

A Service of



Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Agrawal, David R.; Zimmermann, Laura V.

Working Paper The Effects of Adopting a Value Added Tax on Firms

CESifo Working Paper, No. 11469

Provided in Cooperation with: Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Agrawal, David R.; Zimmermann, Laura V. (2024) : The Effects of Adopting a Value Added Tax on Firms, CESifo Working Paper, No. 11469, CESifo GmbH, Munich

This Version is available at: https://hdl.handle.net/10419/308365

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU



The Effects of Adopting a Value Added Tax on Firms

David R. Agrawal, Laura V. Zimmermann



Impressum:

CESifo Working Papers ISSN 2364-1428 (electronic version) Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute Poschingerstr. 5, 81679 Munich, Germany Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de Editor: Clemens Fuest https://www.cesifo.org/en/wp An electronic version of the paper may be downloaded • from the SSRN website: www.SSRN.com

- from the RePEc website: <u>www.RePEc.org</u>
- from the CESifo website: <u>https://www.cesifo.org/en/wp</u>

The Effects of Adopting a Value Added Tax on Firms

Abstract

This paper studies the effects of transitioning from a system of sales taxes to a value-added tax (VAT) on firm-level outcomes. We construct a dataset of product- and state-specific tax rates before and after India gradually switched from a sales tax to a VAT. Exploiting staggered state-level adoptions, we first show that following the transition, effective tax rates declined substantially and complexity as measured by various proxies generally also fell. We then show that sales increased by 57% in the medium run. The reform resulted in increased earnings for workers and higher amounts of capital and digital accounting.

JEL-Codes: H210, H250, H260, H710, O170, O230.

Keywords: value added tax, sales tax, production efficiency, firms.

David R. AgrawalLaura V. ZimmermannDepartment of EconomicsDepartment of EconomicsUniversity of California, Irvine / CA / USA
david.agrawal@uci.eduUniversity of Georgia, Athens / GA / USA
lvzimmer@uga.edu

This Version: October 2024

Agrawal is also a Fellow of CESifo. A previous version of this paper was circulated as 'Production Responses from the Decentralized Adoption of a Value Added Tax in India'. We are grateful to Mohamad Burjak, Ali Chandrani, Nicole Cruikshank, Sanjukta Das, Phillip Jones, Chinmay Korgaonkar, Al Luna, Andres Mira, Bailey Palmer, Madhuree Patel, Rodrigo Saurin, Jawad Ali Shah, Stephanie Stewart, and Marleena Tamminen for their exceptional research assistance in the creation of the tax databases. We thank Bibek Adhekari, Pradeep Agrawal, Michael Best, Brant Callaway, Pinaki Chakraborty, Wei Cui, William Dougan, Ira Gang, John Gardner, Mohammad Hoseini, William Hoyt, Michael Keen, Li Liu, Mahesh Purohit, Rasoul Ramazani, Johannes Spinnewijn, Sandip Sukhtankar, Tejaswi Velayudhan, Mazhar Waseem, Caroline Weber, and David Wildasin for helpful comments along with anonymous officials in the Delhi GST Commissionerate. We thank seminar participants at Clemson University and IIM Lucknow. Financial support from the International Growth Centre (IGC) and the University of Kentucky Kalam Grant is gratefully acknowledged. Laura Zimmermann also acknowledges financial support through the Terry-Sanford and Bonbright Center Research Award at the University of Georgia.

1 Introduction

While many of the early adopters of the value-added tax (VAT) were European countries, most recent adopters have been developing countries that switched from sales taxes (ST). Although there is a literature that studies the trade-off between production and evasion responses to value added taxes (Best et al. 2015), there is very little evidence of the effect of the transition from a sales tax to a VAT on firms. In this paper, we study how the transition to a value added tax can benefit firms in a developing country. One of the countries that recently adopted a VAT is India, which gradually moved from a sales tax to a VAT system. To what extent did the transition allow firms to produce more goods, increase earnings, invest in more capital, and eliminate some of the frictions long regarded to constrain firms?

Both before and after the transition, each Indian state levied commodity-specific tax rates on hundreds/thousands of itemized products. Tax rate information is published in "schedules," which are long lists of commodities with their corresponding tax rates. Product descriptions and categories also change over time. We construct a comprehensive harmonized dataset of product-specific and state-specific tax rates before and after the reform, which allows us to identify tax changes at the firm-level. Using published tax laws and official online archives of tax change notifications, we digitize thousands of pages of tax rules. We thus know tax rates on every product, in every state, in every post-reform year and in select pre-reform years. We then harmonize the products over time and across states.

We combine these tax data with firm-level data from the Annual Survey of Industries (ASI). The ASI provides representative repeated cross-sections of firms, focusing predominantly on manufacturing firms, broadly defined. The dataset also includes the precise products each firm produces and uses as material inputs. As the tax schedules do not conform with standard product codes, we match the tax rates to the ASI dataset using a multi-step process involving fuzzy-matching algorithms. The linkage of inputs and outputs in the ASI then allows us to create firm-specific effective tax rates before and after the switch to VAT.

Using our tax dataset, we first document the effect of the reform on tax complexity

and tax rates. These stylized facts help inform researchers about what the reform actually achieved, and how the Indian fiscal system varies across place and time. First, although the number of tax rate schedules declined, the products on these schedules became more detailed. Using these variables and other proxies for complexity, we construct a complexity index. This index declines following the reform. Next, we document that the adoption of VAT led to lower statutory tax rates in most states, but this effect was heterogeneous across firms even within the same industry. Matching inputs and outputs in the ASI allows us to document, under some assumptions, that firm-specific tax rates also fell substantially.

Given the gradual implementation of VAT between 2003 and 2008, we use a staggered difference-in-differences approach (Gardner, 2021). We study a variety of outcomes, focusing first on total sales and material input usage by the firm. The estimated effective tax rate changes are then used to translate the quantity responses to elasticities. Finally, we explore reasons why firm output grew in response to the reform by studying the effect on labor earnings, capital investment, and technology usage by the firm. The wide range of outcomes that we study paints a comprehensive assessment of how the shift to a VAT affects firms.

We find a 57% increase in firm output by three years after the tax reform. This translates into an elasticity of -0.88 with respect to the effective tax rate. This increase in output is coupled with a similar increase in material inputs, likely due to the reduction of taxing business inputs. Firms that realized the largest tax rate decreases see the largest absolute production responses and changes in inputs. But among the set of firms that saw large tax decreases, the responses are smaller and more delayed when the complexity index indicates that the VAT system remained highly complex. Thus, tax rates mattered, but tax changes also interact with tax design and imperfections of the VAT system. We then explore how the reform influenced the allocation of resources within the firm including what inputs to use. We find that the shift to VAT—and the overall lower tax rates that came with the reform increased the earnings of workers and the amount of capital/land used in production.

Although the large and salient reform led to big effects, there are reasons why it could have

fallen flat. In a "textbook" VAT, the tax is broad-based, features a single tax rate, allows for full refunds on input purchases, and is implemented at the national level. India's adoption of VAT was arguably its largest reform affecting firms, but it was imperfect. The VAT system was decentralized, featured many tax differences across commodities within a state, was not broad-based because it excluded services, did not eliminate double taxation, had incomplete refunds and credits, and did not resolve issues regarding interstate transactions. Slemrod and Velayudhan (2022) write "a research agenda for the VAT should consider the efficiency...implications of real-life, imperfect VATs".¹ We provide initial evidence along this dimension by showing how VAT complexity dampens the effects. Although India may have been unable to take advantage of the self-enforcing properties of the VAT, incentives to evade were likely stronger under the sales tax.² If firms report the same information in the ASI as on tax returns, some of our response could then be a compliance effect.

This paper contributes to our understanding of the costs and benefits of major public reforms in developing countries. The VAT reform was one of the most important recent reforms affecting firms in India—mainly because it affected a large share of economic activity. Many other policy reforms studied in India focus on tariffs or trade liberalization (Goldberg et al., 2010b,a; Topalova and Khandelwal, 2011), which necessitate studying particular industries. In contrast, this reform heterogeneously affected all industries and firms.

Second, the transition from a sales tax to a VAT has been implemented by many developing countries in the last three decades based on the theoretical prediction that it would remove inefficiencies (Adhikari, 2020). But evaluations of its effects are scarce, perhaps because it is difficult to find credible research designs, as the reforms are usually implemented at the national level. The case of India is thus a great opportunity. Two exceptions are Keen and Lockwood (2010), who focus on aggregate data, and Smart and Bird (2009), who study the Canadian transition. In the context of India, Hoseini and Briand (2020) study the same reform as we do, but only focus on the effect on supply chain linkages and do not exploit tax-rate information. We study a wide range of outcomes and provide evidence on whether those theoretical predictions are realized by firms. Although a number of features of the Indian context are different from other countries, our study provides evidence that tax reforms can have important benefits in developing and emerging markets.

2 India's Commodity Tax System

India has long had a complex commodity-tax system (Bagchi and Team, 1994; Chelliah and Committee, 1992). State governments have the power to levy taxes on intra-state sales, whereas inter-state sales are taxed by the central government.³ Services cannot be taxed by the states and instead were taxed under the Central Value Added Tax (CENVAT). Given the difficulty of enforcing direct taxes in a low-capacity environment, indirect taxes have been an important source of revenue, making up 61 percent of states' own-tax revenue (Sury, 2015).

2.1 The Pre-Reform Sales Tax System

Most states adopted independent sales tax systems in the 1940s or 1950s. Unlike most other commodity tax systems around the world, the Indian sales tax system was distinguished by its rate schedules, which featured a large number of product-specific tax rates within a given state. States had detailed schedules with product-specific tax rates for hundreds or thousands of itemized products.⁴ While goods considered to be necessities were mostly exempt or subject to lower tax rates than other products and luxuries, the same good was often taxed differently across Indian states. All goods not listed under a specific schedule are subject to a "default tax rate", which was 8 to 10 percent in many states (Purohit, 2001a).

Unlike the retail sales tax in the United States, most states levied sales tax at the first point of sale in their state, defined as the sale by the first registered dealer in the state. Each state had its own rules guiding which dealers had to formally register and pay taxes, typically based on turnover, but also often subject to additional criteria such as industry. Only a few states had a last-point tax, which taxes the sale by the last registered dealer either to an unregistered dealer or to a consumer (Purohit, 2001a,b). By the 2000s the state sales tax systems differed widely, due to a mixture of administrative convenience and ad hoc measures to increase revenues (Sury, 2015). In addition to its complexity, problems with tax cascading and tax competition between states became widespread. The first point of sale required very high rates to raise the same amount of revenue as would be generated under an equivalent retail sales tax. This narrow base and arguments over what the first point of sale was, led many states to tax inputs in the production process, which in turn were generally denied credits against the manufacturing stage resulting in the double taxation of industrial inputs and machinery. According to the Bagchi report, created by Indian tax experts, inputs were often subject to their own tax schedule, were denied credits on technical grounds, turnover taxes were not deductible, and no relief was provided to taxes on plants and equipment (Bagchi and Team, 1994).

The double taxation of inputs and outputs resulted in tax cascading, leading to effective tax rates that were well in excess of statutory rates. Moreover, taxation at the first point of sale created production inefficiencies by encouraging firms to shift production to later stages in the supply chain in order to reduce the tax burden. The first-point-of-sale tax also resulted in high tax rates, which introduced reasons to shift production to a lower taxed but less efficient production process. These factors resulted in a tax system with large productive inefficiencies, which in turn have important welfare implications (Keen 2012; Keen 2014).

With respect to tax competition, time-, industry- and location-specific tax incentives were common, and some states engaged in "rate wars" by lowering tax rates to attract firms to their state (Sury, 2015). At a time of increasing state fiscal deficits, this led to large tax revenue losses among states (Ministry of Finance, 2003). This incentivized states to create new commodity taxes to raise additional revenue, but those led to further double taxation. Bagchi and Team (1994) refer to the commodity tax system as "archaic, irrational and complex—according to knowledgeable experts the most complex in the world". The Bagchi report recommended the switch from a state sales tax to a state VAT regime as the most realistic step towards a more efficient commodity tax system.

2.2 The VAT Reform Process and the Post-Reform System

Implementing a switch to a VAT required getting state governments on board with suggested changes, because many states were highly protective of their right to levy taxes (Purohit, 2001b). While state governments acknowledged the need for harmonization of their widely differing systems, finding an agreement proved to be difficult in part because states were worried about the revenue uncertainty from the move to a completely new tax system.

These issues led to repeated delays in the process. In 2002, the central government reduced its demands for harmonization across the states. It now sought to harmonize a few key features of the VAT system by proposing a model VAT act, while allowing state governments freedom to adjust the model act to state-specific concerns. The central government also agreed to compensate states for any reform-related losses in the first three years (Ministry of Finance, 2002, 2003, Sury 2015). April 1, 2003 was chosen as the VAT start date, but only one state, Haryana, actually switched to the VAT on that date. As researchers (Hoseini and Briand, 2020) and the central government (Ministry of Finance, 2005) note, the other states delayed their start dates due to quasi-random factors, including the pre-set timing of elections and the intense legislative processes required for the reform. Many states introduced the VAT system on April 1, 2005, but in several states the VAT adoption took until 2006, 2007 or 2008. Appendix Table C.1 lists the implementation dates.

The VAT system in every state was established in a state-specific VAT Act, passed in the state legislature. Like the sales tax, the Act included schedules with lists of goods and their tax rates. Most states had an exempt category, as well as tax rates of 1, 4 and a general tax rate of 12.5 percent. But many states also levied higher tax rates on items such as diesel and alcohol. Due to the model act, most states therefore saw a reduction in the complexity of their tax system as measured by the number of tax rates, as well as some harmonization of tax rates with their neighbors. But tax rate changes after the VAT introduction also potentially led to a new divergence of state tax rates from tax competition. We explore the time path of tax complexity and rates with our tax database below.

Despite the initial fears of a transition to VAT, tax revenue increased substantially after the reform, and the central government called the implementation "quite encouraging." Tax revenue for the financial year 2005-06 increased by 13.8 percent, with even higher increases in subsequent years (Ministry of Finance, 2007, 2008, 2009). Only eight states requested compensation for losses in 2005-2006. Despite its revenue success, the new system retained a number of shortcomings, such as many product-specific tax schedules, a lack of harmonization across states, and the complete non-taxation of services, leading to calls for further reform. In addition, although tax cascading was reduced by the reform, it was not entirely eliminated due to delays in credits against the CENVAT, the exclusion of major sectors from the CENVAT and denial of these sectors to credits on state VAT as well as the inability to credit the Central Sales Tax on interstate trade (Keen 2014). Thus, the overall effect of the VAT reform on output remains an empirical question despite anecdotal evidence that the transition was an improvement of the tax system.

