Efficiency Effects of Tax Reform in India and the Interaction with Vertical Integration Decisions *

Juan S Holguin

August 3, 2026

Abstract

In this paper, I evaluate the efficiency effects of a tax reform executed in India between 2003 and 2008 intended to transition from a sales tax to a VAT (value-added tax) scheme. Moreover, I explore the interactions between the vertical integration plants' decisions and the reform effects on productivity-related variables. To conduct this analysis, I initially estimate a product-level production function to recover Total Factor Productivity (TFPQ) and marginal costs. Subsequently, standard causal estimation methods are employed to assess the impacts of the tax policy on vertical integration, marginal costs, and TFPQ. The principal findings are as follows: 1) the tax policy serves as an incentive for vertical integration; 2) it induces a broad decrease in marginal costs while leaving TFPQ unaffected; 3) vertically integrated plants demonstrate an increase in both marginal cost and TFPQ compared to their non-vertically integrated counterparts. I interpret my results as evidence of tax reform solving distortions in the input market. Notably, this new environment proves more advantageous for non-vertically integrated plants, elucidating the heterogeneity in the reform effects observed between the two types of plants.

1 Introduction

Value-added taxes (VATs) have become the prevailing tax structure worldwide. More than 170 countries have adopted a VAT system, which has doubled compared to 25 years

*I thank David Rivers, Lisa Tarquinio and Salvador Navarro for their useful comments, suggestions, and constant support. I thank the Indian government for granting me access to the data.

ago. Diverging from the "sales tax" model, VAT is levied at every phase of the production chain. Yet, businesses are eligible for a full refund of taxes on inputs utilized in crafting their final product. The theoretical merits of a VAT system stem from the notion that sales taxes inevitably establish input-price distortions in the input markets. These distortions increase as the production chain length increases, which affects final-good prices. In contrast, when the tax is levied at the final consumption stage (as in the VAT scheme), the outcome yields similar tax revenue collection, and simultaneously, it alleviates potential distortions in both final goods and input prices.

Under this argument, transitioning from sales taxes to VAT benefits firms by reducing the distortions they face along the production chain. A standard view is that after the transition of sales taxes to VAT, incentives to integrate for firms vertically reduce substantially. Specifically, under a sales tax scheme, firms have incentives to produce their inputs to minimize taxable stages. Hence, under sales taxes, vertical integration offers firms a mechanism to internally coordinate efficiently and potentially overcome the double marginalization they face under this tax scheme, giving them more room to compete in the final output price. This incentive disappears under a VAT scheme, allowing firms to allocate their resources to activities in which they are more productive rather than attempting to minimize their tax liability through producing a sub-optimal variety of products, including inputs.

Given that firms reduce their vertical integration decisions, it is plausible to expect a generalized increase in productivity-related measures after transitioning from a sales tax to a VAT scheme. However, little evidence suggests that vertical integration reduction is observed after this tax scheme transition. Indeed, Singh (2019) finds that in the Indian economy, vertical integration prevalence increased after the VAT implementation in 2003. The argument behind this result is that if firms are not vertically integrated under VAT, they must provide receipts for their input purchases to claim their corresponding input tax credit. However, for the final firm in the production chain, since final consumers are not obligated to request receipts, there is an incentive to avoid providing a receipt and thus evade the final VAT payment. Consequently, vertical integration complicates enforcement efforts, as the government may lose track of business-to-business transactions. As a result, retail firms can take advantage of the lack of cross-reporting enforcement by the final consumer (Kopczuk and Slemrod (2006)).

In this paper, I evaluate if the Indian economy evidences productivity gains from VAT in a context where vertical integration increased after a sales tax to VAT reform. Moreover, I study the interaction between tax reform, the vertical integration firms' decisions and its consequences on productivity-related variables. One plausible rationale for this potential interaction is that if actual gains are derived from avoiding vertical integration under a VAT tax scheme, one should observe that vertically integrated firms evidence lower productivity gains after the reform than those not vertically integrated. That is because firms that choose to remain vertically integrated make sub-optimal decisions regarding their goods portfolio production. In a counterfactual scenario where these firms are not vertically integrated, they can focus their production efforts more efficiently on their main outputs instead of producing the required inputs.

This analysis is relevant to understand the unintended effects of a VAT transition policy. As was noticed by Singh (2019), vertical integration in India increased after the VAT reform. If vertical integration decision is associated with tax evasion, as Singh (2019) points out, vertically integrated plants' productivity could be potentially lower, leading to misallocation dynamics, which implies achieving sub-optimal economic aggregates. This paper does not explore misallocation in a structural sense, but it provides reduced form evidence pointing in that direction. This paper is a departure point to analyze the unintended effects of VAT reforms on misallocation.

