper-middle-income countries, but mixed and insignificant impact on the lower-middle-income and the low-income countries. Thus, I conclude that the results are driven by the VAT adoption itself and not by other contemporaneous events.

7.4 Restricted Donor Pools

In the main analysis, I run the analysis using a worldwide sample of countries. The main advantages of doing so are that it includes more countries in the donor pool so the probability of finding a synthetic country that matches the treated country increases and it also increases the size and the power of the test. However, one disadvantage of using a worldwide sample of countries is that I cannot rule out the possibility of some far-fetched country comparisons, although this possibility is minimized in practice by including region dummies and income group dummies in the estimation and by requiring the synthetic control to match the initial level of the outcome variable and its important predictors.

I use two alternative control groups where I restrict potential controls to the same income group or the same geographic region to ensure the existence of a common support between treated and control countries. This approach has two main advantages. First, it avoids biases caused by interpolating across countries with very different characteristics. That is, even if I am able to find a synthetic unit with good pre-treatment fit, interpolation biases may be large if the linear factor model used in the estimation of the synthetic control does not hold over the entire set of regions in any particular sample. This happens if the relationship between the outcome variables and the predictors is highly non-linear and the combination of two extreme donor units is used to construct a synthetic unit that has the average value. In the similar spirit, it also controls for unobservable characteristics associated with the level of economic development or geography and any other secular changes over time that might affect countries from different income groups or different geographic regions differently.

The results restricting the control group to be from the same geographic region are presented in Figure 15. I was able to analyze only 8 countries from the high-income group, and 1 each from the rest of the income groups because of the restrictions imposed to the donor pool. Even so, I find the same differential pattern that I find in the main results. For the high-income countries, the impact of VAT adoption is positive, ranging from 2.3 percent to 8.8 percent higher compared to the synthetic group and they become statistically significant at the 10 percent or better level from year 4 onwards. Although much cannot be said about other income groups because of small number of countries included, I find that the upper-middle-income country has positive impact, the lower-middle-income country has mostly negative impact and the low-income country has negative impact of VAT adoption.

The results restricting the control group to be from the same income group are presented in Figure 16. I was able to analyze only 8 countries from the high-income group, 2 from the upper-middle-income group, 3 from the lower-middle-income group, and 1 from the low-income group because of the restrictions imposed to the donor pool. Even so, I find the same differential pattern that I find in the main results. For the high-income countries, I find an immediate positive and statistically significant impact at the 5 percent or better level. The estimates range from 2.4 percent to 8.7 percent higher compared to the synthetic group. Although much cannot be said about other income groups because of the small number of countries included, I find that the upper-middle-income group has positive impact after some time lag, while the lower-middle-income and the low-income group have mostly negative impacts of VAT adoption.

7.5 Conventional Difference-in-Differences

I also estimate the impact of VAT reforms using the conventional difference-in-differences (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. Doing so allows me to test whether the parallel trends assumption is violated or not and to assess the dynamic nature of the 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, 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 17. The Figure in left panel includes country fixed effects, and year fixed effects while the figure in right panel also includes country-specific linear time trends. The figure in the top row is obtained by running regression on the sample of 33 treated countries used in the baseline SCM estimates, thus the effect of VAT comes from the variation in the timing of VAT adoption. The figure in the bottom row adds 24 countries that contributed to the construction of the synthetic controls in the SCM results. Since 23 out of 24 control countries eventually adopted a VAT it implies that I am still exploiting the variation in the timing of VAT adoption to find its effects. In all specifications, the standard errors are clustered at the country level. 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. The results in the right panel indicate that VAT adoption has positive and significant impact on economic efficiency. These effects increase with time for first few years and then start decreasing. For instance, GDP per worker is $407 higher in the treated group compared to the control group in the first year after the reform, which increases up to $840 five years after the reform. The estimates from year 2 onwards are significant at the 5 percent level, but they become insignificant from year 7 onwards.[26] The results for 57 countries are presented in the bottom row. The differential trends between the treated group and the control group are more prominent than the estimation with just 33 treated countries; however, results controlling for country-specific linear trends are very similar.

7.6 Additional Robustness Tests

I also conduct several additional robustness tests in the Appendix. First, I construct the income classification system using a data-driven approach. That is, I divide the countries into 4 quartiles each year, according to their GDP per Worker, and use the first quartile as the low-income, the second quartile as the lower-middle-income, the third quartile as the upper-middle-income, and the fourth quartile as the high-income group. Next, I use this classification system to run the SCM and to aggregate the effects into different groups. I obtain very similar results.

Second, in the main analysis, I use a cross-validation technique to minimize the risk of over-fitting by dividing the pre-treatment data into training period and validation period. I test whether my results are sensitive to the use of a cross-validation technique by following a more traditional approach and using the entire pre-treatment data to find the synthetic control. The results remain unchanged.

Third, I use four alternative measures of economic efficiency to test whether my main

[26] I also run the analysis that includes all the covariates used in the main SCM estimation, the results are similar.

results are dependent to the use of output-side GDP per worker from PWT 8.1 as a measure of economic efficiency. The alternative measures are output-side GDP per capita from PWT 8.1, expenditure-side GDP per worker from PWT 8.1, output-side GDP per worker from PTW 8.0, and output-side GDP per capita from PTW 8.0. I find that the results using GDP per worker are slightly smaller in magnitude and they become significant after a year compared to GDP per worker. Other than that the results are largely unchanged.

Fourth, I control for the initial level of tax burden by adding tax as a share of GDP as one of the control variables while estimating the impact of VAT adoption on GDP per worker. The results for the high-income, lower-middle-income and low-income remains very similar. However, the estimates for the upper-middle-income countries become statistically insignificant (although the point estimates remain positive and 6 out of 10 estimates have p-values less than 0.2).

Fifth, I iteratively re-estimate the baseline model excluding in each iteration one of the countries that received positive weights in the baseline model. This allows me to assess to what degree the main results are driven by any particular donor country. The leave-out-out average synthetic units very closely track all 13 high-income countries, indicating that any one particular donor country is not driving the main results. The leave-one-out averages also closely track most of the upper-middle-income countries (two exceptions) and most of the lower-middle-income countries (three exceptions), but about half of the low-income countries are sensitive to the inclusion of specific donor countries. I conclude that the baseline estimates are not very sensitive (especially for richer countries) to the inclusion of any particular country.

Sixth, it is well-known that policies of these magnitudes are often implemented with some time lag. This time lag can potentially create an anticipation effect that could bias the estimates. I test and correct for the anticipation effect by running the synthetic control algorithm by assuming that the VAT was adopted two years prior to the actual VAT adoption year. In doing so, I am constructing the synthetic unit with data before the anticipation might reasonably be expected to have any impact. The results show that the post-VAT GDP per worker trajectory of the treated group and the synthetic group closely track each other up to the actual VAT adoption year and then they diverge, with the magnitude of the divergence very similar to the baseline estimates. This suggests that there were no anticipation effects affecting the baseline results.

Seventh, I conduct in-time placebo experiments by assuming that the country of interest adopted a VAT five years (i.e., halfway through the baseline pre-treatment period) before it actually did and rerunning the SCM. The synthetic control group almost exactly reproduces the trajectory of GDP per capita of the high-income group. There is some gap between the treated group and the control group for the rest of the income groups. However, these gaps are much smaller than what I find in the main results. Thus, I conclude that in contrast to impact of the actual VAT adoption, the placebo VAT adoptions do not have any sizable effect.

To conclude, I find the baseline estimates to be largely robust, and they are discussed in detail in the Appendix.

8 Conclusions

This paper analyzes the impact of replacing sales and turnover taxes with VAT on economic efficiency by combining information from comparative case studies obtained with a synthetic control methodology recently developed in Abadie and Gardeazabal (2003) and expanded in Abadie, Diamond and Hainmueller (2010, 2015). The procedure involves identifying the causal effects by comparing the trajectory of post-treatment 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 two potential mechanisms (capital accumulation and total factor productivity) through which a VAT can affect GDP per worker and find that both the channels are important. I again 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 VAT do not necessarily translate into practice. In particular, the impact of 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, in order 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; Carrillo, Pomeranz and Singhal, 2014;

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 since my results indicate sizable gains in economic efficiency from adopting a well-designed and well-enforced VAT.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of Californias Tobacco Control Program." *Journal of the American Statistical Association*, 105(490).

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*, 93(1): 113–132.

Abdallah, Chadi S., and William D. Lastrapes. 2012. "Home Equity Lending and Retail Spending: Evidence from a Natural Experiment in Texas." *American Economic Journal: Macroeconomics*, 4(4): 94–125.

Adhikari, Bibek, and James Alm. 2015. "Evaluating the Economic Effects of Flat Tax Reforms Using Synthetic Control Methods." *Tulane Univeristy Working Paper*.

Alm, James, and Abel Embaye. 2013. "Using Dynamic Panel Methods to Estimate Shadow Economies Around the World, 1984 - 2006." *Public Finance Review*, 41(5): 510–543.

Alm, James, and Asmaa El-Ganainy. 2013. "Value-Added Taxation and Consumption." *International Tax and Public Finance*, 20(1): 105–128.

Andre, Francisco J, Alejandro Cardenete, and Carlos Romero. 2010. "Designing Public Policies: An Approach Based on Multi-Criteria Analysis and Computable General Equilibrium Modeling." Chapter Economic Policy Using Applied General Equilibrium Models: An Overview. Springer.

Ballard, Charles L., John Karl Scholz, and John B. Shoven. 1987. "The Value-Added Tax: A General Equilibrium Look at its Efficiency and Incidence." *The Effects of Taxation on Capital Accumulation*, 445–480. University of Chicago Press.

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

Bauhoff, Sebastian. 2013. "The Effect of School District Nutrition Policies on Dietary Intake and Overweight: A Synthetic Control Approach." *Economics and Human Biology*.

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

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

Bird, Richard, and Pierre-Pascal Gendron. 2011. *The VAT in Developing and Transitional Countries. Cambridge Books*, Cambridge University Press.

Boeters, Stefan, Christoph Bohringer, Thiess Buttner, and Margit Kraus. 2010. "Economic Effects of VAT Reforms in Germany." *Applied Economics*, 42(16-18): 2165 – 2182.

Bohn, Sarah, Magnus Lofstrom, and Steven Raphael. 2014. "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?" *The Review of Economics and Statistics*, 96(2): 258269.

Bovenberg, Lans A. 1987. "Indirect Taxation in Developing Countries: A General Equilibrium Approach." *Staff Papers: International Monetary Fund*, 34: 333–73.

Brennan, Geoffrey, and James M. Buchanan. 1977. "Towards a Tax Constitution for Leviathan." *Journal of Public Economics*, 8(3): 255 – 273.

Bye, Brita, Birger Strom, and Turid Avitsland. 2012. "Welfare Effects of VAT Reforms: A General Equilibrium Analysis." *International Tax and Public Finance*, 19(3): 368–392.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal. 2014. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement." National Bureau of Economic Research Working Paper 20624.

Casanegra de Jantscher, Milka. 1990. "Value Added Taxation in Developing Countries." , ed. M. Gillis, C. Shoup and G. P. Sicat, Chapter Administering the VAT, 171190. Washington, DC: World Bank.