2.3 The Tax Database

Documenting the effect of the reform on output first requires knowing the effect of the reform on the tax system. Yet, no standardized database exists for all of India's state-specific tax rates. We assemble the first harmonized database, which creates a standardized listing of tax rates by products for all states, a database of registration thresholds for firms, and a database of proxies of tax complexity. All states provide information on their state-specific VAT systems on the website of their Commercial Taxes Department. This information typically includes the state's VAT Act, a list of tax schedules with specific products, and an archive of tax change notifications.⁵ The states do not maintain a database that provides a panel of tax rates over time, and the exact version of the VAT Act is sometimes unknown.⁶

We first digitize the tax rate schedules for each state for a given point in time, which gives us a cross-sectional snapshot of the product-specific VAT rates. We then trace any tax rate changes from information on the schedules or from the notifications in the online archive, and note the old and new tax rates as well as the exact date the changes went into effect. We then harmonize the tax rate information over time, including when product descriptions are split, combined or otherwise altered over time. Products may also appear or disappear from the schedules as they switch out of or into the default tax rate. While all commodities that do not fall under the default tax rate are explicitly listed, any remaining products are not mentioned in the schedules. In our harmonized dataset, the number of products and their exact description track every product that ever appears in a list for the entire time range. This gives us a state-product panel database from the VAT introduction until 2016, although the event study design will only exploit information 4 years after each reform.

We supplement these databases with information from two collected volumes of state VAT systems, one from 2005 and one from 2007 (Sangal, 2005; Sangal and Goel, 2007). These books contain snapshots of the VAT systems of any states that had implemented the reform by the publication date, as well as details on all notifications since the VAT introduction. They allow us to cross-check the tax rates and registration thresholds.

Information on the state-specific sales tax regimes unfortunately is typically not available from online sources in any complete or consistent form. Similar to the VAT regime, we therefore rely on two publications containing multiple volumes of state documents and tax rate schedules from 2000 and 2002 (Commercial Law Publishers 2000; Commercial Law Publishers 2002). We digitize this information to get a database of prevailing sales tax rates in 2000 and 2002, before any state implemented the tax reform. Unfortunately, the publications are missing for some of the small states or Union Territories, but this does not affect many firms in our database.⁷ Since both the ST and VAT schedules do not contain product codes, we harmonize the sales and VAT lists over time using a variety of fuzzy matching algorithms and manual matching of product names. See Appendix A for details.

The VAT Act and sales tax documents also contain information regarding state-specific registration thresholds for firms. We code up all general thresholds as well as any additional rules, such as different thresholds for specific industries or firm characteristics. In addition to the VAT registration thresholds, we track turnover tax thresholds for which firms can pay a (generally) lower rate on their turnover rather than the VAT. We also track other thresholds for which firms are required to pay additional fees or other taxes. Like the tax rate changes, we record any changes to the registration rules over time. Although these thresholds will not be the main subject of our analysis, our data collection of these thresholds make one thing clear: thresholds are higher under the VAT regime than under the sales tax regime. We also digitize proxies for complexity such as the number of pages in the tax code.

We create the first extensive panel database of state-specific commodity tax rates.

2.3.1 A Snapshot of Tax Rates and Tax Categories

Our constructed tax databases give us the unique opportunity to systematically explore tax rates and the complexity of the Indian tax system before and after the reform.

Figures 1a and 1b show two possible proxies for complexity: the number of tax rate schedules and the number of specific products itemized. The number of tax rates/products under the sales tax system is depicted by red diamond symbols. The number of tax rates/products in the first year of the VAT system, which depends on each state's implementation start date, is shown in orange circles, and the tax rates 4 years after the reform, the last year for which we will be able to report event study coefficients, are plotted as green triangles.

Under the ST system, states had anywhere between 6 and 23 different tax rates, with Bihar and West Bengal being the most complex by this metric. Almost half of the states in the sample have more than 10 tax rates. Model VAT acts that attempted to induce conformity across states implied that the transition to VAT reduced the number of tax rates to generally between four and ten rates. But this was not the case everywhere: the VAT switch allowed some states with few tax groups to use the reform as a policy window to increase the differentiation in tax rates. Figure 1a shows, however, that the number of rate schedules under the VAT system generally featured few changes after the initial reforms.

With few exceptions, the number of products itemized in the schedules increased sub-

stantially immediately after the VAT reform (Figure 1b). Thus, the VAT induced a smaller number of tax rates, but more specific listing of products. On the one hand, the increase in the number of products could be viewed as a decline in complexity as it created less ambiguity over the tax rate that a product was subject to. On the other hand, increasing the number of itemized products could be viewed as an increase in complexity if firms need to search through more laws or notifications to identify the appropriate tax rate.

State tax rates vary dramatically. Figure 1c shows changes in the average tax rate of all the goods specified in the schedules over time for the same three time periods. The calculation keeps the number of products under the VAT regime constant, so any tax rate changes over time are not driven by changes in product categories. Most states saw a decrease in the average tax rate with the introduction of the VAT, from an average (statutory) sales tax rate of 6.49% across states to a mean VAT rate of 5.87%. By four years after the reform, the average tax rate fell further in many states. The big outlier is Tamil Nadu, which levies extremely high tax rates on alcohol, leading to very high average tax rates. Because the average tax rate is calculated over all the goods listed on the schedule, the average gives more weight to more detailed schedules. This is potentially misleading because all states have a default tax rate that applies to all goods that are not explicitly mentioned in any of the schedules. This tax rate is just one entry on the list, but may apply to many products.

Figure 1d plots the default rate separately as well. It highlights the role of a centralized model act, which induced near uniformity in the default tax rate across states after VAT adoption. Although a centralized model act was designed to elicit tax harmonization across states, the share of products subject to this rate declined after adoption (Appendix Figure C.3). As the tax schedule became more detailed in the products listed, more products came out of the default category, possibly mitigating the ability of the model act to compensate imperfections in the VAT from multiple rates. The most extreme example is Karnataka, which had approximately 50% of products (weighted by sales) subject to the default rate under the sales tax regime, but had less than 10% subject to it four years after the reform.

2.3.2 Thresholds

In addition to studying tax rate changes and changes in complexity, it might also be useful to know how tax evasion changes as registration thresholds change. While bunching at the threshold may pick up both real responses and evasion responses, any bunching mostly identifies evasion. The reason is similar to Best et al. (2015): below the threshold firms face a very low rate turnover tax while above the tax they face the generally higher tax rate.⁸

However, we are unable to identify any bunching—even descriptively—except at the very highest of thresholds across states. Unfortunately, the share of firms in our dataset near the threshold is limited, given that the ASI generally includes larger rather than smaller firms. Thus, the ASI is not the ideal dataset for answering this question. For this reason our subsequent empirical models will focus on the effect of tax rate and complexity changes on firm behavior. In our main analysis, we include both exempt and non-exempt firms.

3 Empirical Methods

3.1 Firm Data and Matching with Tax Information

The firm data comes from the Annual Survey of Industries (ASI), a well-known survey of establishment-level data from Indian manufacturing firms (e.g., Allcott et al. 2016; Hsieh and Klenow 2014). India generally defines "manufacturing" much more broadly than other countries, so the database contains firms producing a wide variety of products at various stages. The ASI is widely considered to be the best source of high-quality data on firms in India. We use the version of the database that contains firm identifiers from the years 1998-99 to the year 2012-13, but given the data are representative within state-year we exploit the repeated cross-section feature to account for the fact that not all firms are surveyed every year. Large firms with more than 100 employees (more than 200 employees for 1998-1999) are interviewed every year, whereas smaller firms are sampled every three to five years.⁹ Sample weights make observations in any given year representative of the state distribution of firms registered under the Factories Act. This excludes small firms with fewer than 10 workers (or fewer than 20 workers for factories that do not use electricity), which are not required to register and therefore are not in our dataset. While this is a well-known limitation of the database, excluding about 80 percent of the workforce in the manufacturing sector (Hsieh and Klenow, 2009), these small firms contribute only about one third of total output in the sector (Sincavage et al., 2010). The survey contains information on sales revenue, employment and input costs, as well as detailed industry identifiers. Firms are asked to list the specific goods used as main inputs as well as the products they produce.

This detailed information on firm inputs and outputs makes it possible to merge tax rate information from our created databases. Because the sales tax and VAT schedules do not conform to any product classification system and may contain state-specific products and spelling differences, we have to use name matching algorithms as well as extensive manual matching of firm information and tax databases. To do this, we first use fuzzy match algorithms to match products in the tax data to the firm data. This may yield false matches or non-matches, which we best resolve by manually checking every matched item and then manually matching all remaining products. See Appendix A for details.

3.2 Empirical Estimation Method

We use temporal and cross-state variation to study production responses from the tax system shift. We exploit the staggered implementation of the VAT across states in India to implement a difference-in-differences design (DID) relying on variation in treatment timing across the states. With such a large reform, it is reasonable to assume that there are dynamic treatment effects, for example due to learning about the tax law and a compounding growth effect. The main focus of our analysis is therefore on an event study setup.

To avoid the identification concerns that have been raised about two-way fixed effects (TWFE) specifications in applications like ours, we use the estimator proposed by Gardner (2021). It consists of an intuitive two-step procedure and is based on the observation that

the bias in TWFE specifications arises from estimates of the fixed effects. Appropriately estimating the fixed effects therefore allows for causal ATT estimates. To do this, in stage 1, the outcome variable of interest is regressed on time and state fixed effects for units that have *not yet* been treated. This amounts to estimating:

$$z_{f(s),t} = \nu_t + \zeta_s + \varepsilon_{f(s),t}.$$
(1)

where $z_{f(s),t}$ is an outcome of interest for firm f, which is located in state s, and observed at time t. Our main outcomes of interest are sales and the value of inputs, which are deflated using state-industry-time specific price indexes. Let ν_t be time fixed effects and ζ_s be state fixed effects. The ASI dataset surveys a representative unbalanced sample of manufacturing firms in each state in every year. We use sampling weights in all of our specifications, which allow us to treat the data as a representative repeated cross-section of firms in each state.

After estimating (1), the outcome variable for the entire sample is then residualized. In stage 2 of the Gardner approach, the residualized outcome variable can then be regressed on a set of event-time dummies to obtain unbiased coefficients. Standard errors in this approach are corrected for the fact that stage 2 contains a generated regressor. We estimate:

$$\widehat{\varepsilon}_{f(s),t} = \sum_{y=-5}^{-2} \pi_y \mathbf{1}\{t - t_s^* = y\} + \sum_{y=0}^{3} \gamma_y \mathbf{1}\{t - t_s^* = y\} + \eta_{f(s),t},\tag{2}$$

where $\hat{\varepsilon}_{f(s),t} = z_{f(s),t} - \hat{\nu}_t - \hat{\zeta}_s$ is the residualized outcome variable of interest and hats denote estimated fixed effects. The indicator variables $\mathbf{1}\{t - t_s^* = y\}$ measure time relative to the date of VAT adoption t_s^* (i.e., y is event time). Thus, π_y represents the differential evolution of $z_{f(s),t}$ in the treated states in the years prior to the reform while γ_y represent the differential evolution of $z_{f(s),t}$ for the years after the reform. The year (immediately) prior to the event is omitted from the regression so all coefficients are relative to that year.¹⁰ Estimating the overall ATT using a simple treatment indicator interacted with a post-reform indicator proceeds analogously, which we present for the main outcomes of interest.¹¹ Data in the ASI are based on firm financial years which start in April. We classify the omitted year -1 as the financial year prior to the VAT reform where the *entire* financial year was subject to the sales tax regime. Thus, the first year of the reform, event year 0, includes some states that switched to the VAT at the start of the financial year (and are treated for the full year) and some states that switched to the VAT later in the financial year (and are partially treated). We do this so that the omitted year is not contaminated by partial treatment; this implies that event year 0 may show a smaller response partially because some firms are only exposed to the VAT reform for a limited number of months.

Furthermore, as we do not have any never-treated states in continental India, we focus on a time interval relative to the treatment event that allows us to always rely on not yet treated observations as the control group (Malkova, 2018). This includes 5 pre-treatment periods and four post treatment periods. However, as we show in the results section, our production results are robust to including the never treated observations.¹² Given the limited number of never treated observations, and because many states adopt within a two-year period, the long-run effects are identified relative to a small number of states. Standard errors are clustered at the state level as treatment status occurs at this level.

To yield causal estimates, the parallel trends assumption needs to be satisfied, potentially conditional on covariates, and there should be no anticipation effect. One way of satisfying these assumptions is for the event timing to be (conditionally) random. As discussed in the institutional details section, there were many haphazard reasons why some states adopted the VAT earlier or later and it is unlikely that firms anticipated the reform date.¹³ To further test the plausibility of the assumption of parallel trends, following Bailey and Goodman-Bacon (2015), we regress economic and political variables on date of reform adoption to see whether they are systematically correlated. We find no significant evidence that variables like GDP growth, population, state tax revenue, civil unrest, or the nature of the political alignment with the central government explain the timing (Appendix Table C.2). Critically, the only correlates related to adoption concern election timing and the political shares of

parties in power.¹⁴ As Hoseini and Briand (2020) note, the timing of elections was predetermined in that time period and quasi-random, leading to a political incentive for state governments run by the political opposition to adopt right before an election to blame the national government for any issues, while other state governments had an incentive to adopt right after an election. Appendix Figure C.4 shows that our results are robust to including the socio-economic and political covariates in our empirical design.

Another concern is that adoption timing is correlated with levels or pre-reform trends of our outcome variable, but we do not find a significant correlation (Appendix Table C.3).

4 The Effect of the VAT Reform on Complexity and Rates

The introduction of the VAT attempted to reduce complexity and reform tax rates, but it is unknown if this worked. We provide causal evidence of the effect of the reform on complexity and tax rates to shed light on mechanisms influencing our subsequent reduced-form estimates.

4.1 Complexity

Figure 2 shows DID event studies for the number of tax rate categories and the number of itemized products, each of which proxies for complexity. Recall that we focus on four years after because that is the maximum length we can document given adoption times. The figures formalize the prior qualitative analysis: the number of tax rate categories declines by four on average (30%), while the number of items increases by approximately 100 (20%).

The multidimensional nature of tax complexity makes using a single proxy for it of limited value. Thus, next, we consider a wide range of proxies for complexity including the number of tax schedules, the number of items, the number of pages in the law, the number of subsections of the law, the share of items in the default category, and the number of schedules and products that change tax rates. One might be tempted to study the effect of each of these variables individually, but such an approach is then challenging to interpret. Each variable may affect complexity differently. Instead, we take a data-driven approach that utilizes all of the information available from these proxy variables.