To validate my hypothesis, I exploit a staggered adoption of VAT across Indian states between 2003 and 2008 to measure the impact of transiting from a sales tax to a VAT on vertical integration and its implications for productivity-related measures such as TFPQ and marginal costs. To this end, I use the Annual Survey of Industries (ASI). This plant-level database includes information about outputs' prices and quantities produced by each plant and inputs' prices and quantities used by each plant. The advantage of using this database is twofold. First, having information about price and quantities allows me to use production function estimation methods to consistently recover product-level markups and back out marginal cost and TFPQ estimates. Particularly, separate information about price and quantities prevents the "output price bias", which is prevalent in production functions estimations using revenue data (Bond et al. (2021)). Second, input usage information allows the construction of an indirect measure of vertical integration following Orr (2022), which is my main variable of interest.

As a first pass, I compute a generalized difference-in-difference estimate of the effects of the reform on my vertical integration indicator. In this step, I follow Cengiz et al. (2019) approach that solves the widely recognized issue of estimating policy effect when standard two-way fixed-effects are used in the context of a staggered implemented policy (Gardner (2022); Callaway and Sant'Anna (2021)). Estimates suggest that changing the tax scheme from a sales tax to a VAT increases the prevalence of vertically integrated plants by around 7%.

Using a production function estimation approach, I subsequently estimate the markup and marginal cost at the product level. The importance of the marginal cost analysis in this context rests in that it helps to understand the complete picture of the policy change effects. For example, evaluating the policy implications regarding marginal costs helps to understand how plant markups change due to this policy adoption. Since markups are a valuable tool to approximate the market competition level in the economy, my analysis allows me to disentangle whether this policy induces changes in this variable through tougher price competition or changes in cost efficiency. In this step, I followed De Loecker et al. (2016) to recover consistent estimates of the product-level production function output elasticities. These estimates are critical to recovering markups, as De Loecker and Warzynski (2012) show. With markups on hand, it is straightforward to back out marginal costs, considering that the markups are defined as the ratio between price and marginal costs. In this step, I also computed a generalized difference-in-differences-in-differences combined with instrumental variables to evaluate how the policy affected the plant's efficiency dimension approximated by its marginal costs and TFPQ. Results show that the transition from sales tax to VAT reduced plants' marginal cost by 11%

In the subsequent phase, I explore the potential interplay between my findings on vertical integration, marginal costs, and TFPQ. The rationale behind this potential interaction is that if plants opt for vertical integration to circumvent the distortions implied by the sales tax, substantial advantages in vertical integration after the tax reform may not materialize. The aftermath of the reform might yield a possibly adverse effect on vertical integration. This is because plants could redirect their production endeavours toward manufacturing their most productive outputs rather than expending resources on producing inputs that, according to my dataset, involve incurring elevated marginal costs. To quantitatively evaluate this, I exploit the panel-data structure of my dataset that allows me to identify the vertical integration status across years. The results suggest that, after the reform, vertically integrated plants exhibit higher marginal costs than their non-vertically integrated counterparts. However, they also exhibit higher TFPQ than their non-vertically integrated plants.

Combining the previous results, a conclusion that elicits is that the differential effect comes through scale and input-price effects, and it does not come from more efficient production processes within non-vertically integrated compared to vertically integrated plants. One possible interpretation for these results is related to the initial hypothesis: vertically integrated plants do not take full advantage of the improved input market after VAT, which leads them to face higher input prices and, consequently, reduce their input usage relative to their non-vertically integrated counterparts, after the reform.

In what follows, I will shortly discuss the related literature and describe the tax reform details. Finally, I will explain this paper's empirical methods and discuss the main results.

1.1 Related Literature

My paper follows closely Singh (2019). In this paper, the author evaluates the effects of the transition from sales tax to VAT in India on vertical integration. Her findings suggest that VAT incentivized firms to vertically integrate, which the author attributes to tax evasion motives. In my paper, I extend these findings to analyze the effects of vertical integration increase on prices and productivity. A deeper understanding of the joint effects of vertical integration and productivity allows us to identify if VAT taxes generated an economic misallocation, where firms decide to vertically integrate at the cost of lower productivity and, potentially, higher prices, which erodes consumer welfare.

My paper also contributes to the existing literature on the effects of the Indian Tax reform, whose findings show that this policy change was particularly positive in multiple economic dimensions. For example, Agrawal and Zimmermann (2022) explores the effects of the reform on firms' sales and tax complexity rates, suggesting that, consequently, firms' sales and input purchases surged. Moreover, the reform helped to reduce the complexity of the tax system. Hoseini and Briand (2020) evaluate how the reform impacted firm investment and value-added, suggesting positive effects in both dimensions. On the same line, Aneja et al. (2021) found a positive effect of the reform on capital acquisition and revenue productivity. My paper contributes to this literature by showing that the reform affected TFPQ, marginal costs and markups. Exploring these outcomes is essential as it gives a broader perspective of the reform effects. Notably, it helps to understand if markups change after the policy and if the change is driven by technical improvements or tougher price competition. To the best of my knowledge, although there is evidence relating other relevant Indian reforms, such as the Trade Liberalization episode and efficiency improvements De Loecker et al. (2016), there is no evidence linking the Indian tax reform to marginal cost or total factor productivity changes. Finally, in this paper, I explore the interplay between vertical integration status and the reform effect, which to the best of my knowledge, has not been studied before.