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

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *The Review of Economics and Statistics*, 95(5).

Cizek, P., J. Lei, and J.E. Ligthart. 2012. "The Determinants of VAT Introduction : A Spatial Duration Analysis." Tilburg University, Center for Economic Research Discussion Paper 2012-071.

Cnossen, Sijbren. 1998. "VATs in CEE Countries: A Survey and Analysis." *De Economist*, 146(2): 227.

Coolidge, Jacqueline. 2012. "Findings of Tax Compliance Cost Surveys in Developing Countries." *eJournal of Tax Research*, 10, no. 2: 250–287.

DeAngelo, Gregory, and Benjamin Hansen. 2014. "Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities." *American Economic Journal: Economic Policy*, 6(2): 231–57.

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

Dorsett, Richard. 2013. "The Effect of the Troubles on Gdp in Northern Ireland." *European Journal of Political Economy*, 29(C): 119–133.

Ebrill, Liam, Michael Keen, and Jean-Paul Bodin. 2001. *The Modern VAT.* International Monetary Fund. Washington D.C.

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

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

Ferede, Ergete, and Bev Dahlby. 2012. "The Impact of Tax Cuts on Economic Growth: Evidence from the Canadian Provinces." *National Tax Journal*, 65(3): 563 – 594.

Gordon, Roger, and Wei Li. 2009. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." *Journal of Public Economics*, 93(78): 855 – 866.

Hines, James R. 2004. "Might Fundamental Tax Reform Increase Criminal Activity?" *Economica*, 71(283): 483–492.

Inklaar, Robert, and Marcel P. Timmer. 2013. "Capital, Labor and TFP in PWT8.0." Groningen Growth and Development Centre, University of Groningen.

International Country Risk Guide. 2011. "IRIS-3: International Country Risk Guide, 1982-1997." *The Political Risk Services Group.*

International Tax Dialogue. 2005. "The Value Added Tax: Experiences and Issues." International Tax Dialogue.

James, Kathryn. 2015. *The Rise of the Value-Added Tax.* Cambridge University Press.

Jones, Larry E, and Rodolfo E Manuelli. 1990. "A Convex Model of Equilibrium Growth: Theory and Policy Implications." *Journal of Political Economy*, 98(5): 1008–38.

Keen, Michael, and Ben Lockwood. 2006. "Is the VAT a Money Machine?" *National Tax Journal*, 59(4): 905 – 928.

Keen, Michael, and Ben Lockwood. 2010. "The Value Added Tax: Its Causes and Consequences." *Journal of Development Economics*, 92(2): 138 – 151.

Keen, Michael, and Stephen Smith. 2006. "VAT Fraud and Evasion: What Do We Know and What Can Be Done?." *National Tax Journal*, 59(4): 861 – 887.

Kneller, Richard, Michael F Bleaney, and Norman Gemmell. 1999. "Fiscal Policy and Growth: Evidence From OECD Countries." *Journal of Public Economics*, 74(2): 171–190.

Lee, Young, and Roger H. Gordon. 2005. "Tax Structure and Economic Growth." *Journal of Public Economics*, 89(56): 1027 – 1043.

Levchenko, Andrei A., Romain Rancire, and Mathias Thoenig. 2009. "Growth and Risk at the Industry Level: The Real Effects of Financial Liberalization." *Journal of Development Economics*, 89(2): 210 – 222. New Approaches to Financial Globalization.

Lucas, Robert Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics*, 22(1): 3–42.

Mankiw, N Gregory, David Romer, and David N Weil. 1992. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics*, 107(2): 407–437.

Marshall, Monty G, and Keith Jaggers. 2013. "Polity IV Project: Politial Regime Characteristics and Transitions, 1800-2012."

Martinez-Vazquez, Jorge, and Richard M. Bird. 2011. "Value-Added Tax: Onward and Upward?" In *The Elgar Guide to Tax Systems*. Chapter 6. Edward Elgar.

Metcalf, Gilbert E. 1995. "Value-Added Taxation: A Tax Whose Time Has Come?." *Journal of Economic Perspectives*, 9(1): 121 – 140.

Mikesell, John L. 2014. "Misconceptions about Value-Added and Retail Sales Taxes: Are They Barriers to Sensible Tax Policy?" *Public Budgeting & Finance*, 34(2): 1–23.

Montalvo, Jos G. 2011. "Voting After The Bombings: A Natural Experiment On The Effect Of Terrorist Attacks On Democratic Elections." *The Review of Economics and Statistics*, 93(4): 1146 – 1154.

Nannicini, Tommaso, and Andreas Billmeier. 2011. "Economies in Transition: How Important Is Trade Openness for Growth?" *Oxford Bulletin of Economics and Statistics*, 73(3): 287–314.

Nellor, David. 1987. "The Effect of the Value-Added Tax on the Tax Ratio." *IMF Working Paper*, 1–28.

Oakland, W. H. 1967*a*. "The Theory of the Value-Added Tax: I. A Comparison of Tax Bases." *National Tax Journal*, 20: 119 – 136.

Oakland, W. H. 1967*b*. "The Theory of the Value-Added Tax: II. Incidence Effects." *National Tax Journal*, 20: 270 – 281.

Persson, Torsten, and Guido Tabellini. 2007. "The Growth Effect of Democracy: Is It Heterogenous and How Can It Be Estimated?" National Bureau of Economic Research Working Paper 13150.

Piggott, John, and John Whalley. 2001. "VAT Base Broadening, Self Supply, and the Informal Sector." *American Economic Review*, 91(4): 1084 – 1094.

Pomeranz, Dina. 2015. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review*, 105(8): 2539–69.

Rebelo, Sergio. 1991. "Long-Run Policy Analysis and Long-Run Growth." *Journal of Political Economy*, 99(3): 500–521.

Ring, Raymond J., Jr. 1989. "The Proportion of Consumers' and Producers' Goods in the General Sales Tax." *National Tax Journal*, 42(2): 167 – 179.

Ring, Raymond J., Jr. 1999. "Consumers' Share and Producers' Share of the General Sales Tax." *National Tax Journal*, 52(1): 79–90.

Riswold, Scott. 2004. "VAT in Sub-Saharan Africa - A Critique of IMF VAT Policy." *International VAT Monitor.*

Sachs, Jeffrey D., and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity*, 26(1, 25th A): 1–118.

Schneider, Friedrich. 2005. "Shadow Economies Around the World: What Do We Really Know?" *European Journal of Political Economy*, 21(3): 598 – 642.

Shoven, John B, and John Whalley. 1984. "Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey." *Journal of Economic Literature*, 1007–1051.

Smart, Michael, and Richard M. Bird. 2009. "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): 591–609.

Smith, Brock. 2015. "The Resource Curse Exorcised: Evidence From a Panel of Countries." *Journal of Development Economics*, 116(C): 57–73.

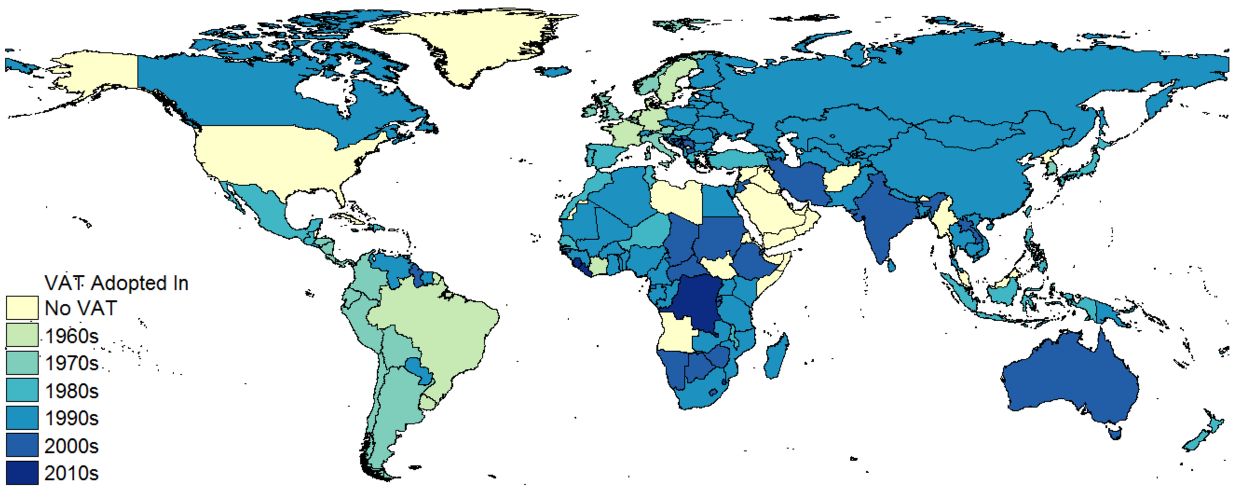
Tait, Alan A. 1988. *Value Added Tax: International Practice and Problems.* Vol. 24, International Monetary Fund.

The President's Advisory Panel on Federal Tax Reform. 2005. "Simple, Fair, and Pro-Growth: Proposal to Fix America's Tax System." Chapter Value-Added Tax, 192–222.

Ufier, Alex. 2014. "Quasi-experimental Analysis on the Effects of Adoption of a Value Added Tax." *Economic Inquiry*, 52(4): 1364–1379.