While there are many other data-driven approaches, we rely on Lubotsky and Wittenberg (2006) who develop an index creation procedure, which minimizes biases from measurement error, in the presence of multiple proxy variables. Following their approach allows us to construct a single proxy for complexity, $c_{f(s),t}^{\rho}$:

$$c_{f(s),t}^{\rho} = \left[\sum_{k=1}^{K} \beta^{k} \frac{cov(z_{f(s),t}, c_{s,t}^{k})}{cov(z_{f(s),t}, c_{s,t}^{1})}\right]^{-1} \sum_{k=1}^{K} \beta^{k} c_{s,t}^{k},$$
(3)

where $c_{s,t}^k$ denote each of the K proxies for complexity. Each of the β^k are the coefficients from a fixed effect regression of $z_{f(s),t}$ on all of our proxy variables. Further, $c_{s,t}^1$ is the (standardized) variable for the number of tax rates, which we use to normalize the index so that changes in it can be interpreted in standard deviations of that variable. Further details of the index creation procedure are found in Appendix B.1. The construction of an index allows us to address the fact that single proxy variables may explain different aspects of complexity and may change complexity differently. In our setting the use of the index does not guarantee it is a "perfect" measure of complexity since it can only capture information contained in the proxies we have access to. Including all proxy variables for complexity in a multivariate regression informs us of the relative noise in each proxy as well as the correlation in noise across variables. This is superior to simply averaging or aggregating all the variables but also relies on our variables being valid proxies for complexity.

However, the index's correlation with compliance or administrative costs may be positive or negative. In high-income countries, longer laws are usually associated with more compliance costs (Slemrod, 2005), but especially in a low-income country context, sometimes more complexity may be desirable (Tran-Nam and Evans, 2014) to provide taxpayer guidance or limit the discretion bureaucrats may have in administration. This brief discussion highlights that there is not a consensus over the definition of tax complexity. Slemrod (2005) notes that "some argue that spelling out the tax ramifications of as many as possible situations reduces uncertainty and therefore complexity, with others are taking the opposite position." Our view is similar to Slemrod (1996) who argues that the best measure of tax complexity is the total resource cost of taxation: the government budget for tax administration plus the cost to taxpayers and any third parties of compliance. Given our paper is about firms, our proxies are designed to capture the resource costs of the firm. This conceptual discussion makes it clear that there is no one way to measure complexity, and in different contexts, complexity may have larger effects on administrative rather than compliance costs. Although our purpose is to capture costs of the firm, it may also capture the costs of the tax authority.¹⁵

The final panel of Figure 2 makes it clear that our index of complexity has a large change after the reform. Given the index is standardized, the change is almost a one standard deviation change by three years after the reform.

4.2 Tax Rates

Turning to tax rates, using the tax schedules on their own is potentially misleading because it may not capture the experience of the typical firm very well. Matching the productspecific tax rates to firms in the ASI dataset allows us to rectify this. In Figure 3, we show event studies for firm-specific statutory tax rates on inputs and outputs before and after the reform. To do this, we calculate the weighted average of statutory tax rates on all inputs, $\tau_{f(s),t}^{I}$, [respectively outputs, $\tau_{f(s),t}^{O}$] a firm uses, where the weights are proportional to the product's share of total inputs [outputs] in the firm. To shut down any endogenous changes in production patterns as a result of the VAT reform, we hold production constant over time at the pre-reform level.¹⁶ This implies that the firm-specific tax rate can only change due to changes in the statutory tax rate of one or more of the inputs. As a cautionary note, the pre-reform coefficients do not provide a very strong test of parallel trends in this specific instance. Because we only have sales tax information for two pre-reform years, we need to impute tax rate information for other years using those years. This is not a concern for the post-reform period where we have high-frequency information on tax rate changes. As Figure 3 shows, the average statutory input and output tax rates facing a typical firm in the ASI dataset fall substantially. Relative to the average input tax rate of 7.31% in the year before the reform, the tax rate drops by over 1 percentage point in the initial implementation year, with a further decline over the next few years. By year four of the reform, input tax rates for the typical firm have declined by about 2.5 percentage points relative to the pre-period, a reduction of 34%. A qualitatively similar pattern is visible for the output tax rate. The average firm-specific output tax rate declines by about 1.5 percentage points relative to the pre-reform mean of 7.16% immediately following VAT adoption. Four years after, the decline is almost 4 percentage points, implying a reduction of over 50%.

These results only represent changes in the statutory tax rate on inputs or outputs. Absent double-taxation of inputs, the statutory tax change could also be interpreted as an effective tax change. However, this would not be appropriate in the Indian context due to the extensive amount of tax cascading. Given that the Indian system featured extensive cascading (Chelliah and Committee, 1992), it is likely that the percent change in prices may be substantially higher than the statutory tax changes we document.

To address this issue, we exploit the linking of inputs and outputs in the ASI to construct effective tax changes. For firm f in state s, we construct an effective tax rate as:

$$\tau_{f(s),t}^E = \tau_{f(s),t}^O + s_{f(s)} \cdot \alpha_t \cdot \tau_{f(s),t}^I \tag{4}$$

where $s_{f(s)}$ is the (constant) pre-reform share of total production costs that are devoted to material (goods) inputs, and α_t is the factor that inputs are taxed at relative to outputs before and after the reform.¹⁷ Under the sales tax system, the factor α_t is determined by different (unobservable to us) tax schedules for inputs at different stages of production, while in the VAT system cascading is more likely due to imperfections in the credit system. As discussed in appendix B.2, although not perfect, we use Aggarwal (1995) and Poddar and Ahmad (2009) to pin down α_t under the sales and VAT system, respectively. The lack of data on α_t is a limitation that induces measurement error in our effective tax rates. As we do not know how cascading varies by firm, we assume α_t is constant across firms and states. As a robustness check, we will show how varying α_t under the sales tax and VAT changes the magnitude of the elasticities. Further, we conduct simulations at the firm, good, and state level to determine the robustness of our results to measurement error at these different levels. As we will show, the level of α_t is most critical for our estimates: adding idiosyncratic variation across goods, firms or states without changing the mean leaves the point estimates almost unchanged. Changes to the mean value of α_t in the pre- and post-reform, in contrast, affect our estimates quantitatively but leave the qualitative patterns unchanged.

The final panel of Figure 3 shows that the effective tax change was double the tax change on outputs. Interestingly, due to differences in tax rates across products, effective tax changes differ substantially within an industry and state. We conduct an ANOVA for the effective tax rate measure, using state fixed effects and then state by industry effects. State fixed effects explain 11.4% [3.1%] of variation before [after] the reform. With state by industry effects, those values become 43.7% and 50.3%, respectively. There is ample variation within industries of the same state, but the within variation decreases after the reform.

5 The Effect of the Tax Reform on Production

Overall, the prior results show that the VAT adoption lowered the complexity of the tax system on many dimensions. For the typical manufacturing firm, statutory and effective tax rates also declined sharply. Both the effective tax changes and declines in tax complexity, along with any other efficiency benefits of the VAT relative to ST system, could therefore lead to positive impacts on production, which we turn to next.

Figure 4a plots the event study coefficients for sales. The analysis exploits 486,664 firm-year observations from repeated cross-sections. Since a sizable number of firms have low or no sales, we run the specification in levels, though these results are robust to a log specification. To provide an indication of the magnitude of the effect, we also report the

pre-reform outcome mean in event year -1. To test the robustness of the results to different decisions about the composition of the control group, we report two event studies on the same graph. The circular lines plot the coefficients and corresponding confidence intervals for our preferred specification, which drops the small number of never-treated observations in our sample—firms from the Andaman and Nicobar Islands. This Union Territory never adopted a VAT, but also did not have a sales tax. The islands are extremely remote and do not have the status of a full state. These factors are likely to make the Andaman and Nicobar Islands systematically different than other Indian states. In the diamond lines on the graph, we include firms from this area along with the not-yet treated observations.

Figure 4a finds that both of these specifications yield almost identical results. It shows a dynamic increase in sales post treatment. By four years after the adoption of the VAT, sales increase by 79 million rupees. Given the coefficients on the event dummies are relative to the omitted year, this represents a cumulative effect. The dynamic path indicates that the VAT reform shifts up production, but this occurs gradually in the years following the reform. This is consistent with frictions and learning, which are critical to the firm adjusting their production process, as well as compounding effects of the reform. The smaller effect in event year 0 could additionally result from the way we have coded event time where some firms are treated for only a partial year in event year 0. Not until event year 1 are all firms treated for the *full* financial year. The increase in sales is very large relative to the pre-treatment mean of 155 million rupees, implying a percentage medium-run increase in average sales of about 50%. The pre-treatment coefficients, on the other hand, are close to zero, supporting the assumption that the estimation satisfies the parallel trend assumption.

However, when thinking about the dynamics of the reform, it is important to keep in mind that many regions adopt the VAT in close succession to each other. As a result, longerrun effects are estimated off of a smaller number of treatment and comparison units than shorter-run effects. That the number of yet-to-be treated units declines over time is especially important when thinking about the effects in +3, which shows a dynamic path different from other event periods. As a result, when subsequently reporting elasticity estimates, we will do so at different time horizons, emphasizing which results survive when focusing on shorter-run estimates that are identified from a larger number of treated/comparison units.

Appendix Table C.4 reports the coefficients corresponding to Figure 4a. Critically, all the γ_y coefficients are at least an order of magnitude larger than the pre-reform π_y coefficients. The average effect of the reform is an increase of sales of about 52 million rupees or a 34% increase relative to the pre-reform mean. Three years after the reform, output increased by 57%. Percent changes are similar to those using a log specification (Appendix Figure C.5).

Figure 4b presents the analogous results for the value of purchased inputs. Again, including the never treated observations does not change the results meaningfully, so we drop the Andaman and Nicobar Islands in all further results. As in Figure 4a, we see dynamic treatment effects that increase with length of exposure to the reform. By year four of implementation, the value of inputs has risen by 39 million rupees (60%) relative to the pre-reform mean. Column 2 of Table C.4 reports the corresponding coefficients and the ATT estimate.

5.1 Discussion of Magnitudes

The magnitudes of the identified effects are very large, even accounting for the fact that the tax changes were substantial and that the reform also bundled together numerous policy changes that are all likely to lead to firm growth. In this section, we explore several possible explanations that could lead to such large effects: outliers and the exclusion of informal firms, composition effects of firms, and tax evasion being more intense under the ST regime.

5.1.1 Outliers and Biases From Excluding Informal Firms

As we use a level specification, one concern is that our results might be driven by very large firms. An alternative specification is to use logs or, given that sales are close to zero for a subset of firms, the inverse hyperbolic sine transformation. As Appendix Figure C.5 shows, our results are robust to these specifications. Our results are also robust to winsorizing our outcome variables of interest. In results not presented here, we explore heterogeneity by prereform average sales quartiles.¹⁸ We find that the sales response to VAT adoption is large for all quartiles relative to the mean at event time -1, suggesting that the overall results are not driven by a few outliers or by the highest quartile.

With respect to firm size, using our threshold database, we study what happens to firms that are exempt from the VAT and ST regime, respectively. Here we separately focus on firms that are always below the sales tax threshold or the VAT threshold, which minimizes possible confounding effects from the firms being directly treated by the reform later in time. In the sample that are always under the sales tax exemption threshold, we find no or small negative effects, although these firms are extremely small in size and not well represented in the data. When focusing on firms below the higher VAT thresholds, we identify positive but insignificant effects. These firms could benefit from trading with firms above the threshold, but also from switching from paying the sales tax to being exempt from VAT.

Recall the ASI is representative of the state distribution of firms registered under the Factories Act. The exclusion of (smaller) formal tax-registered firms and informal not-tax-registered firms may create biases. To further explore this issue we use the GDP data series for manufacturing (GSVA/NSVA) from the Niti Aayog website. This series contains information on firms that are registered under the Factories Act and those that are not. The registered component of the series is derived from the ASI while data for unregistered informal units is derived from quinquennial follow-up surveys of the Economic Census—these units represent 35% of sectoral output in 2002. Appendix Figure C.6 shows that the pattern of response is qualitatively the same as identified in the ASI, but total output in the sector increased by a smaller amount: 20%. The smaller response could be due to smaller effects among unregistered informal firms, especially if these firms trade with each other to avoid the regulatory burden of trading with a formal firm. Alternatively, as discussed below, part of the ASI response could be a compliance effect not found amongst informal firms.

5.1.2 Composition Effects and Integration

With respect to composition effects, we cannot directly measure entry/exit from the dataset. However, although it is often argued that incentives to vertically integrate are stronger the more tax cascading exists, it may also be possible that firms are vertically integrating in order to avoid the scrutiny of inter-firm transactions under the VAT.¹⁹ To attempt to address this, we run an event study on firm age. We see a mild increase in firm age—though statistically insignificant—of approximately 6 months relative to a pre-reform mean of 15 years. Our increases in output could therefore capture mild effects from firm composition.

5.1.3 Survey Versus Tax Return Data

Finally, although the ASI is a survey, firms could fill out the survey to match information reported on tax returns. Then, because the sales tax had stronger incentives to evade, some of our effects could be due to misreporting rather than a real response. To attempt to address this issue we explore the effects by proxies for evasion. We use data from input-output tables to determine the share of output by sector for private final consumption. We also explore heterogeneous effects using the backward and forward linkages constructed in Hoseini and Briand (2020). The basic idea is that firms in industries with more sales to final consumers should have more scope for evasion under the VAT—but this proxy is imperfect because there is likely substantial within industry heterogeneity and because the firm's position in production networks may be highly correlated with factors other than evasion. The results in the Appendix Figure C.8 are generally statistically indistinguishable from one another, but firms with fewer business-to-business sales as measured by forward linkages appear to have larger effects, though they are also slightly larger firms. Given the imperfect nature of these proxies and the inability to match our data to tax records, we conclude that it is possible some of our responses could be due to a decline in underreporting resulting from the enforcement mechanisms of VAT relative to the sales tax.

Of course, issues of measurement error in survey data are well-known in the individual

setting (Bollinger et al., 2019) and recent work has compared individual survey data to administrative data (Bollinger et al., 2018). Overall, this literature finds meaningful differences between survey and administrative data on average, and our results should be interpreted in the context of those differences. Further, because our results rely on survey data, it is impossible to know to what extent there was a similar change in the amounts actually declared in tax forms—but Appendix Figure C.7, which shows the dynamics of tax revenues, indicates larger percent increases than our main event studies. There are many reasons to believe that firms may be more compliant in a VAT system—thanks to its design features—than in the prior system, but there is substantial uncertainty about the magnitude in our setting and little evidence on the extent to which these compliance effects manifest in survey data.

5.2 Limitations of the Empirical Design

In this section, we discuss some potential limitations of our analysis. The empirical approach relies on variation across states over time, which requires the SUTVA assumption that each state's potential outcomes are unaffected by the treatment status of other states. As with most subnational public finance papers, this may not be satisfied for several reasons: the reform may affect interstate trade, firms may relocate production from late-adopting states to early adopting states, and tax competition may occur between states. We acknowledge this is likely true, but attempt to highlight the expected magnitudes of these failings below.

5.2.1 Interstate Trade

According to Van Leemput (2021), Indian domestic trade as a percent of GDP in the manufacturing sector is 9%, much lower than the 30% in the U.S. Van Leemput (2021) estimates trade barriers in the U.S. are about one-fifth of those in India, concluding "This is suggestive of relatively large trade frictions across Indian states, which corroborates the relatively small domestic trade flows presented in the [data]." Critically, these trade frictions were likely similar both before and after the reforms because most states levied and continued to levy entry taxes on interstate trade (Van Leemput and Wiencek, 2017). Given this, we conclude that the reform may affect comparison states via trade, but this effect is small.