2 Tax Reform

The VAT implementation replacing sales taxes was a significant reform in India. This reform was discussed from the late 90s to 2001 when the VAT was approved. ROY et al. (2010) summarized the main objectives of this reform based on public India's government documentation. Mainly, the reform is expected to avoid the cascading tax burden and simplify the tax system. Due to these improvements, India's government expected two main effects: an efficiency increase to rationalize the tax burden and a price decline to remove the cascading effect. Moreover, the structure of the tax reform intended to generate a revenue-neutral effect, a principle that drove the determining process of new tax rates.

After five years of deliberation, the states agreed to replace the sales tax scheme designed at a state level with a uniform national-level VAT. Still, each state could choose specific features of its scheme, keeping basic rules imposed at the national level, such as three basic rates: For inputs, the base VAT rate was 4%, for gold and silver was 1%; and for the rest of commodities, was 12.5%, (the election of this particular rate aimed for the revenue neutrality). States move away from these taxes heterogeneously, with average VAT tax rates of 5.87% by the first year after the reform (Agrawal and Zimmermann (2022)). Finally, the chosen VAT scheme was destination-based, meaning the rate is based on the consumers' location.

The adoption of the policy was staggered across states. Haryana was the first adopter on April 1, 2003. Most Indian states introduced the VAT in 2005, and there was a small set of late adopters after 2005 (Table A.1 in the appendix A summarizes the adoption schedule). One possible concern that arises when reduced form methods I use throughout this paper are applied is that the staggered adoption responds to endogenous conditions that may be related to the outcomes of interest. However, as noticed by Hoseini and Briand (2020), the staggered adoption obeys mainly exogenous factors such as election schedules and administrative processes. This makes me confident about using the empirical methodology discussed below.

3 Empirical Methods

3.1 Data

The primary data set used in this paper is the Annual Survey of Industries from 2001 to 2010. I decided to use this period as, first, it covers the tax reform period. Second, product codes are uniform throughout this period. Since 2011, India has moved from a 5-digit product classification (ASICC) to a 7-digit product classification (NPCMS). It is hard to map both classifications one-to-one. After a manual checking procedure, I found incoherent product correspondences, which prevented me from using the publicly available correspondence files. Finally, there are no considerable gains in using an extended database as India started to discuss a new tax reform during the 2010s, which may create confounding components in my policy effects identification.

ASI covers all manufacturing plants in India that have ten or more workers. The sampling design in this survey works as follows: from 2001 to 2004, all plants with 100 or more workers, or those located in the six (it was reduced to the five less developed by 2008) less industrially developed, were eligible for a census and were required to fill the survey. A random sample is applied to the remaining plants. After 2004, plants belonging to the strata (State by 4-digit of NIC-04) having less than or equal to 4 units are also considered Census Sector units. All my estimations were corrected using sample weights.

3.2 Potential Vertical Integration Measurement

Measuring directly vertical integration using the ASI database is challenging, as ASI does not record information about the inputs a plant produces and uses in its production process. Hence, I follow Orr (2022) approach to measure *"Potential Vertical Integration"* that exploits the fact that the ASI records information about the goods that each plant produces and the inputs that each plant employs to produce its goods' portfolio.

The *"Potential Vertical Integration"* measure involves two steps. First, I keep the single-product plants and identify the set of inputs each plant purchases to produce for each 5-digit product code. This step aims to measure the inputs required on average to produce a product. Note that I need to use single-product plants only to be sure that the set of inputs I observe corresponds to producing the only product the plant produces. Secondly, I use the multi-product firm database to identify the plants that produce a pair (k, g) of products where at least 10% single-product producers of k also buy input g. That set of plants will be identified as *potentially vertically integrated*.

3.3 Measuring Tax Policy Effects on Vertical Integration

To measure the effects of the tax policy on vertical integration, I appeal to standard reduced-form methods. One first natural approach is to use a standard two-way fixed effects approach as follows:

$$VI_{it} = \alpha_i + \delta_t + \beta \text{Treated}_{it} + \gamma X_{it} + \xi_{it} \tag{1}$$

Where VI_{it} denotes whether a plant i is categorized as vertically integrated at year t using the methodology described in section 3.2, α_i denotes a set of plant fixed effects, δ_t denotes a set of year fixed effects and X_{it} are plant-level control variables such as plant's size and plant's age. As highlighted recently, this regression is inadequate when treatment is staggered (Callaway and Sant'Anna (2021)). To address this, I use a stacked-event design (Cengiz et al. (2019)). The main idea of this technique is to create a "rolling control group" constructing a dataset as follows: First, I create a dataset for each of the cohorts c in which states were treated (2003, 2005, 2006 and 2007). In each dataset, I consider all the plants in the states that transited to VAT in that year as treated. In each dataset, I defined the control group as the firms in states that experienced the VAT transition in later years. As a next step, I create event-time dummies relative to the year of treatment in each separate dataset. My estimating equation becomes:

$$VI_{it} = \alpha_{ci} + \delta_{ct} + \eta_s + \sum_{k=-3}^{k=3} \gamma_k * D^k \times Treated_{ic} + \mu X_{it} + \xi_{it} \quad (2)$$

Where α_{ic} are cohort-plant fixed effects,δ_{tc} are year-cohort fixed effects, D^k is a set of relative event-time dummies, and X_{it} is a matrix of plant-level control variables. The coefficients of interest are the $\gamma_k's$ that indicate the average change in the vertical integration probability of treated plants k years after the state they are located transited to VAT relative to the pre-treatment year.