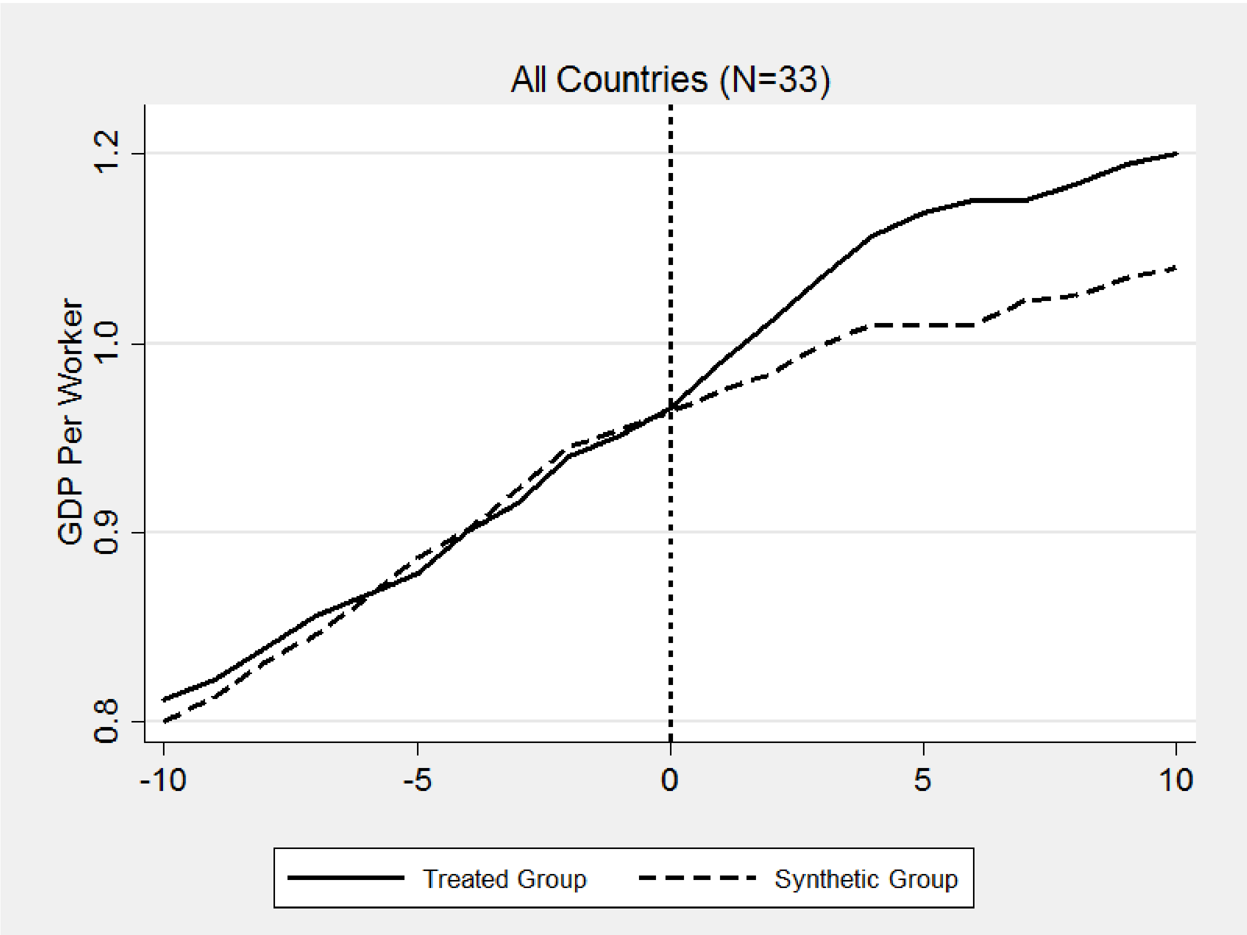
Wacziarg, Romain, and Karen Horn Welch. 2008. "Trade Liberalization and Growth: New Evidence." *World Bank Economic Review*, 22(2): 187–231.

Figure 1: The Spread of the Value Added Tax



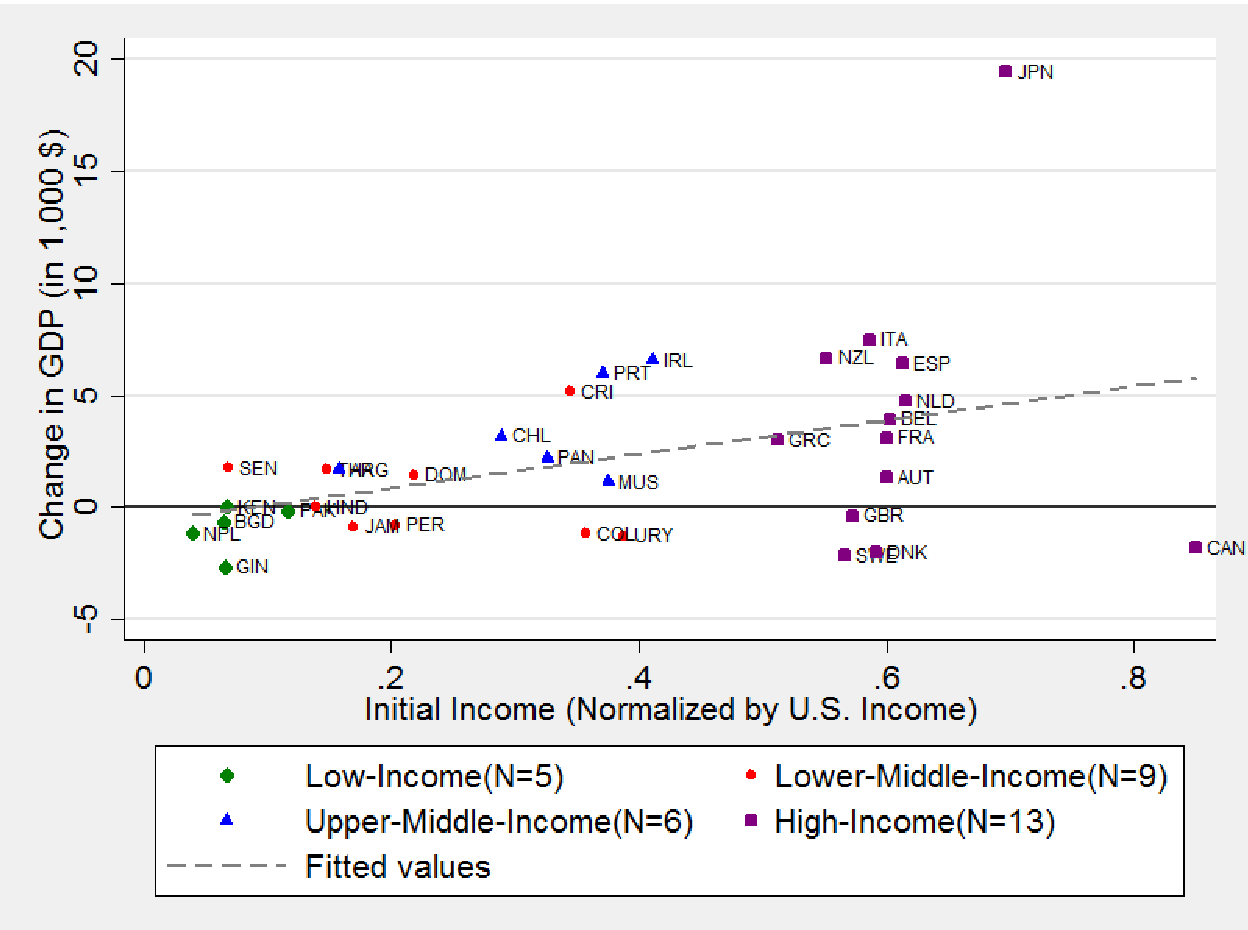
Notes: The map shows the countries that adopted Value Added Tax (VAT) between 1960-2013.

Figure 2: The Overall Impact of VAT Adoption on Average GDP Per Worker



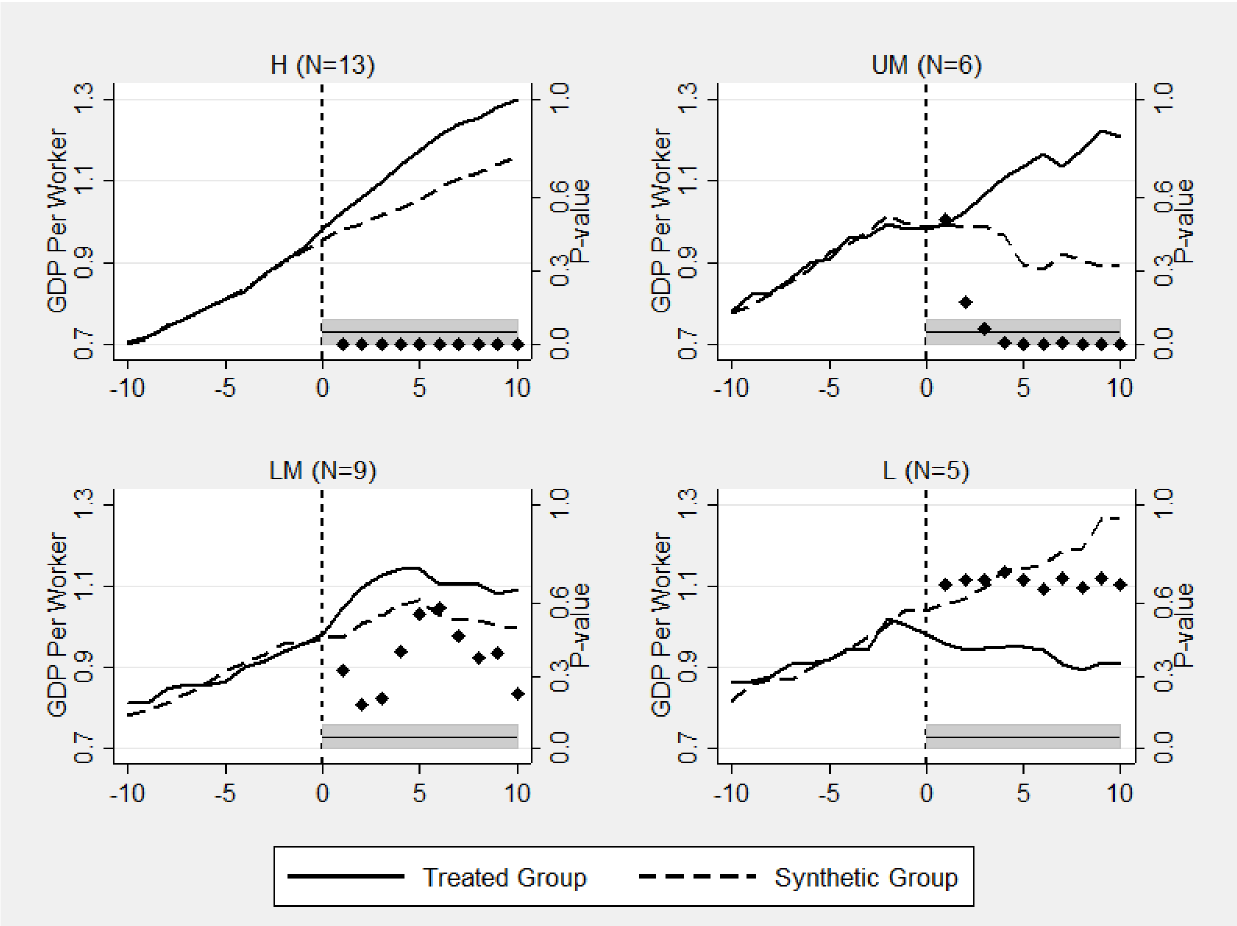
Notes: The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome of its synthetic control. N denotes the total number of countries. The left axis consists of labels for the outcome variable. To avoid the size effect caused due to varying 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.