5.2.2 Firm Mobility / Composition

The large effects we estimate suggest that firms may have an incentive to relocate production from late-adopting to early-adopting states. Unfortunately, the data do not allow us to study firm migration directly, but any firm relocations to early adopting states should manifest in pre-trends of firm age. Appendix Figure C.11 provides circumstantial evidence that relocation is not common, perhaps due to regulatory and cultural barriers across states. This result on firm age is also useful to show that the extensive margin effect—changes in the composition of firms due to firm entry/exit—does not appear to be a major concern.

5.2.3 Tax Competition

The pre-reform period was characterized by "rate wars." Thus, it is likely that tax mimicking across states existed and that later-adopting states may have been influenced by their earlyadopting competitors. Studying the extent and existence of tax competition is challenging, and more so in this setting, where governments have a full vector of commodity tax rates to compete on. To provide some descriptive evidence, we study tax rate convergence across states, acknowledging that some models of tax competition need not predict convergence. We plot the standard deviation of tax rates across states by year in an attempt to show the effect of states adopting VAT on other jurisdiction tax rates (Appendix Figure C.14). In the year when most jurisdictions adopt (2005), this metric falls, but it falls much more so in 2006 when fewer states adopt, perhaps suggesting that earlier adopters may have responded. As noted in Agrawal et al. (2022) the best resolution to this type of competition is to understand the mechanisms underlying strategic tax setting behavior. Because interstate trade openness and firm mobility appear to be low, and because tax-driven cross-border shopping is likely costly due to poor infrastructure, competition likely results from mimicking or yardstick competition rather than mobility of the tax base. It is the latter mobility-induced competition that would be most problematic for SUTVA.

6 Heterogeneity of Effects

Given the substantial heterogeneity of tax changes with states and industries, we might expect to see heterogeneous effects. Because the reform reduced tax cascading, we might imagine the effects to be the largest amongst firms that have the highest effective tax rates in the pre-reform period. Figure 5a extends the analysis in Figure 3c by dividing firms into 4 quartiles based on the (firm-specific) pre-reform effective tax rates. As Figure 5a shows, firms in quartile 1—with the lowest pre-reform effective tax rates—saw the smallest absolute effective tax rate decreases as a result of the reform. Then, firm-specific effective tax rates decrease monotonically for the remaining quartiles. Intuitively, firms facing very high tax rates with substantial cascading had the most room for their tax rates to decline. Figure 5b and 5c reveal the absolute response of sales/inputs increases monotonically with the magnitude of the effective tax change. But firms with higher pre-reform tax rates are also larger firms at baseline, as the pre-reform outcome means show. As a result, the percent changes in output/inputs are relatively constant over the distribution of tax rates. To reconcile this, Figure 5a shows that the *percent* change in taxes is similar over the distribution of tax rates. Taken together, the implied elasticities of output/inputs with respect to the effective tax rate appears to be relatively constant over the tax rate distribution.

Next, Figure 6 shows heterogeneous effects with respect to how our complexity index interacts with tax rate changes. We continue to show effects by high/low pre-reform sales tax rates, as high sales tax firms experienced the biggest effective tax cuts. However, when cutting by high/low complexity we focus on post-reform values of our index in order to capture the extent of imperfections of the VAT after the reform. Higher post-reform values of the complexity index correspond to firms facing more schedules, more items in the schedules, more tax changes, and more pages in the tax code, etc. These proxies for complexity are

likely correlated with deviations from the standard broad-based flat-rated VAT.

Figure 6 shows that highly complex VAT systems saw only small changes in their complexity relative to the prior system. Tax rates generally fell for all four groups, though firms that realized the biggest declines in complexity actually saw slightly higher effective tax rates. When studying production responses, confidence intervals are generally large, but firms with low sales tax rates—and thus changes in tax rates that were not large—saw similar output changes regardless of whether the VAT regime had more or fewer imperfections. But when the sales tax rate was high—and effective tax rates fell substantially—the output response was larger and occurred more quickly if the VAT system saw large declines in complexity relative to the ST. Comparing the diamond and triangular lines yields similar long run changes, but the effects are almost twice as large in percent terms for low VAT complexity firms. The triangular lines indicate that even though these firms saw large tax decreases, the tax system only became mildly less complex along the dimensions our index captures. Thus, these firms needed to comply with an entirely new tax system that was almost as complex as the prior ST system, perhaps slowing the time path of growth. We conclude that the extent of imperfections in the VAT matters, but more so as a drag on growth when tax changes are large. Interestingly, the group that saw large declines in complexity—but increases in effective tax rates—realized small output gains, perhaps indicating that making the VAT less imperfect can be helpful even if tax rates don't change much.

7 Elasticities and Within Firm Characteristics

The prior analysis estimates the overall effect of the tax reform on sales and inputs. In this section, we convert our results into (tax) elasticities using a two-step procedure that exploits variation in tax changes across firms. Recall that due to cascading, we exploit linkages between inputs and outputs in the ASI to calculate the change in the effective tax rate. For purposes of this section, we present both short- and longer-run elasticities over the full event window. To obtain an elasticity, we first estimate the effect of the reform on effective tax rates using the event study approach in Figure 3c. We then combine this with our estimated event studies for various outcome variables, such as sales or inputs, obtained in Figure 4. The event studies in Figure 4 give us the change in an outcome variable, while those in 3c yield a change in effective taxes. We convert each of these to percent changes at the mean using the average outcome and tax rate period -1. Then, the elasticity is obtained as the ratio of the percent change in the outcome variable to the percent change in taxes at each time horizon. Confidence intervals are obtained by using the parametric bootstrap. Although we estimate tax elasticities, quantity responses will also include effects resulting from non-tax changes in the reform such as simplification or changes in tax design.

Table 1 presents the elasticities for different time horizons. The "longer-run" elasticities are larger in absolute magnitudes than contemporaneous/short-run elasticities, but longer-run estimates are identified from a smaller sample of yet-to-be treated units. Columns (1) and (2) report the elasticies for the two main outcome variables of interest, sales and inputs. In particular, a 1% decrease in the tax rate raises both sales and material inputs by approximately 0.24% and 0.88% depending on the time horizon. In interpreting our effects, recall that our elasticities are with respect to the *effective* tax rate, accounting for tax cascading. As our effective tax rates are calculated using only one linkage in the supply chain, if inputs to the firms' inputs are taxed, we may still underestimate the effective tax changes, which would lead us to overestimate the magnitude of the elasticities. We prefer to use the ASI data rather than input-output tables because the ASI linkages can capture within industry heterogeneity but at the expense of further linkages in the supply chain.

As a comparison to elasticities in the literature, we relate our elasticities to those with respect to input tariff changes. Amiti and Konings (2007) find that a 10 percentage point decline in input tariffs raised productivity by 12%. Topalova and Khandelwal (2011) find that a 10 percentage point decline in input tariffs improved productivity by 4.8%, which explains a large proportion of the productivity growth over the period. The VAT reform was a much more extensive reform, likely the largest and most salient recent reform to affect firms in India. Given the large effects, the tax reform likely had general equilibrium implications. Having large sets of firms simultaneously experience growth in factor and product demands will affect many markets, and these effects could in turn feed back to the sample firms studied through shifts in equilibrium factor prices and so on. Our elasticities are inclusive of these general equilibrium effects, which is the relevant concept to evaluate the totality of the reform. Our estimated elasticities, while large, are therefore not unreasonable.

One potential concern with the elasticity estimates is that generating the effective tax rates requires specifying a pre-reform and post-reform value of α_t in (4). In Appendix Table C.6 and Figure C.12, we therefore explore the robustness of the elasticities for a given time horizon to varying α_t . We first report how the elasticity varies when we hold the sales tax α_t constant at unity and vary the $\alpha_t \in [0, 1]$ under the VAT regime. As the VAT α_t declines, the elasticity monotonically tends toward zero. Intuitively, the more cascading is eliminated under the VAT regime, the larger the percent change in tax rates. Similarly, the table shows that if the VAT α_t is held constant at 0.4, but the sales tax value varies, the elasticity tends closer to zero as the sales tax cascading becomes more severe. The table can therefore be used to bound the estimated elasticities for plausible assumptions on the extent of cascading.

While the prior exercise is useful to showing how the results change as the level changes for all firms, it masks any implications of variation in α_t by firm, good, or state. To address this issue, we conduct six simulations where we allow the value of α to randomly vary by firm, primary-production good, and state separately for the pre- and post-reform. To do this, for each group (firm, good, state) we randomly draw a factor from the uniform distribution on the interval [0, 2]. Then each observation is assigned a value of α for the pre- and post-period that equals our baseline value times the random factor. Given the mean random factor is 1, this adds variation but leaves the mean approximately unchanged. We repeat this process one hundred times for each group, re-estimate the elasticity and plot the distribution of the coefficients in Appendix Figure C.13. The figure shows that our results are robust to adding idiosyncratic variation at any level. Note that if variation in the firm-good-state level instead were to systematically change the mean of α_t rather than (just) the variance, this analysis would collapse to the previous bounding exercise from Appendix Figure C.12. Potential firm-goods-state variation in α_t does not introduce additional concerns for our estimates.

To further test the plausibility of the estimates and get at the production mechanisms, we report several other elasticities.²⁰ In particular, in columns (3)-(8) of Table 1 we explore if the increases in production coincide with firms changing capital/labor/land/materials used in the production process. The results indicate that for 1% changes in taxes, worker earnings in column (3) increase by between 0.17% to 1.05% depending on the time horizon. The value of capital used in production, column (4), increases by at most 0.4%. Combining capital, earnings, material inputs, and other costs of production, we see an increase in total costs in column (5). These results are consistent with firms scaling up production and spending more on capital, labor, and materials. One theory that rationalizes such a result, owing to the standard Harberger-style analysis, could be that the incidence of an overall reduction in tax burden is shared with the factors of production. Another theory could be changes in inputs arise because firms experience tax rate changes that impact marginal incentives for various factors differentially across firms depending on inputs used in production. Although the prices of labor and capital are not directly affected, their relative prices with material inputs are, and as discussed above general equilibrium effects could affect factor prices.

Column (6) indicates the increase in the cost of labor is not a result of increasing the number of employees, but rather a result of higher wages/compensation (or hiring more skilled labor), providing some indication on the incidence within the firm. Columns (7) and (8) show the overall increase in capital is partially coming from increases in land and machinery. India has inefficient land use; our results suggest the sales tax amplified this.

In column (9) we study the effect on digital/computerized accounting systems. The reform made the tax system less complex on some dimensions, more detailed on the list of goods, and slightly more standardized. But firms had to comply with a "new" tax. The reform appears to incentivize firms to use computerized accounting systems. This effect may

have been induced by the initial fixed costs of switching from a familiar to an unfamiliar tax.

To conclude, in column (10), we digitize total consumption tax revenue in each state and estimate the elasticity of tax revenue with respect to the tax reform. The implied elasticity is between -0.98 and -1.48 depending on the time horizon, indicating that the transition to the VAT raised tax revenues. Does this mean Indian states were on the wrong side of the Laffer curve? Because the base, evasion, and other aspects of the tax system changed, a negative elasticity here does not necessarily imply that revenue would have increased had states lowered sales tax rates, but kept this tax regime. The increase in tax revenue is likely a result of production gains, but also better compliance and perhaps a broader tax base. However, given that revenue rises as the effective tax rates decline implies that a broader VAT with lower rates than a narrower sales tax is more efficient at raising revenue, consistent with theoretical evidence favoring the VAT. Despite ample theoretical evidence, this question has empirically "received little attention" (Crawford et al., 2010). These revenue effects capture the totality of tax changes from a state-level fiscal perspective.

8 Conclusion

We provide novel empirical evidence on how the switch from a sales tax to a value added tax affects production responses in an environment where the VAT is designed imperfectly. In addition, we contribute to the literature on fiscal federalism in developing countries, where there has been a trend toward more decentralization (Gadenne and Singhal, 2014).

"C-efficiency" is a widely used tool for evaluating how close the VAT is to an ideal VAT. In particular, it compares the VAT revenue actually raised to the one which would be raised under a perfect VAT with a uniform rate, a comprehensive consumption tax base, and no exemptions, under the assumption that such a perfect world does not induce a change in quantity demanded. As shown in Keen (2013), the C-efficiency rate in lower-income countries is well below 50% and substantially lower than in high-income countries. In Africa, for example, C-efficiency is approximately 35%, suggesting that the VAT as implemented falls short of our textbook VAT. Despite these shortcomings, our paper indicates that relative to a cascading sales tax, the VAT remains a highly-recommended tax for developing countries. Even with an unconventional design of the VAT, as in India, the switch to a VAT can have many efficiency advantages especially if cascading is extensive under the sales tax system. This gain in efficiency plausibly arises because of the elimination of inefficient double taxation, reductions in tax rates from the broader base, and reductions in complexity.

While we have studied a broad range of economic outcomes in our analysis using the widely used ASI dataset, it has some limitations. The dataset only covers manufacturing firms above an employee threshold of 10 workers (20 workers without the use of electricity). Many of the excluded small firms are unlikely to be *directly* affected by the reform as they will not have high enough turnover to require registration under the sales tax or VAT. But the exclusion of small and predominantly informal firms limits our ability to study the effects on small client firms that occur via production networks (Boadway and Sato, 2009; Gadenne et al., 2022; Piggott and Whalley, 2011). Finally, the ASI is a survey dataset, which prevents us from studying tax evasion and compliance effects, which could possibly confound our estimates if firms use their tax returns to reply to the ASI. Future research may combine our tax data series with surveys of informal firms or other data on smaller firms to study these questions. In addition, we have not studied consumer outcomes or the equity implications (Bachas et al., 2024) of different consumption tax designs. We hope this paper provides a starting point, both descriptively and causally, for studying the role of imperfect tax design in tax system transitions in low-income countries.

Notes

¹Keen (2013) documents that deviations from a textbook VAT are standard in many countries, including in Europe, but more substantial in emerging and developing markets.

²Adoption of a VAT comes with the standard concern that firms may evade VAT by shifting to a lower-rate turnover tax (Best et al. 2015). See also Kleven and Waseem (2013),
Slemrod et al. (2022), Waseem (2018), and Waseem (2023), for issues of tax evasion.

³This changed with the introduction of a national VAT in 2017.

⁴Appendix Figure C.1 provides an example (one of twenty-five pages) for West Bengal. ⁵For an example of a notification, see Appendix Figure C.2.

⁶Most states include information about notifications that altered tax rates in footnotes. Based on the last tax rate change and a comparison with all individual tax notifications from the archive, it is typically possible to date the posted documents.

⁷Information is unavailable for Meghalaya, Tripura, Dadra and Nagar Haveli, and Daman and Diu. We impute tax rates for the missing states by choosing the most culturally similar state that also has a similar VAT schedule in terms of length and tax rates.