My empirical design relies upon two main assumptions. First, as in a standard difference-in-difference setting, the absence of differential trends between plants experiencing treatment at different periods is required. Secondly, states' timing decision to adopt VAT should not be correlated with the outcome of interest. Regarding the first assumption, equation 2 allows me to test if there were differential trends between plants before the treatment by checking the negative order $\gamma's$. Regarding the second assumption, I rely on previous research that suggests that Indian states' timing of VAT adoption decisions was based mainly on administrative factors and election schedules.

3.4 Productivity, Marginal Cost and Markups Estimates

To measure the impacts of tax reform on technical and cost efficiency, it is necessary to obtain consistent estimates of Total Factor Productivity (TFP) and marginal costs. I decided to use TFP as my technical efficiency measure following Syverson (2011), who argues that alternative productivity measures based on single-factor usage, such as labour productivity, are sensitive to using excluded inputs and relative input prices. Moreover, labour productivity may reflect changes in factor accumulation instead of actual changes in technical efficiency.

The alternative that I employ in this paper to obtain estimates of marginal costs and TFPQ is to estimate a product-level production function, which allows me to recover an estimate of TFPQ initially. With a consistent estimate of output elasticities and TFP, it is possible to recover a markup estimate. Finally, as ASI provides information on prices at the product level, it is possible to back out estimates of marginal costs using the markup estimates and price data. Multiple methods accommodate estimating, at the plant level, either gross-value production functions (Gandhi et al. (2020)) or value-added production functions (Olley and Pakes (1996); Levinsohn and Petrin (2003)). At the product level, recently, multiple methods have been developed to handle multi-product plants (De Loecker et al. (2016) DLGKP; Blum et al. (2022)). The advantage of these methods is that, provided that product-level price data is available, it is possible to recover both markups and marginal product estimates. Moreover, these methods exploit the detailed information on prices and quantities at the product level to avoid the widely recognized issue of the ”output-price bias” that emerges when one uses revenue-based production functions as price-setting behaviour of firms in non-competitive markets makes it impossible to establish a connection between revenue- and quantity-based production functions (Bond et al. (2021)). In the particular case of DLGKP it is possible to avoid the ”input allocation bias” that emerges when input allocations across products in multi-product firms are not observed. This is the standard case in production databases such as ASI. I follow DLGKP to obtain the production function estimates, marginal cost and TFP. In what follows, I will elaborate on the intuition behind estimating the objects of interest following closely the DLGKP exposition.

3.5 Marginal Costs and Productivity Estimation

3.5.1 Theoretical Problem

Consider a production function for a firm f producing a product j at time t

$$Q_{fjt} = F(L_{fjt}, M_{fjt}, K_{fjt})\Omega_{ft} \tag{3}$$

Where Q_{fjt} is physical output, L_{fjt} and M_{fjt} are labour and materials measures (that usually are assumed to be a variable input that the firm can freely adjust), K_{fjt} is a capital measure that is assumed to be a fixed input that faces adjustment costs. Denote the W^{x}_{fjt} as the variable input v price. Finally, Ω_{ft} is the TFP (In this case, TFPQ, as I estimate a quantity-based production function).

Assuming that plants minimize short-run costs, taking output quantity and input prices as given, the first order condition associated with the lagrangian function $\mathcal{L}$ with respect to materials for any product and time is:

$$\frac{\partial \mathcal{L}_{fjt}}{\partial M_{fjt}} = W^{m}_{fjt} - \lambda_{fjt}\frac{\partial Q_{fjt}}{\partial M_{fjt}} = 0 \tag{4}$$

A crucial assumption here is that costs are separable across products, which allows for product-by-product cost minimization. In equation 4, λ_{fjt} denotes the marginal cost of production at a given output level. One can rearrange equation 4 as follows:

$$\frac{\partial Q_{fjt}}{\partial M_{fjt}}\frac{M_{fjt}}{Q_{fjt}} = \frac{1}{\lambda_{fjt}}\frac{W^{m}_{fjt}M_{fjt}}{Q_{f}jt} \tag{5}$$

The left-hand side in equation 5 represents the output-elasticity with respect to materials, which I denote as θ. Moreover, since markup (μ_{fjt}) is defined as the ratio between prices and marginal costs $\left(\frac{P_{fjt}}{\lambda_{fjt}}\right)$, equation 5 can be rewritten as follows:

$$\mu_{fjt} = \theta_{fjt}\left(\frac{P_{fjt}Q_{fjt}}{W^{m}_{fjt}M_{fjt}}\right) \tag{6}$$

Notice that using ASI, the expression ($P_{fjt}Q_{fjt}$) is observed at the product level. However, $\left(W^{m}_{fjt}M_{fjt}\right)$ is observed only at the plant level. Hence, two challenges remain to estimate consistent expressions for TFPQ and marginal costs: first, estimate θ and second, obtain expressions to map the plant-level input expenditures to a product-level counterpart.