Figure 3: The Impact of VAT Adoption on Average GDP Per Worker Across Countries with Different Initial Income



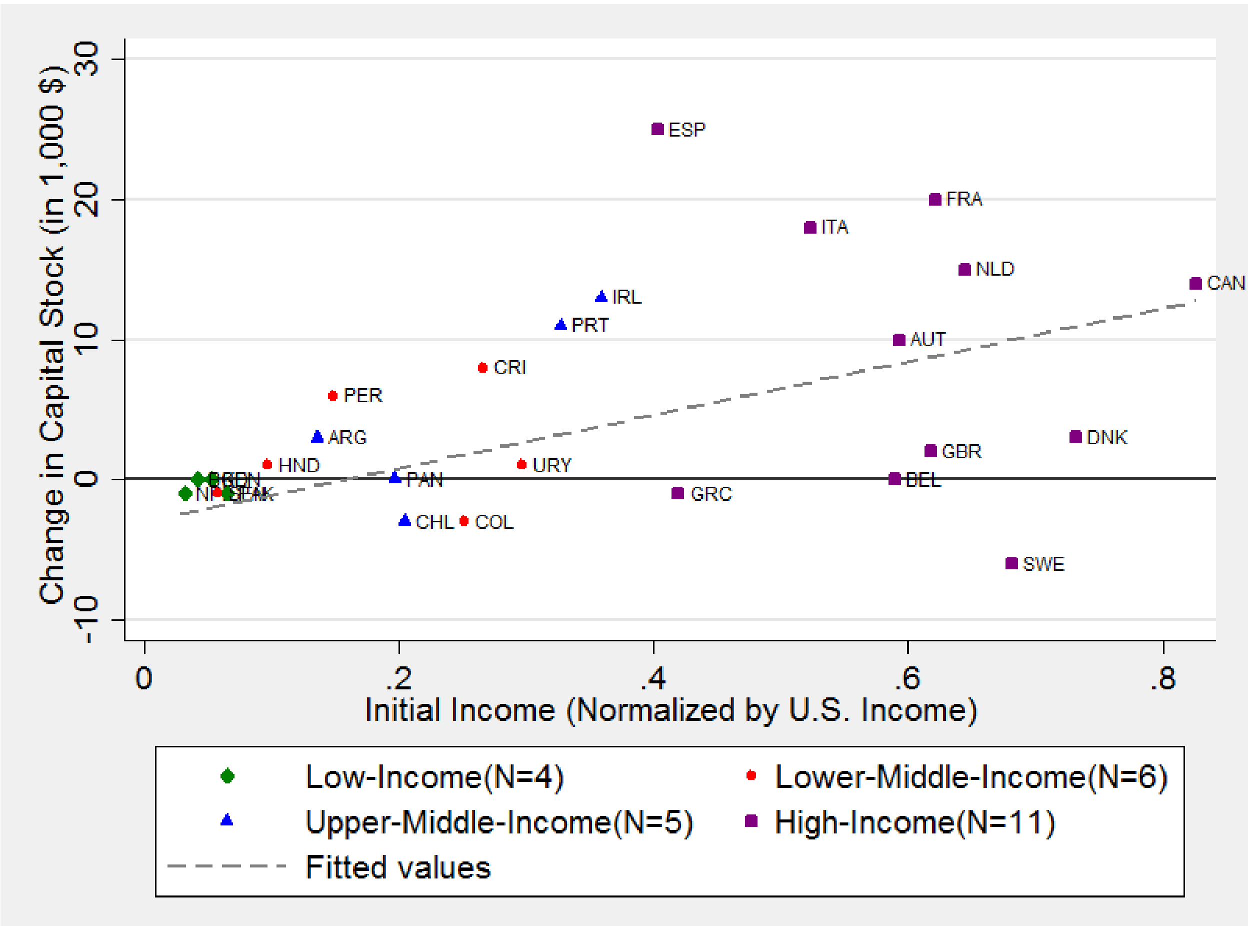
Notes: Country-specific average treatment effects are obtained using synthetic control methods by subtracting the average value of the pre-intervention difference in the outcome variable from the average value of the post-intervention difference in the outcome variable for each country i and its synthetic control. The y-axis consists of labels for the outcome variable, which is expressed in 2005 constant international dollars, and is adjusted for purchasing power parity. The x-axis consists of the income of treated country at treatment year normalized by the income of the U.S in that year.

Figure 4: The Impact of VAT Adoption on Average GDP Per Worker Across Different Income Groups



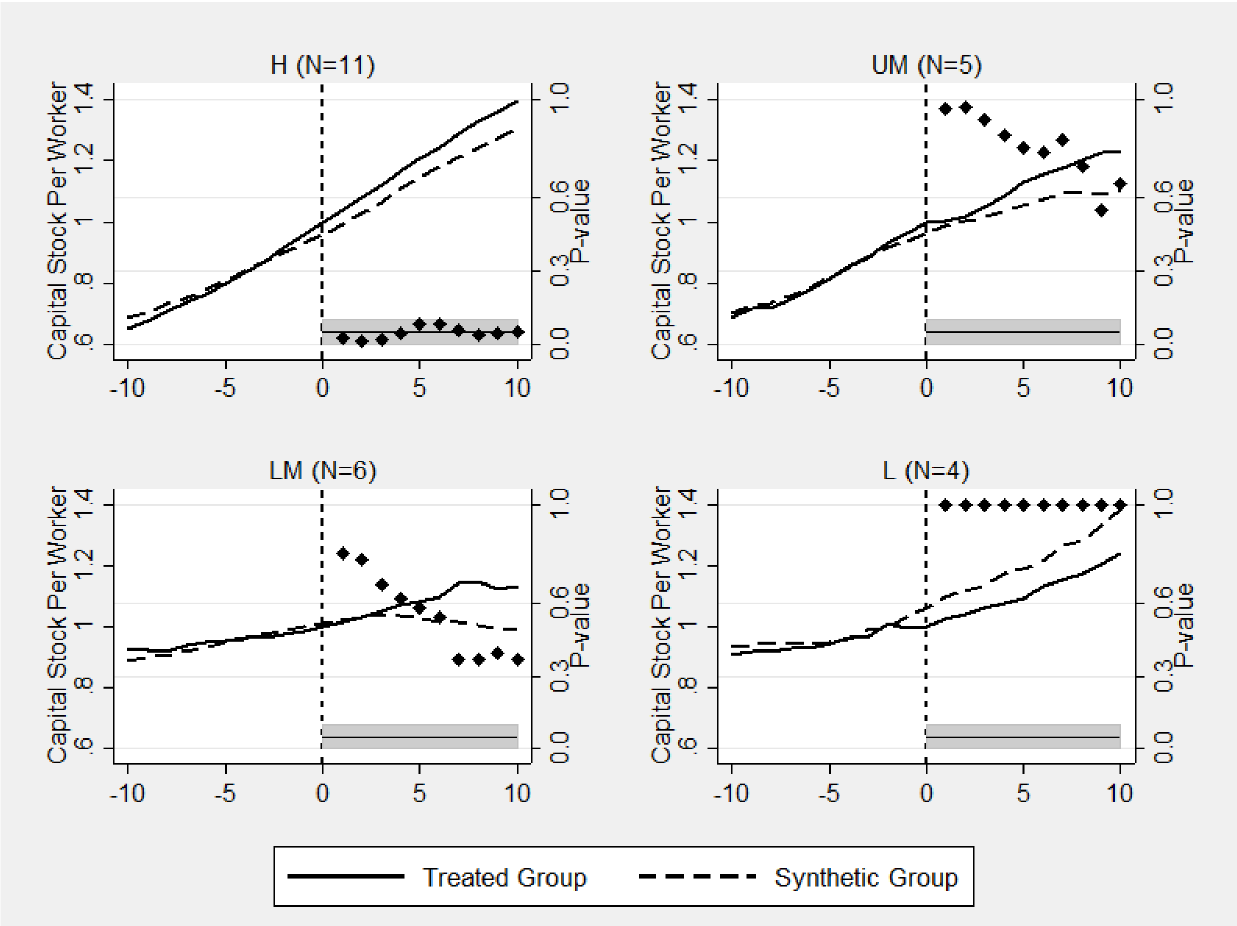
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 5: The Impact of VAT Adoption on Average Capital Stock Per Worker Across Countries with Different Initial Income



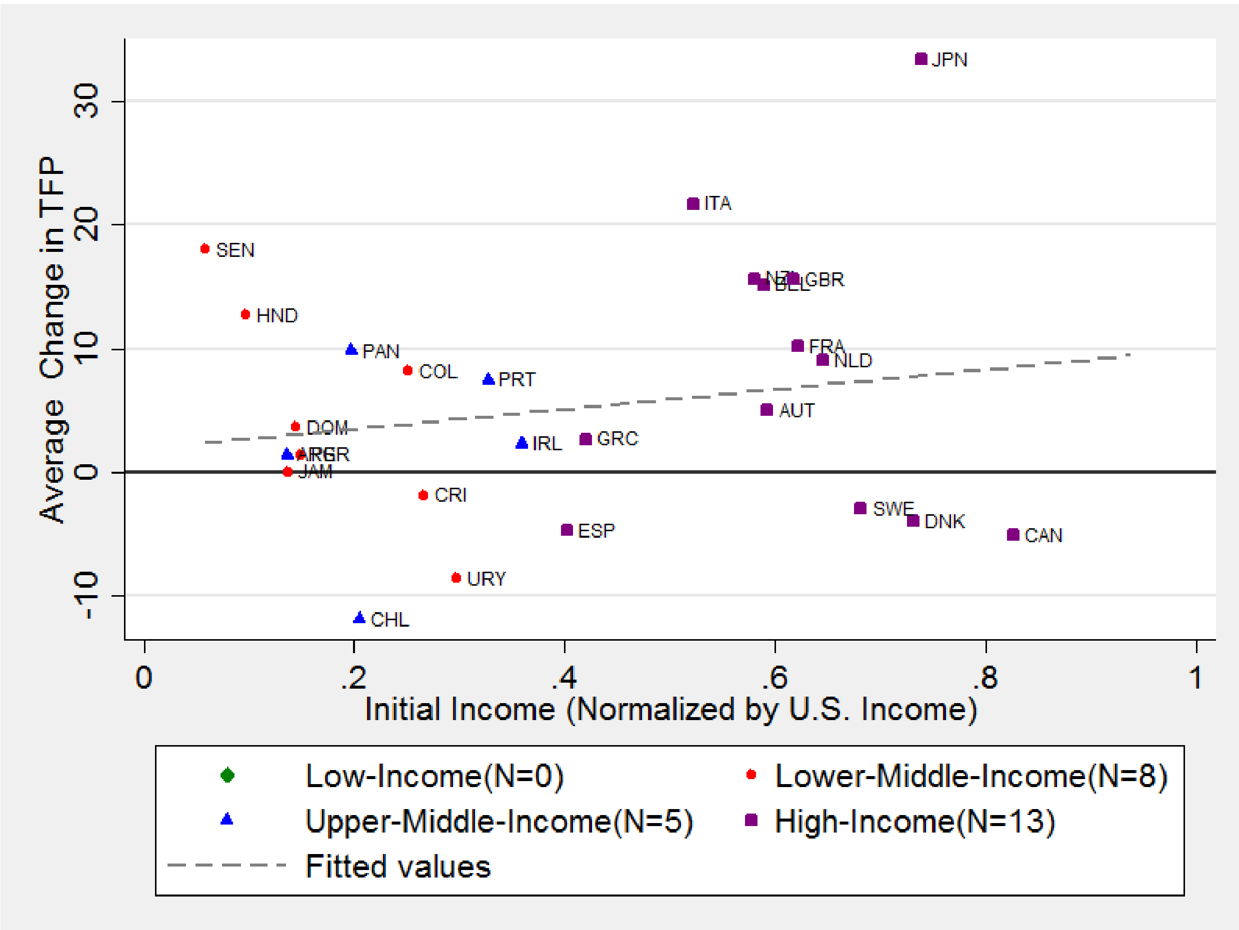
Notes: Country-specific average treatment effects are obtained using synthetic control methods by subtracting the average value of the pre-intervention difference in the outcome variable from the average value of the post-intervention difference in the outcome variable for each country i and its synthetic control. The y-axis consists of labels for the outcome variable, which is expressed in 2005 constant international dollars, and is adjusted for purchasing power parity. The x-axis consists of the income of treated country at treatment year normalized by the income of the U.S in that year.