⁸See also Velayudhan (2018), who studies notches in the CenVAT in India.

⁹An exception are the small states of Manipur, Meghalaya, Nagaland, Tripura, and Andaman & Nicobar Islands, where all firms are surveyed every year regardless of size.

¹⁰We observe a wider event time window than we will report estimates for. Due to the staggered rollout the earlier pre-period observations are unbalanced for many states and the later post-periods occur when all ever-adopting states are treated. While the unbalanced event years are included in the analysis dataset, as they to help estimate state and year fixed effects in the first stage of the Gardner procedure, we only present coefficients for balanced pre-event years (Malkova, 2018).

¹¹ATT estimates are also qualitatively similar when we use the alternative estimators proposed by Callaway and Sant'Anna (2021) and de Chaisemartin and D'Haultfoeuille (2020).

¹²Only the Andaman & Nicobar Islands and Lakshadweep, both small remote archipelagos that are likely very different than continental India, do not adopt. The former has only a few hundred firms our dataset, while the latter has none.

¹³Newspaper reports document the uncertainty about when states would adopt, leaving little room for firms to change behavior in anticipation.

 14 Election timing and seat shares of the two national parties (Congress/BJP) are from the

Election Commission of India. We construct an indicator variable of whether a state government is aligned with the Center by checking whether the parties in the state government are part of the central government (typically a coalition government) in that year.

¹⁵For example, the share of goods in the default tax category also may change the cost of administration while also influencing the costs to firms.

¹⁶Top input and output shares, which account for the lion share of input purchases and production, are relatively constant in the pre-reform period.

¹⁷The value of α_t need not be bounded above by one.

¹⁸We do not formally implement this as a quantile regression due to the concerns mentioned in Athey and Imbens (2006) for DID settings.

¹⁹Gadenne et al. (2022) find evidence for additional supply chain inefficiencies that are created by differential incentives to register for the tax.

²⁰Appendix Table C.5 reports estimated elasticities for additional outcome variables related to the production process that are not emphasized in the text. Event studies for all variables for which we estimate elasticities are shown in Figures C.9 and C.10.

References

- Adhikari, Bibek (2020) "Does a Value-Added Tax Increase Economic Efficiency?" Economic Inquiry, 58 (1), 496–517.
- Aggarwal, Pawan K. (1995) "Incidence of Major Indirect Taxes in India," National Institute of Public Finance and Policy.
- Agrawal, David R., William H. Hoyt, and John D. Wilson (2022) "Local Policy Choice: Theory and Empirics," *Journal of Economic Literature*, 60 (4), 1378–1455.
- Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell (2016) "How Do Electricity Shortages Affect Industry? Evidence from India," *American Economic Review*, 106 (3).
- Amiti, Mary and Jozef Konings (2007) "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," American Economic Review, 97 (5), 1611–1638.

- Athey, Susan and Guido W. Imbens (2006) "Identification and Inference in Nonlinear Difference-in-differences Models," *Econometrica*, 74 (2), 431–497.
- Bachas, Pierre, Lucie Gadenne, and Anders Jensen (2024) "Informality, Consumption Taxes and Redistribution," *Review of Economic Studies*, 91 (5), 2604–2634.
- Bagchi, A. and Study Team (1994) Reform of Domestic Trade Taxes in India: Issues and Options: National Institute of Public Finance and Policy.
- Bailey, Martha J. and Andrew Goodman-Bacon (2015) "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans," *American Economic Review*, 105 (3), 1067–1104.
- Best, Michael, Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem (2015) "Production vs Revenue Efficiency with limited Tax Capacity: Theory and Evidence from Pakistan," *Journal of Political Economy*, 123 (6), 1311–1355.
- Boadway, Robin and Motohiro Sato (2009) "Optimal Tax Design and Enforcement with an Informal Sector," *American Economic Journal: Economic Policy*, 1 (1), 1–27.
- Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak (2018)"The Good, the Bad and the Ugly: Measurement Error, Non-response and Administrative Mismatch in the CPS," Technical report, Working Paper, University of Kentucky.
- (2019) "Trouble in the Tails? What We Know about Earnings Nonresponse 30 years after Lillard, Smith, and Welch," *Journal of Political Economy*, 127 (5), 2143–2185.
- Callaway, Brantly and Pedro H.C. Sant'Anna (2021) "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225 (2), 200–230.
- de Chaisemartin, Clement and Xavier D'Haultfoeuille (2020) "Two-Way Fiexed Effects Estimators with Heterogeneous Treatment Effects," American Economic Review, 110 (9), 2964–2996.
- Chelliah, Raja Jesudoss and Tax Reforms Committee (1992) "Tax Reforms Committee: Final Report," Directorate of Publications, Ministry of Development.
- Commercial Law Publishers (2000) Commercial's All India Sales Tax Tariff: States' Sales

Tax Rates, Delhi: Commercial Law Publishers.

- (2002) Commercial's All India Sales Tax Tariff: States' Sales Tax Rates, Delhi: Commercial Law Publishers.
- Crawford, Ian, Michael Keen, and Stephen Smith (2010) "Value Added Tax and Excises," in *The Mirrlees Review*, 4, 276–421: IFS.
- De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik (2016) "Prices, markups, and trade reform," *Econometrica*, 84 (2), 445–510.
- Gadenne, Lucie, Tushar Nandi, and Roland Rathelot (2022) "Taxation in Supplier Networks: Evidence from India," October, Working paper.
- Gadenne, Lucie and Monica Singhal (2014) "Decentralization in Developing Economies," Annual Review of Economics, 6, 581–604.
- Gardner, John (2021) "Two-Stage Differences in Differences," Working Paper.
- Goldberg, Pinelopi K., Amit K. Khandelwal, Nina Pavcnik, and Petia Topalova (2010a) "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *Quarterly Journal of Economics*, 125 (4), 1727–1767.
- (2010b) "Multiproduct Firms and Product Turnover in the Developing World: Evidence from India," *Review of Economics and Statistics*, 92, 1042–1049.
- Hoseini, Mohammad and Oceane Briand (2020) "Production Efficiency and Self-Enforcement in Value-Added Tax: Evidence from State-Level Reform in India," *Journal of Development Economics*, 144, 102462.
- Hsieh, Chang-Tai and Peter J. Klenow (2009) "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124 (4), 1403–1448.
- (2014) "The Life Cycle of Plants in India and Mexico," *Quarterly Journal of Economics*, 129 (3), 1035–1084.
- Keen, Michael (2012) "Targetting, Cascading and Indirect Tax Design," February, Raja Chelliah Memorial Lecture Series.
 - (2013) "The Anatomy of the VAT," National Tax Journal, 66 (2), 423–446.

(2014) "Lessons for the GST from International Experience: Possibilities and Pitfalls," *India Tax Insights*, 3, 12–14.

- Keen, Michael and Ben Lockwood (2010) "The Value Added Tax: Its Causes and Consequences," *Journal of Development Economics*, 92, 138–151.
- Kleven, Henrik J. and Mazhar Waseem (2013) "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 128 (2), 669–723.
- Lubotsky, Darren and Martin Wittenberg (2006) "Interpretation of Regressions with Multiple Proxies," *The Review of Economics and Statistics*, 88 (3), 549–569.
- Malkova, Olga (2018) "Can Maternity Benefits Have Long-term Effects on Childbearing?Evidence from Soviet Russia," *Review of Economics and Statistics*, 100 (4), 691–703.
- Piggott, John and John Whalley (2011) "VAT Base Broadening. Self Supply and the Informal Sector," American Economic Review, 91 (4), 1084–1094.
- Poddar, Satya and Ehtisham Ahmad (2009) "GST Reforms and Intergovernmental Considerations in India," Government of India Working Paper 1/2009, Ministry of Finance, Department of Economic Affairs.
- Purohit, Mahesh C. (2001a) Sales Tax and Value Added Tax in India, Delhi: Gayatri Publications.
- (2001b) Value Added Tax Experiences of India and Other Countries, Delhi: Gayatri Publications.
- Sangal, Balram (2005) Commercial's All India VAT Manual, Delhi: Commercial Law Publishers.
- Sangal, Balram and Jagdish Rai Goel (2007) Commercial's All India VAT Manual, Second Edition, Delhi: Commercial Law Publishers.
- Sincavage, Jessica R., Carl Haub, and O. P. Sharma (2010) "Labor Costs in India's Organized Manufacturing Sector," *Monthly Labor Review*, 133 (5), 3–22.
- Slemrod, Joel (1996) The Economics Effects of Fundamental Tax Reform, Chap. Which Is

the Simplest Tax Sytem of Them All?: Brookings Institution Press.

- (2005) "The Etiology of Tax Complexity: Evidence from U.S. State Income Tax Systems," *Public Finance Review*, 33 (3), 279–299.
- Slemrod, Joel B and Tejaswi Velayudhan (2022) "The VAT at 100: A Retrospective Survey and Agenda for Future Research," *Public Finance Review*, 50 (1), 4–32.
- Slemrod, Joel, Obeid Ur Rehman, and Mazhar Waseem (2022) "The Role of Withholding in the Self-Enforcement of a Value-Added Tax: Evidence from Pakistan," *Review of Economics and Statistics*, 104 (2), 336–354.
- Smart, Michael and Richard M. Bird (2009) "The Impact on Investment of Replacing a Retail Sales Tax with a Value-Added Tax: Evidence from Canadian Experience," National Tax Journal, 62 (4), 591–609.
- Sury, M. M. (2015) Tax System in India: Evolution and Present Structure, Delhi: New Century Publications.
- Topalova, Petia and Amit Khandelwal (2011) "Trade Liberalization and Firm Productivity: The Case of India," *Review of Economics and Statistics*, 93 (3), 995–1009.
- Tran-Nam, Binh and Chris Evans (2014) "Towards the development of a tax system complexity index," *Fiscal Studies*, 35 (3), 341–370.
- Van Leemput, Eva (2021) "A Passage to India: Quantifying Internal and External Barriers to Trade," *Journal of International Economics*, 131, 103473.
- Van Leemput, Eva and Ellen A. Wiencek (2017) "The Effect of the GST on Indian Growth," unpublished.
- Velayudhan, Tejaswi (2018) "Misallocation or Misreporting? Evidence from a Value Added Tax Notch in India," Working Paper.
- Waseem, Mazhar (2018) "Taxes, Informality and Income Shifting: Evidence from a Recent Pakistani Tax Reform," *Journal of Public Economics*, 157, 41–77.
- (2023) "Overclaimed Refunds, Undeclared Sales, and Invoice Mills: Nature and Extent of Noncompliance in a Value-added Tax," *Journal of Public Economics*, 218, 104783.



Figure 1: Tax Rates and Number of Tax Categories by State and Time

(c) Average Tax Rate



Legend: Panel (a) of this figure shows the number of unique tax rate schedules by state and time. Panel (b) shows the number of products itemized in all of the tax schedules. Panel (c) shows the average tax rate by state where each product receives equal weight. Panel (d) shows the default tax rate by state. The average tax rate is calculated over the items in the tax schedule meaning that very detailed schedules get additional weight relative to less specific schedules within the state. The default tax rate is the rate that applies to all goods not explicitly listed in the schedule. In all figures, we depict the VAT statistics in the first year after the state reform and then four years thereafter; given the staggered adoption across states, this can be different years.



Figure 2: Event Study of Complexity Measures

(c) Complexity Index

Legend: This figure shows the effect of VAT adoption on the number of tax rates (Panel a), the number of products itemized in the schedule (Panel b), and our complexity index (Panel c). The complexity index is constructed using the procedure outlined in Section B.1. To construct this figure, we omit two pre-reform periods, as we only have sales tax data in alternate years. Each of the first two panels shows the mean in period -1, immediately prior to the reform. Because our pre-reform data is only for 2000 and 2002, there is no variation for some event years. For the 3 pre-period years that we can include, some of the variation comes from the fact that these two calendar years occur at different times for different staggered timing groups. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure 3: Event Study of Firm-Specific Statutory and Effective Tax Rate Changes

(c) Effective Tax Rate

Legend: This figure shows the effect of VAT adoption on firm-specific statutory tax rate on inputs (Panel a), outputs (Panel b), and the effective tax rate accounting for cascading (Panel c). To construct this figure, we hold production constant at its value in -1 and trace the tax rates forward and backward for the fixed set of products for each firm, weighting each product's tax rate by its total share of inputs or production. Effective tax rates are calculated as described in Section B.2. The mean is the average value in period -1, immediately prior to the reform. Because our pre-reform data is only for 2000 and 2002. For the pre-period years that we can include, some of the variation comes from the fact that these two calendar years occur at different times for different staggered timing groups. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure 4: Effect of VAT Reform on Firm Sales and Input Purchases

Legend: This figure shows the effect of the adoption of VAT using a set of all open firms. Panel (a) shows the effect of the reform on sales, while Panel (b) shows the effect on purchased value of inputs. The circular line includes never-treated observations (Andaman and Nicobar Islands), while the diamond lines do not include them in the comparison group. Standard errors are clustered at the state level and the figure shows 95% confidence intervals; period -1 is the omitted category so no confidence bands are included. Period -1 is the year prior to the VAT reform, where the entire year was subject to the sales tax regime.



Figure 5: Heterogeneity by Effective Sales Tax Rates

Legend: This figure shows result after segmenting the sample into quartiles of the pre-reform effective sales tax rate faced by the firm. Panel (a) shows the change in effective tax rates by quartile, Panel (b) shows the change in sales output, and Panel (c) shows the change in purchased inputs. The figure estimates the effect using open firms in the ASI database. Period -1 is the omitted category. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure 6: Heterogeneity by VAT Complexity and Tax Rate Interactions

Legend: This figure shows effect of VAT adoption using open firms in the ASI database. We split the sample based on levels of pre-reform effective sales tax rates (comparable to prior figures) but also on levels of post-reform complexity as measured by our complexity index. The focus on post-reform complexity is designed to tease out imperfections in the VAT that remain after the reform. Tax rates and complexity are split using the median value. Panel (a) shows how the complexity index changes over time. Panel (b) shows the changes in the effective tax rate. Finally, Panel (c) and (d) show the effect of the reform on sales and the value of inputs. Period -1 is the omitted category. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.