3.5.2 Estimation

To estimate the production function described above, consider first a log version adding a measurement error and unanticipated shocks to output denoted by ϵ_{fjt}

$$q_{jft} = f\left(l_{fjt}, m_{fjt}, k_{fjt}; \beta\right) + \omega_{ft} + \epsilon_{fjt} \tag{7}$$

Lower case variables represent the log of the variables described in equation 3.

Following De Loecker et al. (2016), I estimate the production function exploiting data for multi-product firms that record revenue and quantities sold at the product level. The objective is to obtain consistent estimates of f, which implies the vector β estimation that determines the output elasticities with respect to the physical inputs.

Employing a quantity-based production function mitigates concerns regarding the output-price bias described above. However, potential biases arise when input prices and product-level input allocations are not observable, which is often the case with production databases such as the ASI. This implies that the product-level inputs, l_{fjt}, m_{fjt} and k_{fjt}, are typically unobservable and that the vector $\left(\hat{l}_{ft}, \hat{m}_{ft}, \hat{k}_{ft}\right)$ of firm-level deflated input expenditures is observed instead.

To understand the issue, consider the relationship between product-level input quantities and firm-level for any variable input expenditures denoted by x,

$$x_{fjt} = \rho_{x_{fjt}} + \hat{x_{ft}} - \hat{w}_{x_{fjt}}, \tag{8}$$

where $\rho_{x_{fjt}}$ is the (log) share of input expenditures attributable to product j and $\hat{w}_{x_{fjt}}$ denotes the deviation of the unobserved (log) firm–product-specific variable input price from the (log) industry-wide variable input price index used to deflate input expenditures. The same principle applies to each component of the deflated input expenditures vector. Consequently, (7) can be written as:

$$q_j = f\left(l_{ft}, m_{ft}, k_{ft}\right) + A\left(\rho_{fjt}, l_{ft}, m_{ft}, k_{ft}\right) + B\left(w_{fjt}, \rho_{fjt}, l_{ft}, m_{ft}, k_{ft}\right) + +\omega_{ft} + \epsilon_{fjt}, \tag{9}$$

Where functions A and B are unobservable due to the absence of information about ρ_j and p_j, the vector of (log) input expenditure shares and price deviations by product. The bias emerges from the correlation between the unobservable productivity ω and A and B via the deflated input expenditures.

To address the bias coming from unobserved product-level inputs, De Loecker et al. (2016) rely on single-product firms, for whom $\rho_{x_{fjt}} = 0$. Thus, the bias for these firms comes only from function B. To mitigate this bias, the authors argue that high-quality products produced using high-quality inputs command higher final prices. This theoretical connection establishes a positive relationship between unobserved product-specific input prices and observed product-specific output prices. I employed a second-degree polynomial to approximate the relation between input and output prices. After controlling this, I estimated the production function assuming a trans-log production function using the instrumental variables approach proposed by Wooldridge (2009).

I can retrieve the output elasticity related to the intermediate input θ_m and ρ_m using the results. Following De Loecker and Warzynski (2012), I can recover markups as follows

$$\mu_{fjt} = \theta_m \cdot \frac{P_{fjt} Q_{fjt}}{e^{\rho_{fjt}} W^M_{ft} M_{ft}} \tag{10}$$

Where $e^{\rho_{fjt}}$ is the share of input expenditures attributable to the product j. ρ_{fjt} is retrieved by solving a within-plant system of equations. To execute the estimation algorithm, I use the deflated. intermediate inputs as m, the deflated wage bill as l and k is calculated using the perpetual inventory method.

3.6 Relation Between Tax Reform, Vertical Integration and Efficiency

Due to the nature of the tax reform, its effects may vary depending on whether the plants are vertically integrated. The rationale behind this distinction is as follows: following the reform, economic incentives for vertical integration diminish since plants no longer need to engage in vertical relationships to circumvent the cascading effects inherent to the sales tax scheme. If plants remain vertically integrated after the reform, they will make sub-optimal decisions regarding their goods portfolio production. This is because producing only their primary outputs rather than their primary outputs and a set of inputs implies a higher degree of specialization and, consequently, greater technical efficiency. Therefore, I anticipate that firms that maintain vertical integration after the reform will experience lower productivity gains than those not vertically integrated.