Figure 6: The Impact of VAT Adoption on Average Capital Stock Per Worker Across Different Income Groups



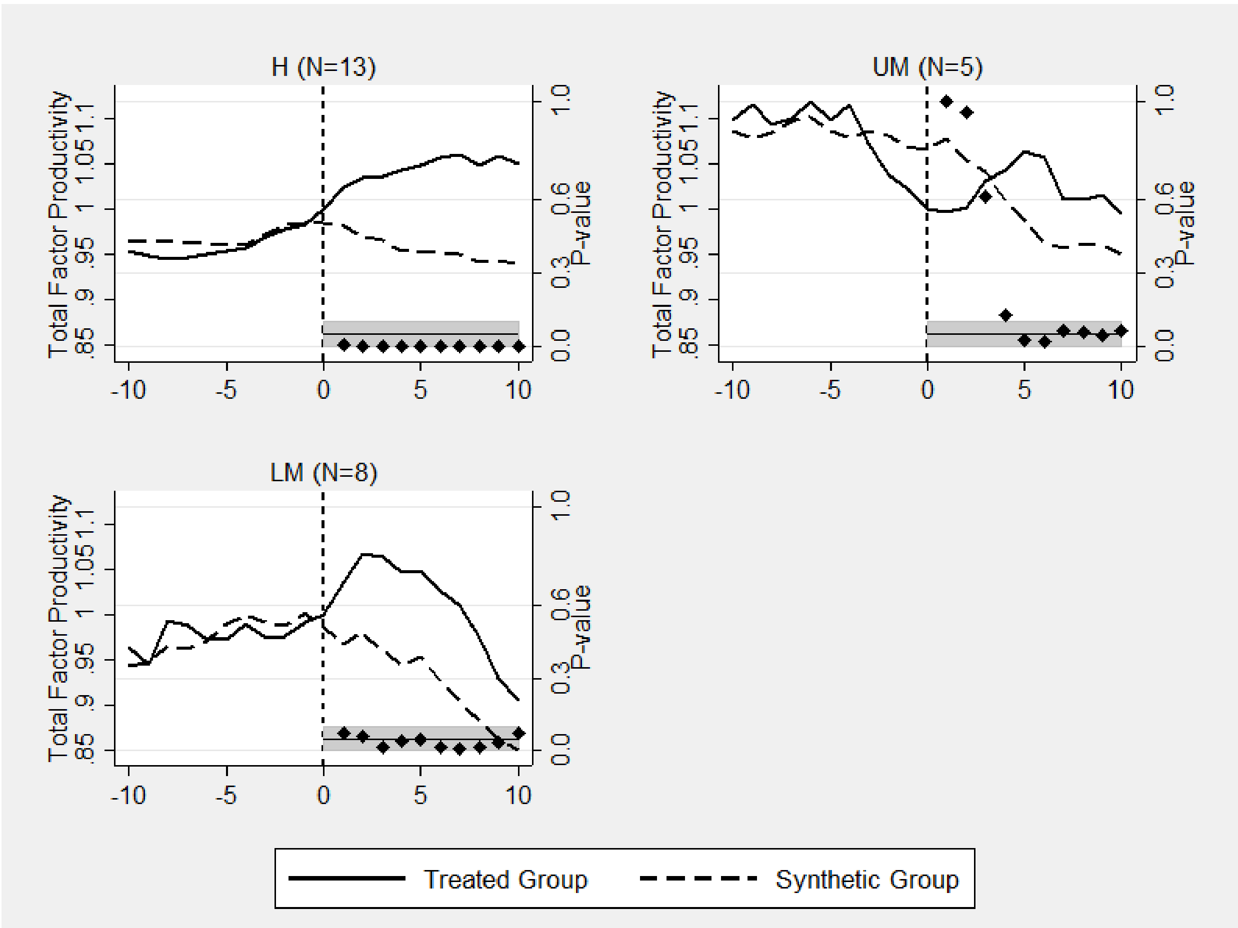
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 7: The Impact of VAT Adoption on Average Total Factor Productivity Across Countries with Different Initial Income



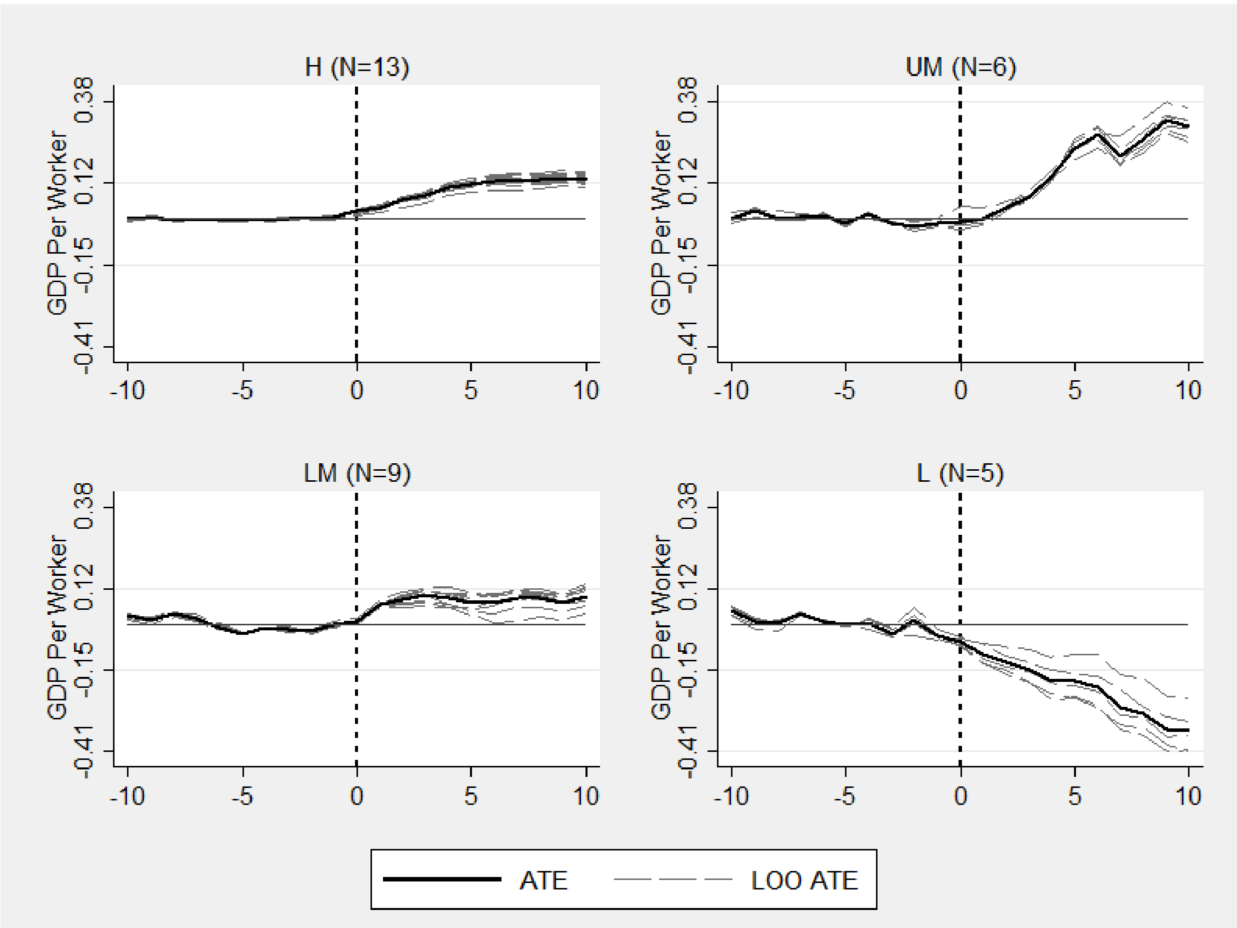
Notes: Country-specific average treatment effects are obtained using synthetic control methods by subtracting the average value of the pre-intervention difference in the outcome variable from the average value of the post-intervention difference in the outcome variable for each country i and its synthetic control. The y-axis consists of labels for the outcome variable, which is normalized using TFP of the United States. The x-axis consists of the income of treated country at treatment year normalized by the income of U.S at that year.

Figure 8: The Impact of VAT Adoption on Average Total Factor Productivity Across Different Income Groups



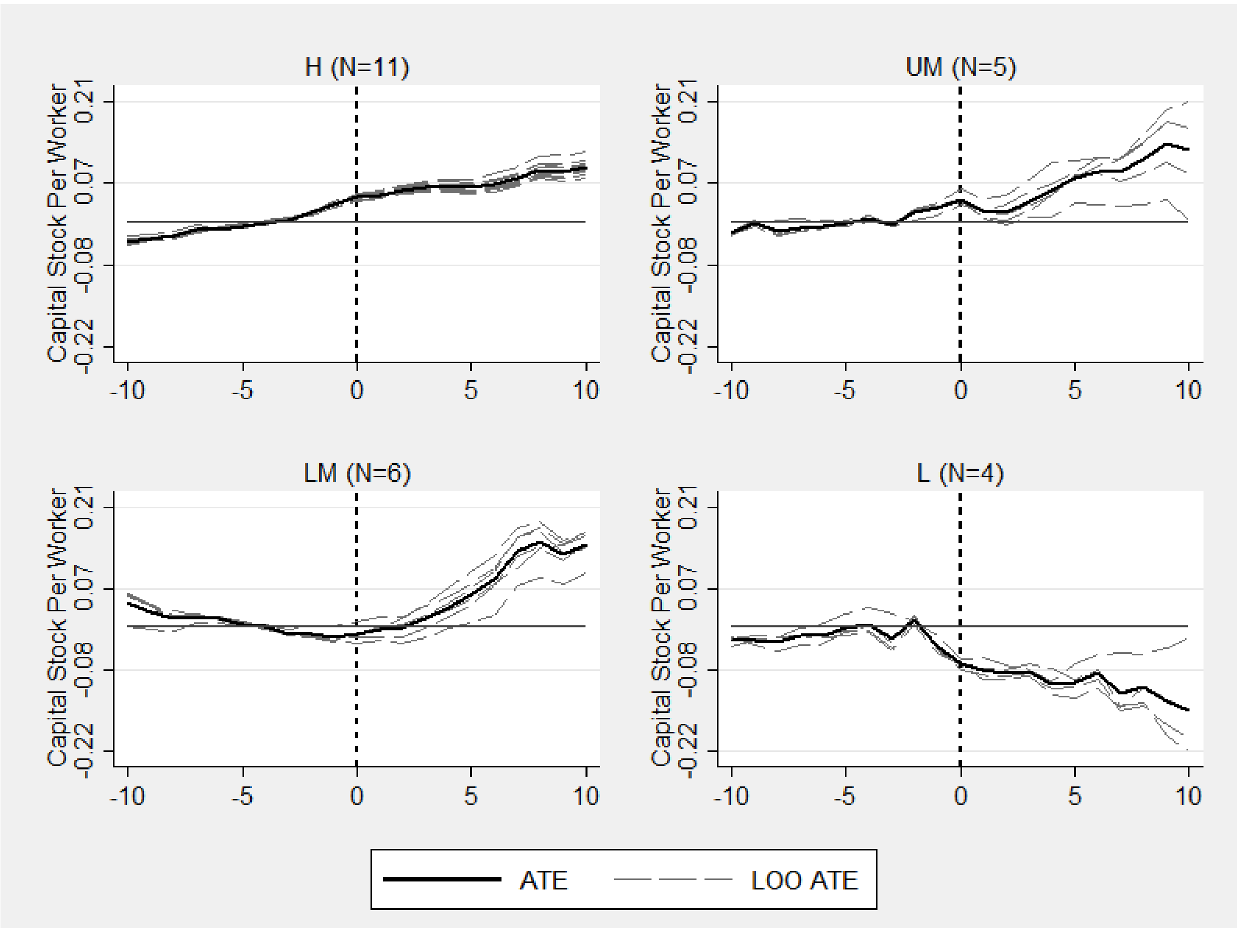
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 9: Average Effect versus Leave-One-Out Average Effects of VAT Adoption on GDP Per Worker Across Different Income Groups