	Panel A				
	(1)	(2)	(3)	(4)	(5)
Outcome	Sales	Inputs	Earnings	Capital	Cost
Elasticity 0	-0.237	-0.245	-0.174	-0.424	-0.335
CI 0	[768;.249]	[764 ; .227]	[53;.149]	[842 ;085]	[913;.18]
Elasticity 1	-0.597	-0.539	-0.590	-0.409	-0.744
CI 1	[-1.134 ;116]	[923 ;204]	[-1.031 ;204]	[725 ;131]	[-1.365;192]
Elasticity 2	-0.880	-0.882	-0.939	-0.229	-0.999
CI 2	[-1.229;587]	[-1.21;612]	[-1.256;683]	[481;.009]	[-1.367;697]
Elasticity 3	-0.761	-0.869	-1.046	-0.365	-0.866
CI 3	[-1.013 ;555]	[-1.138;653]	[-1.403 ;753]	[686 ;066]	[-1.142;643]
		Panel B			
			Panel B		
	(6)	(7)	Panel B (8)	(9)	(10)
Outcome	(6) Employees	(7) Land	Panel B (8) Machinery	(9) Digital	(10) Tax Revenue
Outcome Elasticity 0	(6) Employees 0.0401	(7) Land -0.120	Panel B (8) Machinery -0.539	(9) Digital -0.0974	(10) Tax Revenue -0.975
Outcome Elasticity 0 CI 0	(6) Employees 0.0401 [166 ; .254]	(7) Land -0.120 [726 ; .464]	Panel B (8) Machinery -0.539 [-1.013 ;168]	(9) Digital -0.0974 [289 ; .076]	(10) Tax Revenue -0.975 [-1.822 ;35]
Outcome Elasticity 0 CI 0 Elasticity 1	(6) Employees 0.0401 [166 ; .254] -0.0367	(7) Land -0.120 [726 ; .464] -0.590	Panel B (8) Machinery -0.539 [-1.013 ;168] -0.421	(9) Digital -0.0974 [289 ; .076] -0.205	(10) Tax Revenue -0.975 [-1.822 ;35] -1.262
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1	(6) Employees 0.0401 [166 ; .254] -0.0367 [229 ; .152]	(7) Land -0.120 [726 ; .464] -0.590 [-1.107 ;129]	Panel B (8) Machinery -0.539 [-1.013 ;168] -0.421 [781 ;102]	(9) Digital -0.0974 [289 ; .076] -0.205 [42 ;01]	(10) Tax Revenue -0.975 [-1.822 ;35] -1.262 [-1.942 ;690]
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1 Elasticity 2	 (6) Employees 0.0401 [166 ; .254] -0.0367 [229 ; .152] 0.0163 	(7) Land -0.120 [726 ; .464] -0.590 [-1.107 ;129] -1.057	Panel B (8) Machinery -0.539 [-1.013 ;168] -0.421 [781 ;102] -0.240	(9) Digital -0.0974 [289 ; .076] -0.205 [42 ;01] -0.435	(10) Tax Revenue -0.975 [-1.822 ;35] -1.262 [-1.942 ;690] -1.399
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1 Elasticity 2 CI 2	 (6) Employees 0.0401 [166 ; .254] -0.0367 [229 ; .152] 0.0163 [198 ; .231] 	(7) Land -0.120 [726 ; .464] -0.590 [-1.107 ;129] -1.057 [-1.438 ;744]	Panel B (8) Machinery -0.539 [-1.013 ;168] -0.421 [781 ;102] -0.240 [451 ;044]	(9) Digital -0.0974 [289 ; .076] -0.205 [42 ;01] -0.435 [581 ;316]	(10) Tax Revenue -0.975 [-1.822 ;35] -1.262 [-1.942 ;690] -1.399 [-2.099 ;827]
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1 Elasticity 2 CI 2 Elasticity 3	 (6) Employees 0.0401 [166 ; .254] -0.0367 [229 ; .152] 0.0163 [198 ; .231] 0.00125 	(7) Land -0.120 [726 ; .464] -0.590 [-1.107 ;129] -1.057 [-1.438 ;744] -1.412	Panel B (8) Machinery -0.539 [-1.013 ;168] -0.421 [781 ;102] -0.240 [451 ;044] -0.307	(9) Digital -0.0974 [289 ; .076] -0.205 [42 ;01] -0.435 [581 ;316] -0.446	(10) Tax Revenue -0.975 [-1.822 ;35] -1.262 [-1.942 ;690] -1.399 [-2.099 ;827] -1.475

 Table 1: Elasticities with Respect to Effective Tax Rates

We convert our event-study coefficient estimates for each outcome at time horizons 0, +1, +2, and +3 into elasticities using a two stage procedure that estimates first estimates the effective tax change event study. Changes in tax rates and outcomes are converted to percentages using the pre-reform mean in -1. 95 percent confidence intervals are obtained using the parametric bootstrap. Column (1) uses total sales, column (2) total inputs, column (3) total employee earnings, column (4) total value of capital, column (5) total costs, column (6) is the number of employees, column (7) the value of land, column (8) the value of plants and machinery, column (9) share of firms using digital accounting, and column (10) state tax revenue. All monetary values are in millions of rupees.

A Online Appendix: Data (Online Publication Only)

A.1 VAT Dataset Creation

A.1.1 Digitizing Tax Rate Schedules

We start the creation of the VAT databases by digitizing the tax rate schedules for each state. All states provide information on their state-specific VAT systems on the website of their Commercial Taxes Department. We take the most recent tax schedule information that is available, which is typically either a separate pdf document or part of the VAT Act, and digitize the entries in state-specific spreadsheets. Most states have multiple schedules, one for each tax rate (e.g. zero-rated, 1 percent, 4 percent, etc.), one or more schedules for items with special tax rates (most commonly for items like alcohol, diesel or tobacco), and one entry for the default tax rate, which applies to all products not explicitly listed elsewhere. Within a schedule, entries are typically numbered consecutively. An entry can consist of just one product or, more commonly, a list of products that belong to a similar category of items. Entries can also contain qualifiers (e.g. items up to a specific monetary value, non-branded or branded items) or additional explanations and descriptions.

This means that entries can differ substantially in their scope. For example, the same schedule for Andhra Pradesh contains specific entries like 'Charcoal', but also far more wideranging entries like 'Agricultural implements manually operated or animal driven, hand operated sprayers including knapsack/backpack power sprayers (powered up to 35 cc engines developing 0.8 to 1.4 HP), dusters and parts thereof.' and items with other restrictions such as 'Unbranded bread and unbranded rusk' or 'Water other than i) aerated, mineral, distilled, medicinal, ionic, battery, de-mineralized water, and ii) water sold in sealed container'. Importantly, there is no classification system of any kind that guides the scope or level of detail of what appears on these lists. While some products and product descriptions are similar across states, likely due to the influence of the VAT Model Act by the central government, many entries are state-specific. States also differ in the length of their schedules.

A.1.2 Tracing Product Description and Tax Rate Changes over Time

Digitization of the tax rate schedules for each state gives us a cross-sectional snapshot of the product-specific VAT rates with product entry descriptions and the corresponding tax rate. But tax rates can and do change over time, and the description of products can be altered as well. We trace those changes carefully using all available information. The tax rate schedule documents which we use for digitization often already contain information on some changes, for example in the form of footnotes on specific items or schedule headings. This information can affect our database in multiple ways. First, a state may change the tax rate on all products that are part of a specific schedule, for example by raising the tax rate from the 4 percent schedule to 5 percent. In this case, a footnote on the schedule heading may give the date of the change as well as information on when the change went into effect. We keep track of both pieces of information when available.

Second, a state may add new products to a schedule that were not previously listed anywhere else. This means that the product was previously subject to the default tax rate, but will now be explicitly listed and receive a typically lower tax rate. Such items are either added to a schedule at the end with a new number, or are added elsewhere in a schedule by being appended to an existing entry of a similar category, or added as a separate item with an item number like 48A to indicate that the entry is similar to item 48. In our database, we create a new cell for these products and list it has having the default tax rate initially, with a tax rate change to the tax rate in the specific schedule once it is explicitly mentioned. As before, we keep track of all of the available date information on the tax rate change.

Third, a state may move products from one schedule to another. Most commonly, this shows up in the schedule as a product being deleted in one schedule on a specific date and then added to a different schedule on the same date. In our database, we keep track of this by noting the date of the tax rate change and the previous and new tax rates based on the schedules the product is listed on.

Fourth, a state may delete products without listing them elsewhere. This is equivalent to

assigning the product to the (typically higher) default tax rate. In our database, we therefore assign these products the default tax rate from the date of the change.

Fifth, a state may adjust the description of an entry. This can involve adding clarifying examples, making the description more precise, or adding or removing qualifiers. In those cases, the footnote information will either list the previous description or will refer to a specific notification. In those cases, we carefully compare the previous and new descriptions. Sometimes, the changes are purely semantic or are meant to provide clarifying examples, without any implications for product scope. In those cases, we ignore the change. Other times, the changes alter the range of products covered by the entry. For example, inserting the word 'unbranded' before an item like bread means that brand-name bread would now de facto be subject to the default tax rate unless it is explicitly added elsewhere. In this case, we break the entry 'bread' into two cells, one for branded and one for unbranded bread. Up until the insertion of the word 'branded', both entries will have the same tax rate information, but afterwards we will trace further tax rate changes separately.

Sixth, a state may split or combine entries. Unless these entries later receive different tax rates, we ignore this information.

Digitizing this information provides us with some information about tax rate changes, but it is typically incomplete. This is because the VAT schedules posted on the website may be outdated, so that more recent tax rate changes are not included, because the schedule document does not list all of the prior tax rate changes (e.g. only focusing on recent changes since the previously published version of the document), or because the tax rate changes are missing key pieces of information like the date of the change or the previously applicable tax rate. We solve those problems by manually going through all of the available VAT notifications on the Commercial Taxes Department websites, typically available as scanned pdf documents, which contain this information, and update our database accordingly.

At the end of this process, we have a state-specific harmonized dataset with all of the available information from the state websites, where we track every product that ever appears in a schedule for the entire time between the VAT introduction and 2016. This is a productlevel panel dataset where the harmonized description comes from adjusting the entries in the steps above for any important changes.

A.1.3 Cross-Check I: VAT Published Volumes

For the purposes of our paper, having high-quality information for the early years after VAT adoption is very important. But occasionally, online links to older notifications are broken or documents corrupted, or it is not always clear whether the online notification archives are complete going back a decade. We therefore cross-check and complement our database with information from two collected volumes of state VAT systems, one from 2005 and one from 2007 (Sangal, 2005; Sangal and Goel, 2007). These books contain snapshots of the VAT systems of any states that had implemented the reform by the publication date, as well as details on all notifications since the VAT introduction. They allow us to cross-check the tax rates at the VAT implementation start date. For the few states for which we miss some notifications in the online archives for a time period, this allows us to considerably narrow the time window of any missing tax rate changes. This mostly concerns the small north-eastern states of Sikkim and Nagaland, where the online archives are missing any information on VAT rate changes before 2010. We have to drop Sikkim in our firm-level analysis because firms from that state were only surveyed in post-reform years.

A.1.4 Cross-Check II: Detecting and Resolving Internal Inconsistencies

Since much of the database creation relies on manual harmonization, which could be vulnerable to human error and subjective judgement calls, we cross-check the information in our database in multiple ways. First, coders received extensive training on the general process as well as on specific examples of problem cases and how to resolve them. In regular group meetings, examples and new problems were discussed, ensuring uniformity across states and coders. Second, coders flagged any cases where judgement calls were required for group discussion. Third, all coding was spot-checked by at least two other persons, and particularly complicated states were coded up separately by two or more coders to allow for comparison. Fourth, for the notifications one person was working backwards in time from the most recent available information and another person was working forwards from the VAT introduction, allowing us to catch inconsistencies or missing information along the way. Fifth, multiple people carefully compared the product descriptions for each state from the 2005 and 2007 books entry by entry to the ones in our database to ensure that the harmonized database included all of the information contained in the books and no information had gone missing when splitting or merging entries.

After this manual quality control, all state databases were subjected to a Stata script checking for internal inconsistencies in the dataset. The script created a product-specific month-year panel. If no tax rate change occurred, the script puts in the tax rate from the previous month. This flags any cases in which an initial tax rate is missing or in which the previous tax rate noted in our database from our tax rate change documentation does not match the imputed tax rate from the previous month for further review. Similarly, the program flags any non-numeric characters. Inconsistencies were then resolved by going back to the specific source information. In the few cases where inconsistencies could not be fully resolved (e.g. because of missing information), this is documented in state-specific readme files and the best available information is included in the database.

A.1.5 Cross-Check III: Cross-State Data Cleaning and Quality Checks

In a final round of data cleaning and data quality checks, all state tax rate databases were screened for other potential issues separately by multiple individuals. First, occasionally 'copy and paste' mistakes could occur, for example if all products in a specific schedule changed tax rates but some products are incorrectly included or excluded, or if a tax rate change applicable to one specific product is accidentally entered in the incorrect row. Since products in the same schedule mostly remain grouped together even through the harmonization process, a careful line-by-line inspection of the databases reveals such problems. From context, it may also be clear that a tax rate change should apply to the row above or below, which can then be confirmed by going back to the notification archive.

Second, products that changed schedules and are therefore not necessarily grouped with similar entries at all times could be 'missing' a tax rate increase. A product could have been moved from the default to the 4 percent schedule at some point, for example, but then the entire 4 percent schedule had a tax rate change to 5 percent, which was not entered for this product because there was no product-specific notification on this tax change. Again, a careful line-by-line inspection of items allowed for detecting and resolving these issues.

Third, each state-specific readme document was updated and checked for any remaining actionable issues.

Fourth, the default tax rate information was double-checked against all information from notifications and media reports. Since the residual tax rate applies to all goods that are not explicitly listed in the schedules, it is important to capture all tax rate changes to the defauly rate. Occasionally, for example, states did not report the default tax rate in the schedules, but instead listed it elsewhere in the VAT Act. Almost all states had a default tax rate of 12.5 percent at VAT implementation, but later increased it by varying amounts. Since such a tax rate change affects so many products, this change was also typically reported in the news, which allowed for cross-checking of the information in our database.

Fifth, VAT schedules include an entry for industrial inputs, which typically receive a subsidized VAT rate, but not all states provide a list of what counts as industrial inputs in the schedules themselves. In such cases, the lists were added and traced over time from the information elsewhere in the VAT Act.

Sixth, the Stata internal inconsistency script was run again to ensure that any changes had not introduced new issues to resolve.

This completes the creation of the VAT databases, containing information from the year of VAT adoption in a given state up to 2016.

A.2 Sales Tax Dataset Creation

We are limited in our ability to trace sales tax rates over time since this information is typically not available from online sources in any complete or consistent form. It is typically possible, for example, to find the original sales tax Act, often going back to the 1950s, but almost no state has updated schedules and legal documents for the time period shortly before the VAT introduction. Similar to the VAT regime, we therefore rely on two publications containing multiple volumes of state documents and tax rate schedules from 2000 and 2002 (Commercial Law Publishers 2000; Commercial Law Publishers 2002).

We digitize this information for 2000 and 2002, keeping track of any tax rate changes or product description changes in the same way as for the VAT datasets above. This gives us a harmonized product-specific panel of tax rates for 2000 to 2002, before any state implemented the VAT. Unfortunately, the publications are missing for some of the small states or Union Territories. Sales tax rate information is unavailable for Meghalaya, Tripura, Dadra and Nagar Haveli, and Daman and Diu. These data gaps do not affect many firms in our database because all of these are either very small north-eastern states or Union Territories.²¹

A.3 Other Information from VAT and Sales Tax Documents

In our analysis, we mainly rely on our tax rate database. However, we also collect information on other aspects of the tax regimes. For both sales tax and VAT systems, the published volumes allow us to code up state-specific registration thresholds for firms, which typically depend on firm turnover. We code up all general thresholds in addition to any additional rules, such as different thresholds for specific industries or firm characteristics. The online notification archives for the VAT regime also allow us to trace any changes to those thresholds over time. In addition to the VAT registration thresholds, we track turnover tax thresholds for which firms can pay a (generally) lower rate on their turnover rather than the VAT. We also track other thresholds for which firms are required to pay additional fees or other taxes.