To empirically contrast this hypothesis, I used a differences-in-differences-in-differences (DDD) approach that allows me to compare the effects of the reform. The DDD approach identifies the impact of the tax reform on my interest variables by comparing the differences between vertically integrated plants and non-vertically integrated firms located in states that adopted the VAT and the differences between vertically integrated and non-vertically integrated firms located in states that have not adopted the VAT yet. I used the same "roll-out" control group design as in section 3.3. The estimation equation, in this case, becomes:

$$
\begin{aligned}
o_{itc} = & \alpha_{ci} + \delta_{ct} + \eta_s + VI_{it} \\
& + \sum_{k=-3}^{k=3} \left(\gamma_k \cdot D^k \cdot Treated_{ic} \right) \\
& + \sum_{k=-3}^{k=3} \left(\kappa_k \cdot D^k \cdot Treated_{ic} \cdot VI_{it} \right) \\
& + \sum_{k=-3}^{k=3} \left(\chi_k \cdot D^k \cdot VI_{it} \right) \\
& + \mu X_{it} + \xi_{it}
\end{aligned}
\tag{11}
$$

The triple difference coefficients κ_k and γ_k are my quantities of interest. They represent the tax reform's effect on vertically integrated firms relative to non-vertically integrated firms and not-yet-treated firms (subscript k stands for the relative period that is analyzed) and the general effect of the reform, respectively. In equation 11, o_{it} represents the outcome of interest (either marginal costs or TFPQ). The marginal cost at the firm level is calculated using a sales-weighed average. α_{ic} denotes a set of plant-cohort fixed effects (or product-cohort fixed effects), δ_t denotes a set of year fixed effects, and X_{it} are plant-level control variables such as plant's size, plant's age and the number of products produced within the plant.

Or in a more standard static differences in differences in differences fashion:

$$
\begin{aligned}
o_{itc} = & \alpha_{ci} + \delta_{ct} + \eta_s + \beta_1 VI_{it} + \beta_2 (VI_{it} \times Treated_{ic}) + \beta_3 (VI_{it} \times Post_{it}) \\
& + \beta_4 (Post_{it} \times Treated) + \beta_5 (Post_{ic} \times Treated_{it} \times VI_{it})
\end{aligned}
\tag{12}
$$

In the standard case described in equation 12, the main object of interest is β_5, which describes the tax reform's effect on vertically integrated firms relative to non-vertically integrated firms and not-yet-treated firms

An identification issue emerges when one wants to estimate equation 11 or equation 12: marginal costs and efficiency-related variables are believed to be crucial factors that drive the plant-level vertical integration decision, which causes that the estimates in equation 11 do not reflect solely changes in marginal costs or TFPQ caused by the vertical integration status. Still, they confound the intermediate effect of marginal costs on vertical integration decisions. I use an instrumental variable estimation method nested in the DDD to solve this. To instrument the vertical integration indicator, I propose to use the proportion of vertically integrated plants in all states different from the state in which the plant is located that share the same industry. This variable captures general trends in vertical integration that are idiosyncratic to the industry. The crucial assumption is that the individual plant marginal cost is independent of the vertical integration decisions within the same industry in different states. Formally, the instrumental variable used takes the following form:

$$IV_{ftsm} = \frac{\sum_{k \neq s; r=m} \mathbf{1}[VI_{ftkr} = 1]}{\sum_{k \neq m; r=m} \mathbf{1}} \tag{13}$$

Here, the IV for plant f in time t located in state s producing in the sector m is the ratio of the sum of all vertically integrated plants that have a presence in all states different from s that produce in the sector m, and the sum of all plants that have a presence in all states different from s that produce in the sector m.

4 Results

I present the results following the same order shown in sections 3.3-3.6.

4.1 Summary Statistics

Summary statistics are presented at the product level in table 1 and at the plant level in table 2

Table 1: Summary Statistics (Product Level)

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
Log (Revenue)	400052	16.22	2.72	1.39	32.16	16.26
Log (Quant. Sold)	400052	8.79	3.99	-6.91	25.95	8.73
Log (Quant. Produced)	400052	8.85	3.99	-5.52	25.92	8.79
Multi-Product	400052	.69	.46	0	1	1
Single-Industry	400052	.64	.48	0	1	1
Vertical Integration	400052	.29	.45	0	1	0

Table 2: Summary Statistics (Plant Level)

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
$Log(L)$	214521	9.35	1.64	2.56	16.44	9.12
$Log(K)$	214521	15.94	2.52	-.92	26.74	15.83
$Log(M)$	214521	8.42	3.37	-5.03	25.19	8.26
Number Items by Plant	214521	1.82	1.34	1	20	1
Multi-Product	214521	.47	.5	0	1	0
Number of Plants	93424					