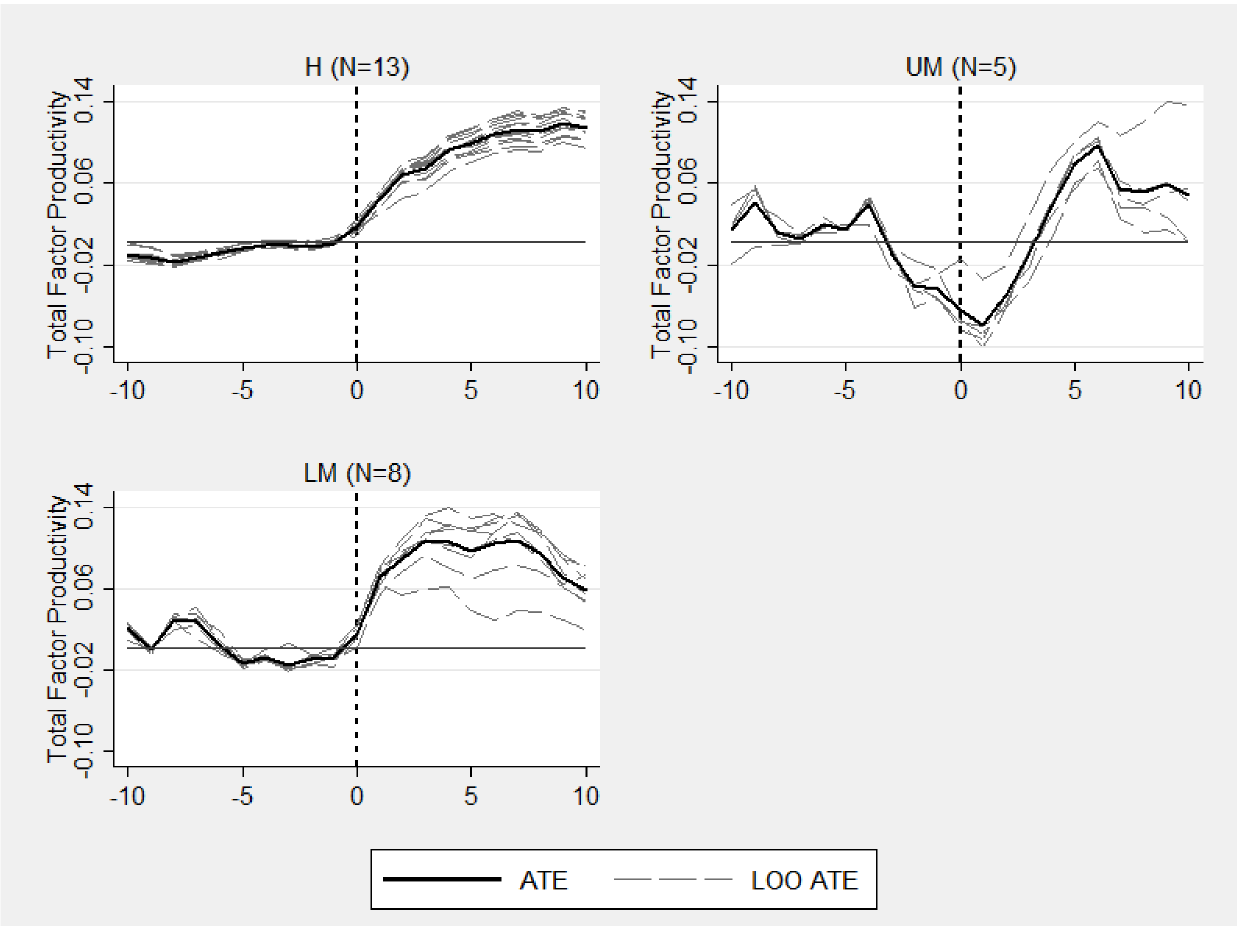
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the difference in the outcome of the group of treated countries and their synthetic controls,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

Figure 10: Average Effect versus Leave-One-Out Average Effects of VAT Adoption on Capital Stock Per Worker Across Different Income Groups



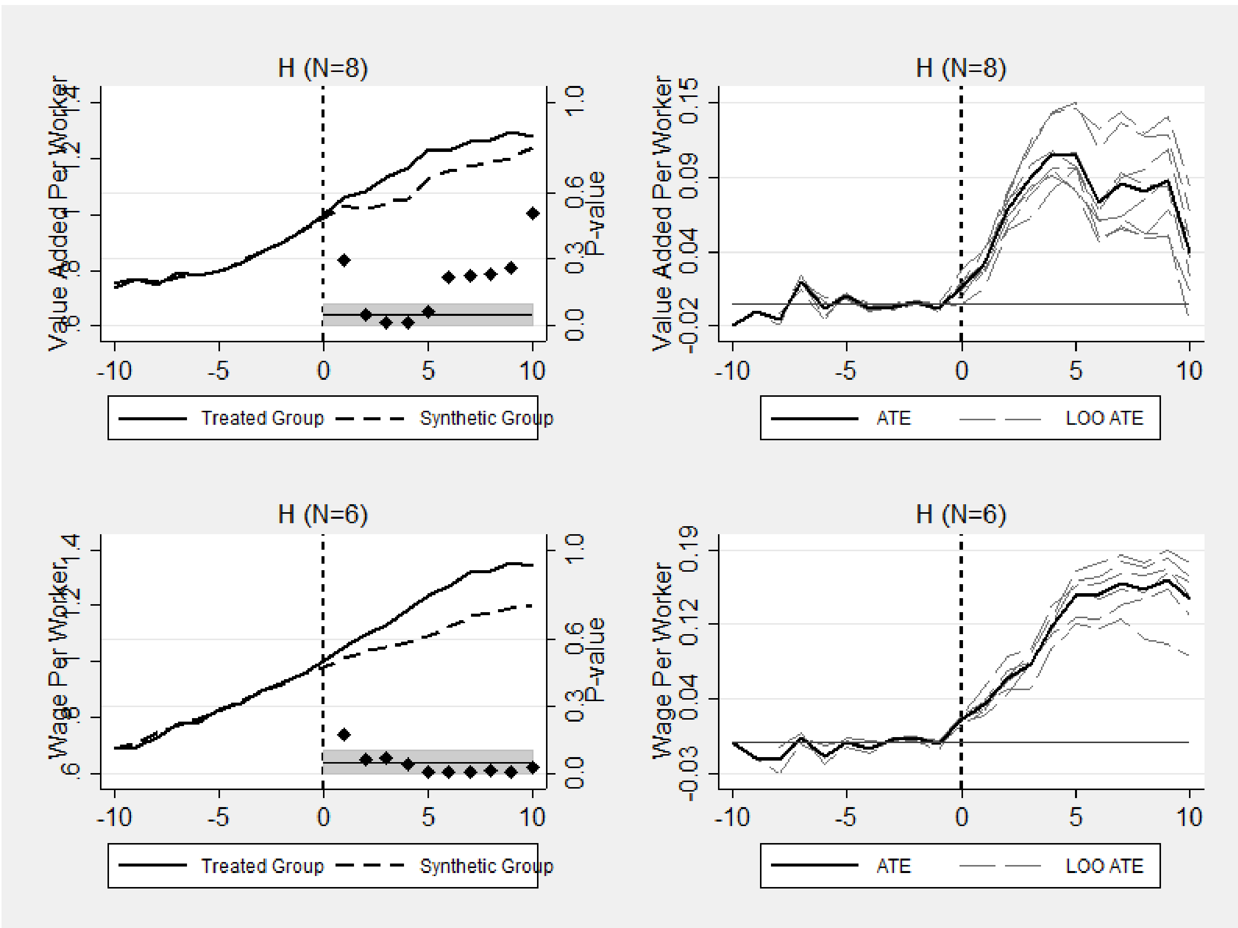
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the difference in the outcome of the group of treated countries and their synthetic controls,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

Figure 11: Average Effect versus Leave-One-Out Average Effects of VAT Adoption on Total Factor Productivity Across Different Income Groups



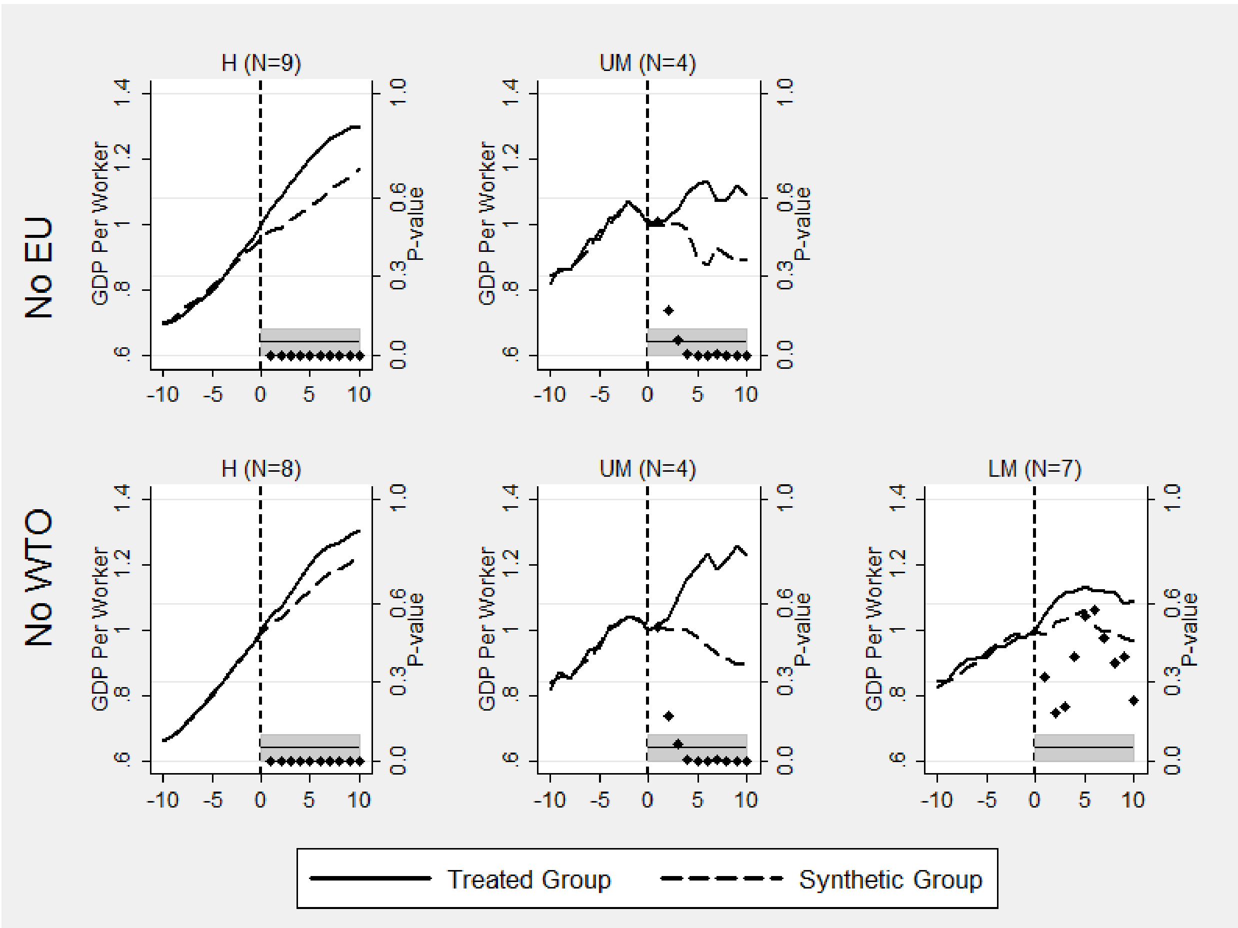
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the difference in the outcome of the group of treated countries and their synthetic controls,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