While we only use measures of tax complexity drawn from our tax rate databases in the

paper, we also use the published volumes to note down other proxies for tax complexity, like the number of pages of the Act, the length of the tax rate schedules, as well as the number of chapters and sub-sections.

A.3.1 Matching Tax Rate Information to ASI Data

Unfortunately, neither the sales tax nor the VAT schedules contain any product classification codes that would allow for an easy merging of the tax rate databases to the ASI data. We therefore need to match the tax databases to the product classification system used in the ASI data. We do so in a multi-step process.

A.3.2 Create Master List of ASI Product Codes and Descriptions

For most of the time period covered in our firm dataset, product classification of inputs and outputs is based on the so-called ASICC system. In 2011, this changes to the NPC-MS classification system, and before that, ASICC codes are occasionally updated or added. We use crosswalks between these classification systems so that we have a harmonized set of products over time. We settle on 3-digit ASICC product codes as the most feasible level of granularity of product information and create a master list of all 3-digit ASICC codes and their descriptions.

A.3.3 Use Fuzzy Matching Algorithm

We now have two sets of lists: A master list of ASICC codes from the firm data, and two sets of tax rate lists for products for each state, one for the sales tax and one for the VAT regime. To find the best matches between list items, we first use the excel add-on fuzzylookup to provide fuzzy matches between the lists, with the similarity threshold set to 75 percent and the maximum number of matches set to 20. One of the challenges of the matching process is that the tax rate database often contains a long list of items as one entry that then needs to be matched to the ASICC database that may just contain one of the items from the list. Fuzzylookup performs quite well in making those matches successfully compared to other fuzzy matching algorithms.

A.3.4 Manual Cross-Check and Internal Consistency Checks

We manually check the fuzzy matching results for false matches. This leaves the unmatched ASI products that need to be matched to tax rate information. The three main reasons for non-matches are spelling differences, differences in product level aggregation (i.e. the tax rate dataset may contain product descriptions at a less or more specific level than the firm data), and products that are not mentioned in the tax schedules and should receive the residual tax rate. Differences in product level aggregation is the most common issue, and whether the tax dataset is more detailed than the firm data differs by state, time period and products, so cannot be easily solved by creating a dictionary of synonyms. We therefore create manual matches, with extensive cross-checking by multiple separate coders.

B Online Appendix: Generating Complexity Index and Effective Tax Rates

B.1 Complexity Index

Lubotsky and Wittenberg (2006) develop an index creation procedure when researchers have access to multiple proxy variables. In our setting, complexity is unobserved but we have multiple proxies for complexity (number of tax rates, number of items, number of schedules, number of pages, etc.) that can then arguably linked to compliance costs.

To summarize the method, let c^* denote the unobserved (underlying) measure of complexity. We have access to proxy variables indexed k = 1..K for complexity, denoted c^k . The proxies we have access to vary at the state-time level. Ultimately, we wish to identify the effect of c^* on our outcome variable, $z_{f(s),t}$ for firm f in state s in year t. To construct the index, we use all K proxies, regressing $z_{f(s),t}$ on the usual fixed effects and all of the proxies $c_{s,t}^k$, then recovering the coefficients on each proxy, β^k . We then aggregate up to construct an index for complexity, denoted $c_{f(s),t}^{\rho}$ using the formula in Lubotsky and Wittenberg (2006):

$$c_{f(s),t}^{\rho} = \frac{1}{\beta^{\rho}} \sum_{k=1}^{K} \beta^{k} c_{s,t}^{k},$$
(B.1)

where

$$\beta^{\rho} = \sum_{k=1}^{K} \beta^{k} \frac{cov(z_{f(s),t}, c_{s,t}^{k})}{cov(z_{f(s),t}, c_{s,t}^{1})}.$$
(B.2)

The denominator of this last expression normalizes the index so that it is in the same units of proxy variable 1. We then standardize the variable so that changes are denoted in terms of standard deviations of that variable. As proxy variables, we use the number of tax schedules, the number of items in the schedules, the number of pages in the law, the number of chapters and subsections, the share of items in the residual category, the total number of products that change tax rates and the number of schedules that change tax rates. As a normalization we use the number of tax schedules, given it is one of the raw variables we depict in the text.

We use the procedure of Lubotsky and Wittenberg (2006) rather than single proxy variables individually because Lubotsky and Wittenberg (2006) show that their index is the one that minimizes any bias from measurement error. Of course, as noted in their paper, one should not include every possible proxy, especially if the variable has a direct effect on the outcome via some other channel other than complexity. For this reason, we explore robustness of the measure to including/excluding different variables. The results are robust and qualitatively similar.

B.2 Effective Tax Rates

As discussed in the text, due to the passage of additional sales taxes at various stages of production, taxes on inputs and outputs under the sales tax regime need not be at the same rate. These additional taxes outside of the sales tax schedules are not observed to us. Under the VAT regime, cascading primarily results from imperfections in the credit system. We model tax cascading and different tax rates on inputs/outputs as a factor of proportionality,

 α_t , on the statutory tax rate. Then, $\alpha_t \tau_{f(s),t}^I$ is the resulting tax rate on inputs. Recall that $\tau_{f(s),t}^I$ is the firm-specific effective tax rate on the goods that the firm uses as inputs—this rate is determined using the rates in the sales and VAT acts. In other words, if $\alpha = 0.5$, the input is taxed at 50% of the rate that the good would taxed at if it were a finished product.

As indicated by the subscripting, we make the strong assumptions that all inputs in the firm have the same factor and that the factor is the same across all states (though it may differ under sales and VAT regimes—explaining why α is only subscripted by t). This assumption, while strong, is necessary given that tax rates on inputs and imperfections in the credit regime are not observed to us at the state, product, or industry level. Although it differs across tax regimes, we assume it is the same for all years within each regime.

The question is what the levels of α_t are before and after the reform. Unfortunately, this cannot be determined in the ASI data nor can it be determined in our tax rate database. The reason is that cascading in the Indian system often results from added ad hoc legislation outside of the main sales tax acts—that create additional product-specific excise taxes or as a result of imperfections in the VAT credit system not apparent at the product or industry level. Thus, for our preferred specification we use alternative aggregated records to determine the value of α_t before and after the reform.

For the post-reform period, we rely on Poddar and Ahmad (2009) who make comparisons of the Indian system to other countries, to argue that Indian VAT regime had a cascading rate of 35-40%. Under the sales tax regime, Bagchi and Team (1994) and Chelliah and Committee (1992) suggest that cascading was more severe, but neither report identifies a precise numerical value concerning the extent of cascading. Thus, we draw on a report by Aggarwal (1995) who calculates effective tax rates on select goods using tax rate data from several states. We match his effective tax rates to our statutory tax rates and then back out a cascading factor. Doing so, yield and average α for the sales tax regime of 0.98. Note that it is possible α could be greater than one if the separate legislation for input taxes had rates hire than the tax rate at the point of sale. To verify the robustness of our results, we allow α_t to vary over different values in the sales tax and VAT regime. This allows us to show how our effective tax rates and elasticities change as a result of changes on the assumptions of the cascading rate. See Appendix Table C.6 and Figure C.12.

It is important to compare our approach to an analysis that exploits input-output tables (De Loecker et al., 2016) instead of the ASI. Input-output matrices can be a useful way to construct effective tax rates, and indeed have been done so in other contexts. India does publish input-output coefficients that can be linked to our data. However, these coefficients come with the standard limitations—that they are derived from stylized assumptions and they eliminate all heterogeneity within industries. Our ANOVA calculations indicate that tax rate heterogeneity, resulting from different product/input mixes, within industries is very important. Further, in the case of India, input-output coefficients are calculated at the national level; regional product coefficients are unpublished, meaning that we cannot not derive state level coefficients from them. The substantial within industry heterogeneity could be correlated with sales tax systems, thus giving way to our preferred method of using the ASI. However, the advantage of the input-output table approach relative to ours is that it would capture all stages of production, where as we only capture the immediately prior round of inputs used in production and the finished product tax rate stage—though again, the Indian system likely specified different implicit tax rates depending on the stage of production. Thus, even if taking an input-output approach, we would need to make the same assumptions on α_t , but would lose within industry heterogeneity. We verify that our results are robust to focusing on industries with few linkages in the production process, which provides evidence that adding many more stages of production comes with limited gains.

- C Online Appendix: Additional Results
- C.1 Additional Tables and Figures

Figure C.1:	Example of	One Page :	from a Tax	Schedule
-------------	------------	------------	------------	----------

SALES	TAX RA	TES IN	WEST	BENGAL	
					-

26.9

SCHEDULE IV

[Sections 10 and 17(1)(c)] GOODS ON SALE OF WHICH TAX IS LEVIABLE AT SUCH RATE AS MAY BE FIXED BY NOTIFICATION UNDER SECTION 18 (SINGLE-POINT LEVY) READ WITH SUB-CLAUSE (a) OF CLAUSE (40) OF SECTION 2 (As amended up to 2002)

PART A

SI. No.	Description of goods	Rate of tax (per cent)
(1)	(2)	(3)
1.	Aerated water including soda water and non-alcoholic beverage (excluding fruit juice, fruit concentrate, fruit squash, fruit syrup and fruit cordial) when sold in sealed containers including sealed polythene bottles, pouches or packets.	Seventeen
1A.	Fruit juice, fruit concentrate, fruit squash, fruit syrup and fruit cordial when sold in sealed containers including sealed polythene bottles, pouches or packets	Seven
2.	Air-conditioner and air-cooler.	Twelve
3.	(a) Aluminium in all its forms, namely, aluminium ingots, slabs, bars, rods, pipes, tubes, wires, coils, sheets, plants, circles, sections, channels, angles, joists, extrusions, including aluminium scraps but excluding aluminium foils	Five
	(b) Aluminium pipes and tubes	Seven
4.	Aluminium foils including aluminium foils backed or interleaved with paper or any other substance.	Five
5.	Betel nuts known locally as supari, whole, broken, perfumed or otherwise treated or of any other form or description whatsoever.	Twelve
6.	Binocular, telescope and opera glass.	Four
7.	Biscuits of all varieties, except those specified in any other Schedule	Ten
8.	Bleaching powder of all varieties and descriptions	Five
<u>9</u> ,	Bulldozer, excavator, pipe-layer and scrapper.	Twelve
10.	[Omitted]	
11.	Cathode ray tubes used in televisions and commonly known as picture tubes.	Fifteen

Legend: From Commercial Law Publishers (2002). This figure shows one page from West Bengal's sales tax schedule. West Bengal's schedules amount to over twenty five pages in length.

GOVERNMENT OF KARNATAKA NO.FD.121 CSL.2006 Karnataka Government Secretariat Vidhana Soudha, Bangalore, Dated:10.05.2006. NOTIFICATION-II In exercise of the powers conferred by section (1) of section 5 of the Karnataka Value Added Tax Act, 2003 (Karnataka Act 32 of 2004), the Government of Karnataka hereby exempts with immediate effect, the tax payable by a dealer under the said Act on the sale of goods specified below, namely:-Coir and Coir Products, but excluding rubberized coir products (1) (2) Honey (3) Hone wax (4) Single moulded Plastic footwear fully made of plastic, costing less than Hundred rupees per pair. (5) Turmeric stems and bulbs, but excluding turmeric in any other form. By Order and in the Name of the Governor of Karnataka (R.S.ITAGI) Under Secretary to Government Finance Department (CT-I) 7

Figure C.2: Example of a Notification of a Tax Rate Change

Legend: From the Government of Karnataka Commercial Taxes Department. Highlight in original document. This figure shows one notification of a tax rate change for Karnataka's VAT system.



Figure C.3: Share of Products Not Subject to Default Rate

(a) Unweighted

(b) Weighted by Sales

Legend: This figure reports the share of products not subject to the default rate at various points in time. To properly report the proportion, we construct a panel version of goods and their tax rates that contains the universe of goods. To do this, we create a master list of all potential products based on the 3-digit ASICC classification system used in the ASI firm data. We then match tax rate schedules by state both before and after the reform to that master list. Some ASICC categories have a unique match to a tax rate in a given state in a given year, others will have multiple tax rate matches if specific products within the ASICC category are subject to different tax rates. To construct the proportion of goods not subject to the default rate, we rely on the most common tax rate for each ASICC code. Panel (a) present the unweighted statistics while Panel (b) presents the statistics weighted by sales.

Figure C.4: Covariate Adjusted Regression



Legend: This figure shows effect of VAT adoption using open firms in the ASI database after adjusting the first stage for covariates. The covariates we include are all of the variables in Table C.2 Column (2), with the exception of "riots" which is missing for many years. Covariates only enter into the first-stage and thus do not utilize post-reform information. Period -1 is the omitted category. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure C.5: Effect of VAT Reform on Firm Sales and Input Purchases (Log Transformed)

Legend: This figure shows effect of VAT adoption using open firms in the ASI database. Period -1 is the omitted category. In Panel (a) we present the results for sales, while in Panel (b), we present the results for inputs. In each figure, we have two alternative specifications. First, for the log series (diamonds), the log takes the natural log of a given variable, thereby dropping 0s from the sample. The alternative is the inverse hyperbolic sine transformation (circles), which yields a similar interpretation as taking logs, but is defined for 0 values. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.

Figure C.6: Effect of VAT Reform on Manufacturing GDP



Note: This Figure shows effect of VAT adoption total GDP in the manufacturing sector available from from Niti Aayog website (GSVA/NSVA). For purposes of making this figure, we present results using the log of manufacturing GDP in each state. These data include both information on registered firms under the Factories Act and unregistered firms. Period -1 is the omitted category. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure C.7: Effect of VAT Reform: Tax Revenue

Legend: Tax revenue comes from the Reserve Bank of India. This figure shows the Gardner (2021) event study where the outcome variable is state consumption tax revenues. Period -1 is the omitted category so no confidence bands are included. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure C.8: Heterogeneity by Backward and Forward Linkages

(a) Backward Linkages

(b) Forward Linkages

Legend: This figures shows the heterogeneity of the effects on sales by backward and forward linkages of the firm's industry. Industry level data on backward and forward linkages are from Hoseini and Briand (2020) and our specifications in this figure cut firms into samples based on if the industry has a high or low amount of linkages. Period -1 is the omitted category so no confidence bands are included. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure C.9: Event-Study Graphs for Elasticity Outcomes

Legend: This figure shows the event study graphs for all outcomes in Table 1. Standard errors are clustered at the state $\frac{1000}{22}$ and the figure shows 95% confidence intervals.