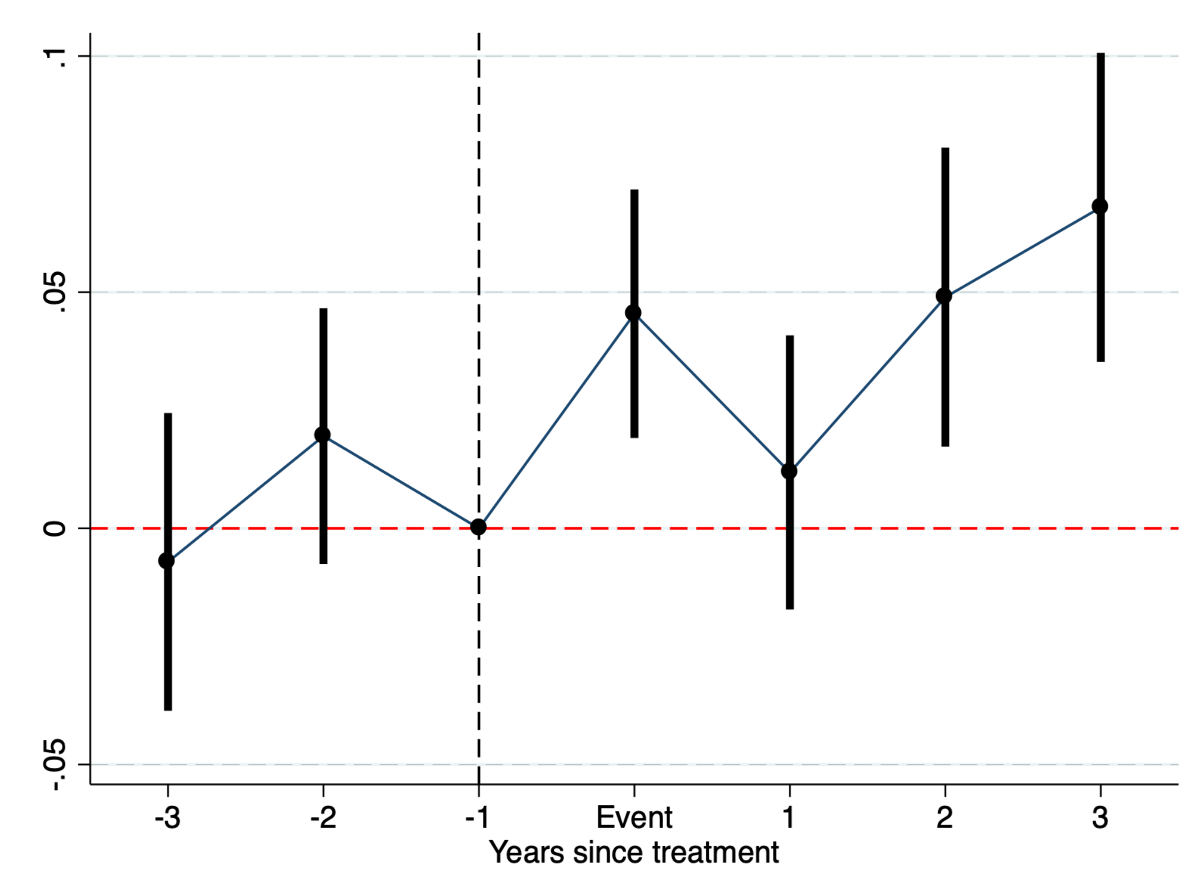
Notes: Labor measured using the deflated wage bill by plant. Capital is measured using the perpetual inventory method. Materials is measured using the deflated expenditure of domestic and imported materials

4.2 Results on Tax Policy Effects on Vertical Integration

Before delving into the investigation of how tax reform interacts with vertical integration, it's essential to examine the general effects of the tax reform on plant decisions regarding vertical integration. My analysis begins by estimating equation 2. In Figure 1, I present the set of coefficients (γ_k) to summarize the dynamic effects of the tax reform on the vertical integration indicator.

The results indicate that the tax reform positively impacted the prevalence of vertical integration in the industries under study. Specifically, on average, five years after the reform's adoption, there is an approximately 7 percentage point increase in the probability of observing a vertically integrated plant relative to the year before the reform is introduced. This finding aligns with the results reported by Singh (2019) for the Indian economy, although Singh used a different measure for vertical integration. My findings also show no evidence suggesting preexisting trends. This lack of preexisting trends suggests that vertical integration decisions are based on the tax scheme change from sales tax to VAT.

Figure 1: The dynamic effect of tax reform on vertical integration



Notes: The estimates presented in this figure correspond to γ_k coefficients in equation 2. This regression limits the analysis exclusively to multi-product firms. I include plant and year fixed effects (saturated by cohort), state fixed effects and control for firm size (approximated by number of workers), firm age and rurality. Standard errors are clustered at the plant level.

My results suggest that contrary to what optimal taxation theory predicts, plants have incentives to engage in vertical integration under a Value Added Tax (VAT) scheme. As Singh (2019) argued, this outcome is likely since vertical integration offers a means to evade taxes under a VAT system. The lack of enforcement requiring final consumers to cross-report their transactions creates an optimal environment for downstream firms to integrate with upstream firms. In this case, vertical integration creates a more complex scenario for tax authorities to monitor business-to-business transactions along the value chain, making it easier for downstream firms to avoid issuing receipts to final consumers and consequently bypass the corresponding tax.