Figure 12: Average Effect and Leave-One-Out Average Effects of VAT Adoption on Manufacturing Industry Value Added Per Worker and Wage per Worker Across Different Income Groups



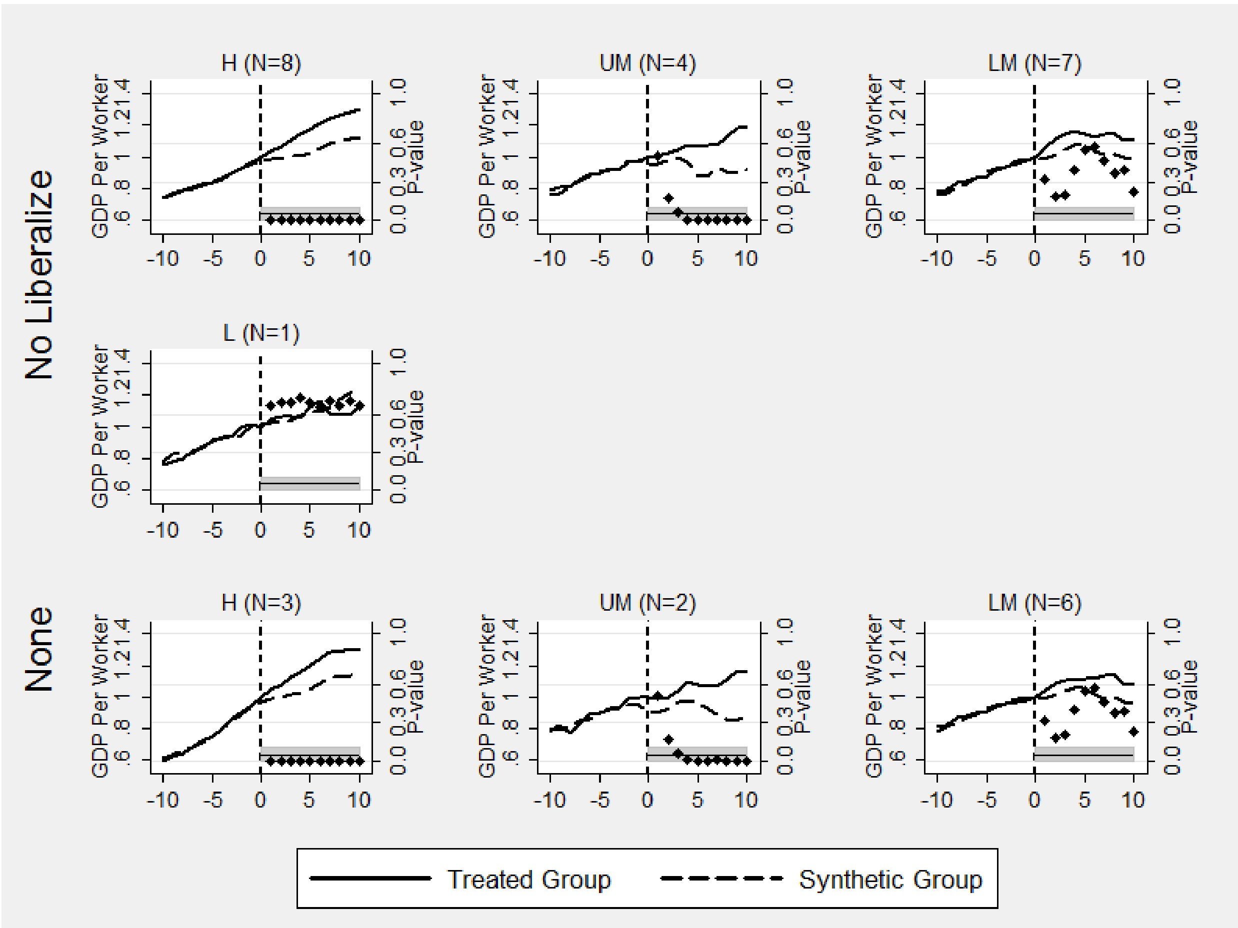
Notes: The results are for high-income countries (H). The outcome variable for the first row is value adder per worker for manufacturing industry, the outcome variable for second row is for wage per worker in manufacturing industry. The first column present the trends in the outcome variables for the treated group and its synthetic control. The second column presents the leave-one-out average, where N denotes the number of countries. 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.

Figure 13: The Impact of VAT Adoption on Average GDP Per Worker Controlling for European Union and World Trade Organization Memberships



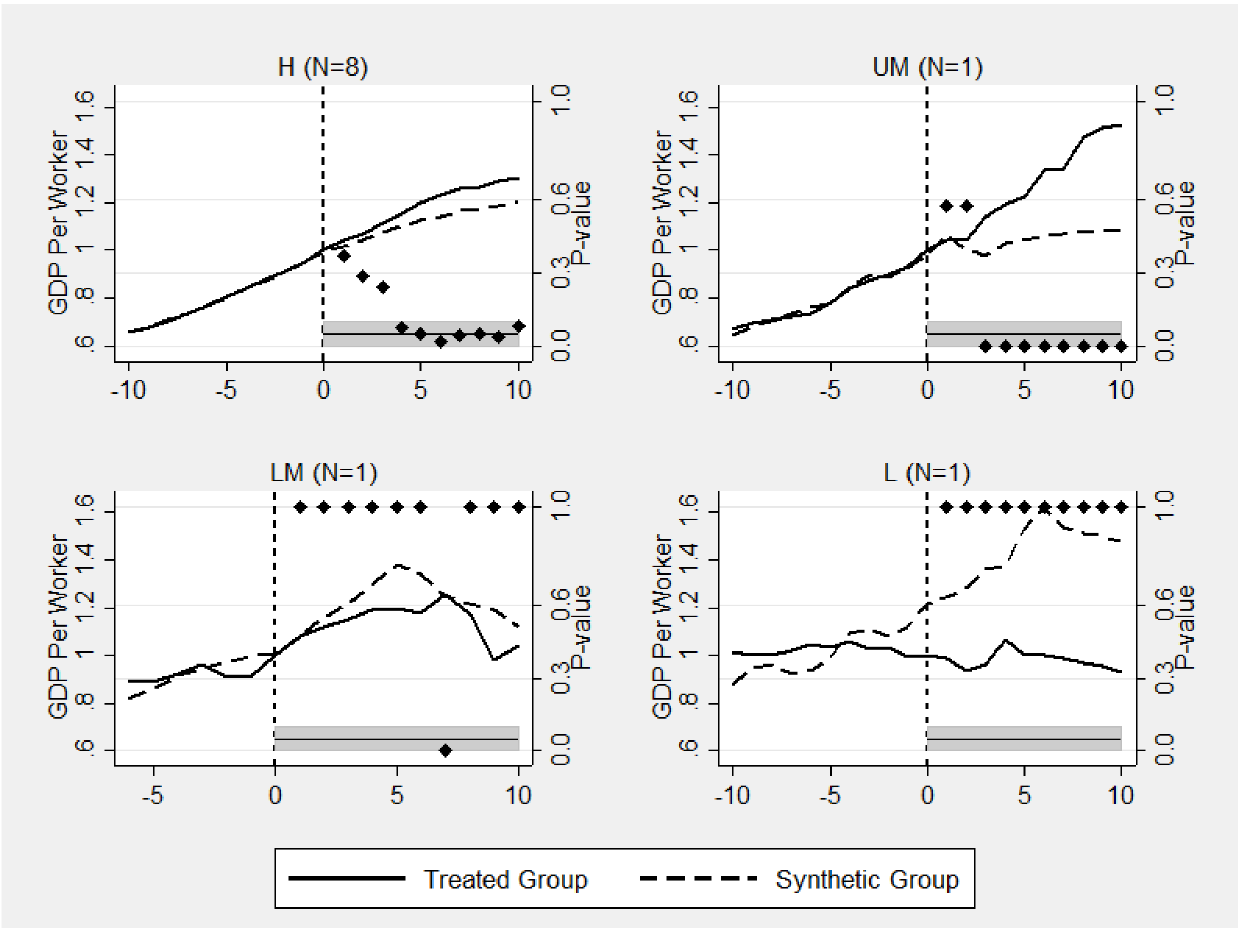
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 14: The Impact of VAT Adoption on Average GDP Per Worker Controlling for Economic Liberalization and All Three Contemporaneous Reforms