Figure C.10: Event-Study Graphs for Elasticity Outcomes Continued

Legend: This figure shows the event study graphs for all outcomes in Table C.5. Standard errors are clustered at the state₂ level and the figure shows 95% confidence intervals.




Legend: This figure shows effect of VAT adoption on firm age using open firms in the ASI database. Period -1 is the omitted category. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure C.12: Effective Tax Rate Changes Under Different Assumptions of α

Legend: This figure shows the range of tax rate changes under extreme assumptions about the extent of cascading in the VAT and sales tax regimes. Our baseline specification, calibrated using sources on cascading, is approximately given by the circular line. Period -1 is the omitted category. Standard errors are clustered at the state level and the figure shows 95% confidence intervals.



Figure C.13: Effective Tax Rates with Variation in α_t by Firms, Goods, and States.

Legend: In this figure, we randomize a factor by which our baseline value of alpha is multiplied by. We conduct three randomization at the firm level, primary good of production level, and state level, separately for each of the sales tax and VAT periods. The first row labeled by "sales tax" are the randomizations where sales tax is randomized and VAT is kept fixed. The second row labeled by "VAT" are those where VAT is randomized and sales tax is kept fixed. Each random factor for each group is drawn from a uniform distribution on the interval [0, 2], such that the mean random factor is 1, approximately preserving our baseline value at the mean. We repeat each randomization 100 times and then ran the Gardner (2021) specification with effective tax rate as outcome variable. The vertical line is the "baseline" coefficient.



Figure C.14: Standard Deviation of Tax Rates Across States

Legend: This figure aggregates effective tax rates to the state by year level and then plots the standard deviation of tax rates by year. The horizontal axis is year, rather than event year, as we are interested in how adoption of simultaneously affects tax rates in other states in that year of adoption (and not in event time).

State	Month	Day	Year
Andhra Pradesh	4	1	2005
Arunachal Pradesh	4	1	2005
Assam	5	1	2005
Bihar	4	1	2005
Chandigarh	12	15	2005
Chhattisgarh	4	1	2006
Dadra and Nagar Haveli	4	1	2005
Daman and Diu	4	1	2005
National Capital Territory of Delhi	4	1	2005
Goa	4	1	2005
Gujarat	4	1	2006
Haryana	4	1	2003
Himachal Pradesh	4	1	2005
Jammu and Kashmir	4	1	2005
Jharkhand	4	1	2006
Karnataka	4	1	2005
Kerala	4	1	2005
Madhya Pradesh	4	1	2006
Maharashtra	4	1	2005
Manipur	8	1	2005
Meghalaya	5	1	2005
Mizoram	4	1	2005
Nagaland	4	1	2005
Orissa	4	1	2005
Puducherry	4	1	2007
Punjab	4	1	2005
Rajasthan	4	1	2006
Sikkim	4	1	2005
Tamil Nadu	1	1	2007
Telangana	4	1	2005
Tripura	4	1	2005
Uttar Pradesh	1	1	2008
Uttarakhand	10	1	2005
West Bengal	4	1	2005

Table C.1: Date of VAT introduction in Indian States

Note: Two states or Union Territories, A28 aman and Nicobar Islands and

Lakshadweep, have never had a sales tax and did not introduce a VAT.

Outcome: VAT Adoption Year	(1)	(2)	(3)	(4)
Aligned with Center	-0.1683	-0.3912	-0.1045	-0.1087
	(0.3237)	(0.4423)	(0.4637)	(0.4158)
Congress Seat Share	-0.0099	-0.0066	-0.0009	-0.0002
	(0.0059)	(0.0066)	(0.0067)	(0.0065)
BJP Seat Share	0.0157**	0.0170*	0.0030	0.0087
	(0.0074)	(0.0091)	(0.0129)	(0.0113)
Time since Last Election	0.0847	0.0533	0.1378	0.1864*
	(0.0827)	(0.1101)	(0.1139)	(0.0890)
Around Election Indicator	0.9262***	0.5611	0.8900*	0.8162*
	(0.3092)	(0.3376)	(0.4281)	(0.4004)
GDP Growth Rate		-0.0459	-0.0407	-0.0448
		(0.0316)	(0.0419)	(0.0397)
Riots		-0.0277	-0.0296	-0.0282
		(0.0304)	(0.0328)	(0.0322)
Population		0.0071	0.0092	0.0092
		(0.0053)	(0.0056)	(0.0063)
Firm Sales			0.0078	0.0079
			(0.0077)	(0.0064)
Firm Inputs			-0.0205	-0.0305
			(0.0278)	(0.0224)
Firm Sales Difference (1998-2002)			0.0045	0.0068
			(0.0091)	(0.0102)
Firm Inputs Difference (1998-2002)			0.0037	0.0154
			(0.0332)	(0.0316)
Sales Tax Revenue				0.0005
				(0.0007)
Observations	27	27	27	26
R squared	0.3329	0.4160	0.5356	0.7343

Table C.2: Determinants of VAT Adoption Timing

Specification follows Bailey and Goodman-Bacon (2015) to explore determinants of adoption timing. Unless noted otherwise, all variables measured in 2002. Aligned with Center is an indicator for whether a state government is aligned with the central government. Congress and BJP are the two big national parties. Time since Last Election measured in years since last state election. Around Election Indicator is equal to 1 before and after a state election, and 0 otherwise, to account for political incentives to adopt right before or after an election. GDP Growth Rate refers to state economic growth rate, downloaded from Niti Aayog website. Information on riots comes from the National Crime Records Bureau, population from the 2001 Census. Firm sales and inputs in million rupees come from the ASI data. Sales Tax Revenue refers to state sales tax revenue in millions of rupees, collected from Reserve Bank of India. Robust standard errors in parentheses. ***99%, **95%, *90%.

	State Sales (in million Rupees)				
	(1)	(2)	(3)		
Time Period	1998	2002	2002 - 1998		
VAT Adoption Year	19.31	25.18	5.87		
	(14.35)	(17.48)	(6.38)		
	State Inputs (in million Rupees)				
	(1)	(2)	(3)		
Time Period	1998	2002	2002 - 1998		
VAT Adoption Year	4.44	3.27	-1.17		
	(5.02)	(6.22)	(2.42)		

Table C.3: State-Level Sales and Value of Inputs on VAT Adoption Timing

Specification follows Bailey and Goodman-Bacon (2015) to explore correlation of pre-reform level and long-difference in our outcome variables with VAT adoption year. Outcome variables of sales and value of inputs are drawn from the ASI and collapsed to the state level using sampling weights. Robust standard errors in parentheses. ***99%, **95%, *90%.

Specification	(1)	(2)
	Sales	Input Value
π_{-5}	1.6969	0.0166
	(1.1214)	(0.5336)
π_{-4}	0.7720	0.6774
	(1.9451)	(0.8160)
π_{-3}	0.1899	0.2278
	(1.5184)	(0.9240)
π_{-2}	-2.8943**	-1.8368**
	(1.4513)	(0.8842)
π_{-1}	-	-
γ_0	11.8163	5.2824
	(12.3572)	(5.2005)
γ_1	46.4043**	18.1049***
	(18.9886)	(5.6135)
γ_2	89.5235***	38.8445***
	(13.3885)	(5.1766)
γ_3	78.9238***	38.9746***
	(8.7024)	(3.6805)
Overall ATT	52.2965***	26.4725***
	(7.8572)	(3.4377)
Observations in -5 to 3	299,320	299,320
Total Observations	486,664	486,664

Table C.4: Event Study Coefficients

This table presents coefficient estimates from the Gardner approach. All models are run using a full set of event dummies; event year -1 is the omitted year. The outcome variable is sales and input values, respectively. The overall ATT is the coefficient on a specification with a treat by post interaction. All regressions are weighted with the sampling weights in the ASI. Standard errors are clustered at the state level. ***99%, **95%, *90%.

			Panel A		
	(1)	(2)	(3)	(4)	(5)
Outcome	Male	Female	Buildings	Transport	Computers
Elasticity 0	0.0533	0.152	-0.362	-0.00136	0.513
CI 0	[13;.247]	[641;.974]	[62;173]	[266 ; .264]	[157; 1.282]
Elasticity 1	0.0247	-0.00287	-0.395	-0.230	0.130
CI 1	[134;.185]	[942;.935]	[611;215]	[474 ;007]	[541 ; .812]
Elasticity 2	0.0894	0.272	-0.419	-0.464	0.449
CI 2	[087;.271]	[93; 1.489]	[622 ;242]	[676 ;281]	[922; 1.845]
Elasticity 3	0.0898	0.221	-0.621	-0.598	0.208
CI 3	[097;.282]	[922; 1.376]	[942 ;337]	[84 ;392]	[-1.108; 1.535]
			Panel B		
	(6)	(7)	Panel B (8)	(9)	(10)
Outcome	(6) Input Share	(7) Labor Share	Panel B (8) Other Share	(9) Top Sales	(10) Top Input
Outcome Elasticity 0	(6) Input Share -0.0198	(7) Labor Share 0.136	Panel B (8) Other Share -0.0429	(9) Top Sales 0.902	(10) Top Input 1.001
Outcome Elasticity 0 CI 0	(6) Input Share -0.0198 [05 ; .006]	(7) Labor Share 0.136 [.015 ; .284]	Panel B (8) Other Share -0.0429 [133 ; .039]	(9) Top Sales 0.902 [.461 ; 1.517]	(10) Top Input 1.001 [.565 ; 1.629]
Outcome Elasticity 0 CI 0 Elasticity 1	(6) Input Share -0.0198 [05 ; .006] -0.0185	(7) Labor Share 0.136 [.015 ; .284] 0.0763	Panel B (8) Other Share -0.0429 [133 ; .039] 0.0636	(9) Top Sales 0.902 [.461 ; 1.517] 0.326	(10) Top Input 1.001 [.565 ; 1.629] 0.387
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1	(6) Input Share -0.0198 [05 ; .006] -0.0185 [06 ; .021]	(7) Labor Share 0.136 [.015 ; .284] 0.0763 [055 ; .215]	Panel B (8) Other Share -0.0429 [133 ; .039] 0.0636 [01 ; .143]	(9) Top Sales 0.902 [.461 ; 1.517] 0.326 [.002 ; .68]	(10) Top Input 1.001 [.565 ; 1.629] 0.387 [.096 ; .715]
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1 Elasticity 2	(6) Input Share -0.0198 [05 ; .006] -0.0185 [06 ; .021] 0.0452	(7) Labor Share 0.136 [.015 ; .284] 0.0763 [055 ; .215] -0.119	Panel B (8) Other Share -0.0429 [133 ; .039] 0.0636 [01 ; .143] -0.0241	(9) Top Sales 0.902 [.461 ; 1.517] 0.326 [.002 ; .68] -0.148	(10) Top Input 1.001 [.565 ; 1.629] 0.387 [.096 ; .715] -0.00923
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1 Elasticity 2 CI 2	(6) Input Share -0.0198 [05 ; .006] -0.0185 [06 ; .021] 0.0452 [.005 ; .088]	(7) Labor Share 0.136 [.015 ; .284] 0.0763 [055 ; .215] -0.119 [242 ;005]	Panel B (8) Other Share -0.0429 [133 ; .039] 0.0636 [01 ; .143] -0.0241 [119 ; .069]	(9) Top Sales 0.902 [.461 ; 1.517] 0.326 [.002 ; .68] -0.148 [249 ;055]	(10) Top Input 1.001 [.565 ; 1.629] 0.387 [.096 ; .715] -0.00923 [092 ; .073]
Outcome Elasticity 0 CI 0 Elasticity 1 CI 1 Elasticity 2 CI 2 Elasticity 3	(6) Input Share -0.0198 [05 ; .006] -0.0185 [06 ; .021] 0.0452 [.005 ; .088] 0.105	(7) Labor Share 0.136 [.015 ; .284] 0.0763 [055 ; .215] -0.119 [242 ;005] -0.232	Panel B (8) Other Share -0.0429 [133 ; .039] 0.0636 [01 ; .143] -0.0241 [119 ; .069] -0.183	(9) Top Sales 0.902 [.461 ; 1.517] 0.326 [.002 ; .68] -0.148 [249 ;055] -0.0878	(10) Top Input 1.001 [.565 ; 1.629] 0.387 [.096 ; .715] -0.00923 [092 ; .073] 0.0380

 Table C.5: Effective Tax Elasticities for Additional Outcomes

This table presents elasticities for additional outcomes. We convert our event-study coefficient estimates for each outcome at time horizons 0, +1, +2, and +3 into elasticities using a two stage procedure that estimates first estimates the effective tax change event study. Changes in tax rates and outcomes are converted to percentages using the pre-reform mean in -1. 95 percent confidence intervals are obtained using the parametric bootstrap. Columns (1) and (2) refer to number of directly employed male and female workers, (3) is value of buildings, column (4) value of transport equipment column (5) value of computer equipment, including software, column (6) is material inputs as a share of total costs, column (7) labor share out of total costs, column (8) share of all other factors (not material inputs, not capital, and not labor) out of total costs, column (9) is the top produced good as a share of total production, and column (10) is the share of the top input out of total costs. All monetary values are in millions of rupees. ***99%, **95%, *90%.

			Panel A			
Varying Value of VAT α	1	0.90	0.80	0.70	0.60	0.50
Sales Tax α at 1	-1.153	-1.097	-1.046	-1	-0.958	-0.919
CI	[-1.701;741]	[-1.598;711]	[-1.508 ;683]	[-1.429;657]	[-1.358 ;633]	[-1.293;61]
Varying Value of VAT α	0.40	0.30	0.20	0.10	0	
Sales Tax α at 1	-0.883	-0.850	-0.820	-0.791	-0.765	
CI	[-1.235 ;589]	[-1.182;570]	[-1.134 ;552]	[-1.09 ;534]	[-1.049;518]	
			Panel B			
Varying Value of ST α	1	0.90	0.80	0.70	0.60	0.50
VAT α at 0.40	-0.883	-0.910	-0.941	-0.976	-1.017	-1.066
CI	[-1.235 ;589]	[-1.279 ;605]	[-1.329 ;623]	[-1.388 ;644]	[-1.458 ;667]	[-1.542 ;694]
Varying Value of ST α	0.40	0.30	0.20	0.10	0	
VAT $alpha$ at 0.40	-1.123	-1.193	-1.281	-1.393	-1.541	
CI	[-1.645 ;726]	[-1.775;763]	[-1.943 ;808]	[-2.169 ;863]	[-2.489;931]	

This table shows how our elasticity of sales changes with respect to different assumptions on α_t , the extent of cascading. Panel A holds fixed that value of α_t under the sales tax regime at its baseline specification of 1, but then allows the value of α_t under the VAT regime to range from [0,1]. Panel B holds fixed that value of α_t under the VAT regime at its baseline specification of 0.4, but then allows the value of α_t under the sales tax regime to range from [0,1]. For simplicity, we show this for event year +2 elasticities of sales, but other variables and time periods scale similarly.