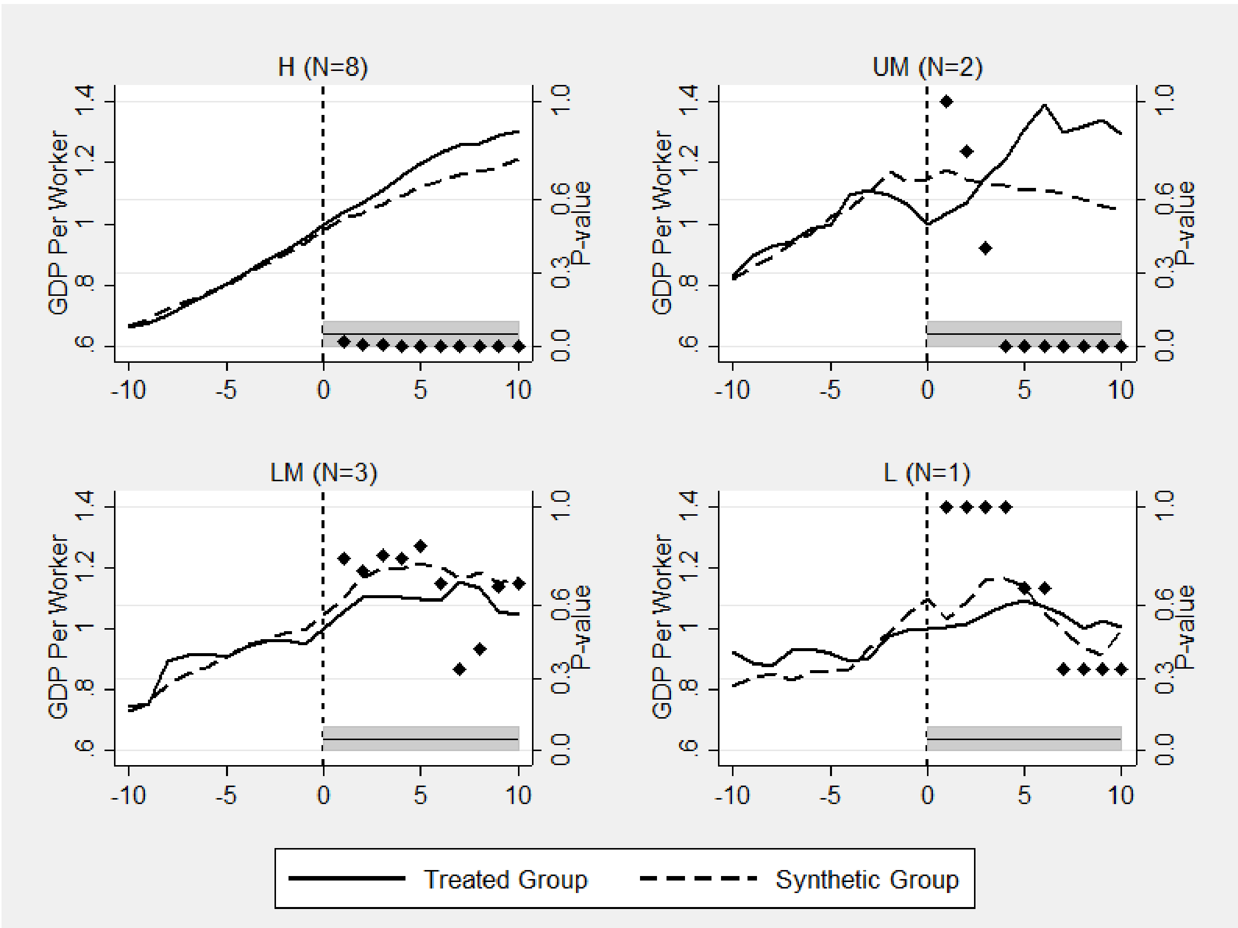
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 15: The Impact of VAT Adoption on Average GDP Per Worker Across Different Income Groups using Control Countries from the Same Region Only



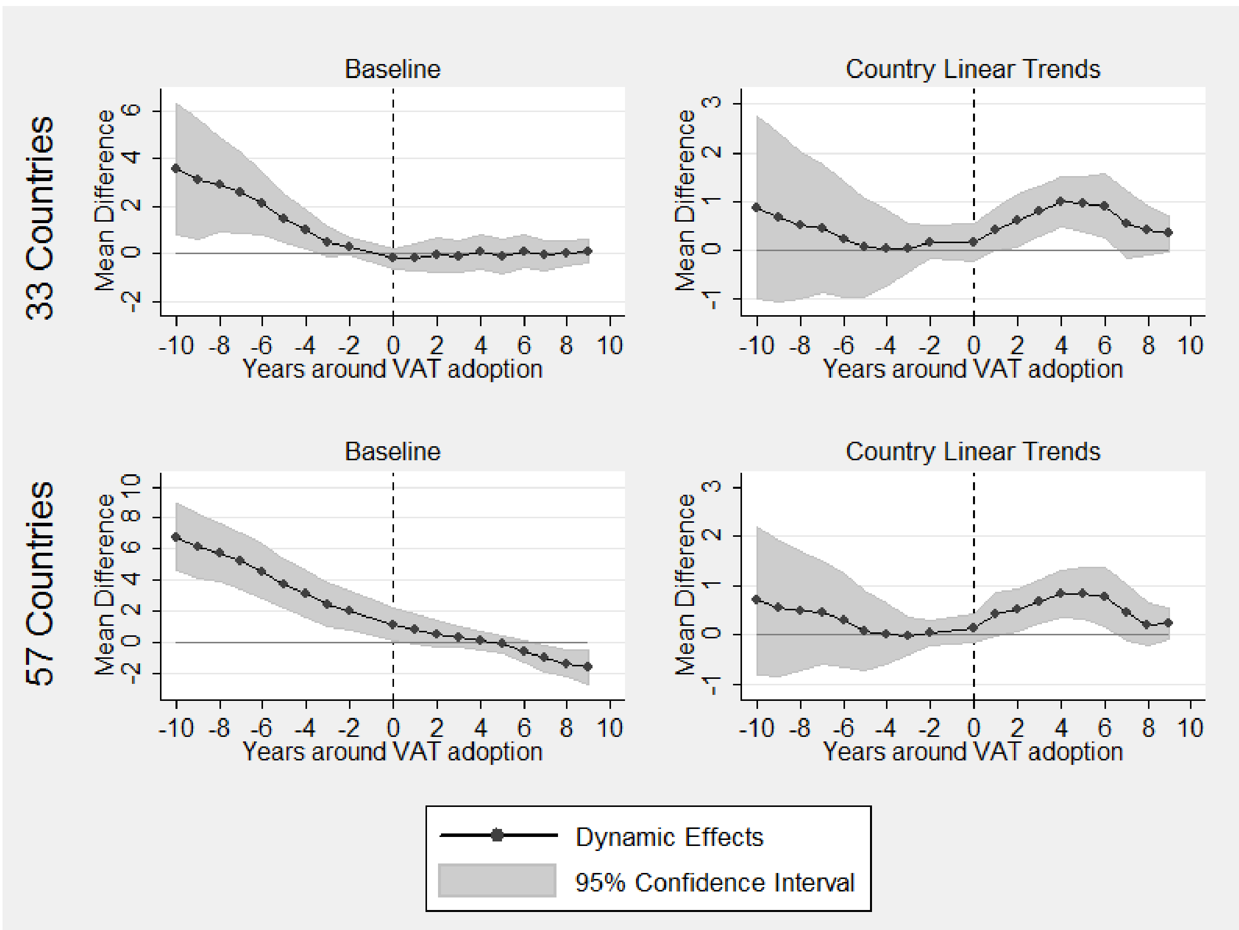
Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). N denotes the number of countries from each income group. The solid black line denotes the outcome of the group of treated countries, and the dashed black line denotes the outcome 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 16: The Impact of VAT Adoption on Average GDP Per Worker Across Different Income Groups using Control Countries from the Same Income Group Only



Notes: The four income groups are H (high-income), UM (upper-middle-income), LM (lower-middle-income), and L (low-income). 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 0.10 is shaded in gray and a horizontal line is drawn to indicate 0.05. 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.

Figure 17: The Impact of VAT Adoption on Average GDP Per Worker Using Conventional Difference-in-Differences Approach



Notes: In the top row 33 treated countries used in SCM are included, while in the bottom row 24 countries that contributed to the construction of the synthetic controls are also included. 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.

Table 1: The Properties of the VAT at the Time of its Introduction in Richer Countries

Country	Year	Coverage	Base	Taxes Replaced
High-Income				
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	multi-rate wholesale tax
Upper-Middle-Income				
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

Notes: The column two (i.e., year) provides information on the year of VAT adoption. In column three (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 four (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 five (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.

Table 2: The Properties of the VAT at the Time of its Introduction in Poorer Countries

Country	Year	Coverage	Base	Taxes Replaced
Lower-Middle-Income				
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
Low-Income				
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: The column two (i.e., year) provides information on the year of VAT adoption. In column three (i.e., coverage): R denotes that VAT is extended through retail stage, W denotes through wholesale stage, and M denotes through manufacturing stage. In column four (i.e., base): G denotes that goods are included the tax base, S denotes services are included in the tax base, ST denotes only selected services are included in the tax base, and CG denotes some capital goods are also included the tax base. In column five (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.

Table 3: The Economic and Political Environment Under Which the VAT is Adopted

Country	**Initial GDP**	**Ratio of Initial GDP**	**Initial Polity2 Score**	**Informal Activity**	**Government Quality**
Austria	33.29	0.60	1.00	9.80	0.93
Belgium	31.52	0.60	1.00	21.90	0.95
Canada	55.24	0.85	1.00	15.70	1.00
Denmark	28.89	0.59	1.00	17.70	1.00
France	30.12	0.60	0.50	15.00	0.92
Greece	31.82	0.51	1.00	27.50	0.66
Italy	32.55	0.59	1.00	27.00	0.74
Japan	44.74	0.70	1.00	11.00	0.91
Netherlands	31.16	0.62	1.00	13.20	1.00
New Zealand	34.13	0.55	1.00	12.40	0.99
Spain	37.93	0.61	1.00	22.50	0.75
Sweden	28.69	0.57	1.00	18.80	1.00
United Kingdom	31.78	0.57	1.00	12.50	0.93
High-Income	**35.00**	**0.62**	**0.96**	**16.46**	**0.93**
Argentina	8.61	0.16	0.60	25.30	0.55
Chile	15.68	0.29	-0.70	19.30	0.61
Ireland	22.13	0.41	1.00	15.80	0.85
Mauritius	28.81	0.38	1.00	22.70	
Panama	18.34	0.33	-0.70	63.50	0.31
Portugal	22.95	0.37	1.00	23.00	0.75
Upper-Middle-Income	**21.19**	**0.35**	**0.46**	**28.16**	**0.62**
Colombia	19.38	0.36	0.80	37.30	0.46
Costa Rica	18.73	0.34	1.00	25.70	0.66
Dominican Republic	12.76	0.22	0.60	31.90	0.52
Honduras	7.73	0.14	-0.10	48.30	0.32
Jamaica	11.04	0.17	1.00	34.80	0.42
Peru	11.34	0.20	-0.70	58.00	0.37
Senegal	3.89	0.07	-0.20	43.70	0.45
Thailand	9.96	0.15	0.90	50.60	0.63
Uruguay	19.39	0.39	0.80	50.60	0.45
Lower-Middle-Income	**12.69**	**0.23**	**0.46**	**42.32**	**0.48**
Bangladesh	4.22	0.06	0.60	35.30	0.20
Guinea	4.80	0.07	-0.10	39.00	0.47
Kenya	4.34	0.07	-0.70	33.20	0.55
Nepal	3.09	0.04	0.50	36.70	
Pakistan	7.54	0.12	0.80	35.70	0.37
Low-Income	**4.80**	**0.07**	**0.22**	**35.98**	**0.40**

Notes: The first column contains GDP per worker (in $1,000s) at the year of VAT introduction, second column normalizes first column using GDP per worker of the U.S., third column contains value of *polity2* at the year of VAT introduction, which ranges from -1 denoting autocracy to 1 denoting democracy, fourth column contains the average share of economy in the informal sector from 1999 to 2007, and fifth column contains the average government quality from 1982 to 1997, which ranges from 0 to 1 and is the average of corruption, bureaucratic quality, and rule of law indices.