The next step implies identifying whether vertical integration is relevant in explaining how plants' TFPQ and marginal costs evolved after the tax reform. This exercise adds evidence to understand if the VAT consequences on vertical integration affected resource allocations in the Indian economy. In the next section, I present the production function estimates discussed in section 3.5 that serve as inputs to recover both plants' TFPQ and marginal costs.

4.3 Results on Production Function Estimates, Marginal Costs and Productivity

I present the output elasticities, markups, and marginal cost estimates in this subsection using the De Loecker et al. (2016) method. I estimate the production function individually for eleven Indian industries.

Table 3: Average Output Elasticity by Sector

Sector	**Observations**	**Labor**	**Materials**	**Capital**	**Returns to Scale**
Food and Beverages	5852	0.09	0.80	0.07	0.96
Textiles	7148	0.06	0.91	-0.004	0.97
Paper	2849	0.21	0.84	-0.08	0.98
Chemicals	4012	0.05	0.79	0.12	0.96
Rubber and Plastic	3194	0.07	0.83	0.09	0.99
Nonmetallic Minerals	2174	0.21	0.64	0.27	1.12
Basic Metals	4478	0.13	0.84	0.05	1.03
Fabricated Metals	751	0.06	0.85	0.14	1.04
Machinery and Equipment	2549	0.23	0.76	0.02	1.01
Electrical Machinery	2341	0.17	0.73	0.07	0.97
Motor Vehicles	772	0.15	0.79	0.23	1.17

Notes: The estimates presented in this figure correspond to the average output elasticity at the plant level using only the single-product plant sample and assuming a translog production function

Table 3 shows the average output elasticity estimated by sector. My results align with the standard production function estimation literature results regarding output elasticities. My results are also close to what De Loecker et al. (2016) reports using a different Indian database for a different period.

I use the firm-specific output elasticity with respect to materials to recover the product level markups as in equation 10. Given that ASI provides information on product-level prices, it is straightforward to notice that using the markups estimate, I can recover an estimate for product-level short-run marginal costs (λ_{fjt}) as follows:

$$\lambda_{fjt} = \frac{P_{fjt}}{\mu_{fjt}} \tag{14}$$

Moreover, with the production estimates in hand, I can recover the ω_{ft} in equation 7 that accounts for firm-level productivity. Since I follow the estimation technique proposed by Wooldridge (2009) to recover the output elasticities, I cannot recover the productivity estimate for single-product firms. This can be judged as a downside in my estimates related to TFPQ, as selection issues may be prevalent.

Before moving to the estimates of the equation 11, I present some stylized facts of my markups and marginal cost estimates to prove basic relations that suggest theoretical consistency in my estimates. First, I present the median markups per sector in table 4. The median markup is heterogeneous across sectors. Table 5 summarizes the main stylized facts regarding the marginal cost estimates. The estimates evidence a negative correlation between marginal costs and sold quantities, coherent with the returns to scale results calculated previously. I also find a negative correlation between marginal costs and within-firm product market share. This is coherent with previous theoretical findings that suggest that within firms, core products have the lowest marginal costs (Mayer et al. (2014)). Finally, short-run marginal costs are negatively related to productivity, which is theoretically expected.

4.4 Results on Heterogeneity by Vertical Integration Status

In this section, I explore the effects of tax reform on marginal costs and TFPQ. First, I investigate the general impact that tax reform caused on marginal cost and productivity for all the plants in my final sample. Secondly, I explore the potential heterogeneity in the effects depending on the vertical integration status. According to the discussion in section 1, a possible mechanism that could explain this heterogeneity is that if avoiding the sales tax distortions is a crucial factor to plants deciding to integrate vertically, vertically integrated plants may exhibit lower productivity and marginal cost gains after the reform relative to non-vertical integrated plants.

Table 4: Median Markups per Sector

Sector	**Median Markup**
Food and Beverages	1.06
Textiles	1.41
Paper	1.22
Chemicals	1.52
Rubber and Plastic	1.33
Nonmetallic Minerals	1.40
Basic Metals	1.19
Fabricated Metals	1.57
Machinery and Equipment	1.45
Electrical Machinery	1.49
Motor Vehicles	1.61

Notes: This table presents the median markups per sector calculated at the product level after trimming the 3rd and 97th percentile of the markup distribution within the sector.

Table 5: Marginal Costs, Quantities, Shares and Productivity

	(1)	(2)	(3)
	Log λ_{fjt}	Log λ_{fjt}	Log λ_{fjt}
Log q_{fjt}	-0.367***		
	(-291.70)		
Product Share		-1.172***	
		(-103.49)	
ω_{ft}			-0.0727***
			(-27.58)
Constant	0.00792**	0.397***	0.0479***
	(2.86)	(70.14)	(10.25)
N	279648	184848	184848

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents the results of a regression between the log marginal costs residualized by-product code-unit fixed effects and the residualized log of quantities sold, product share and the log of productivity. In column 1, I consider all plants. In columns 2 and 3, I confine my attention only to multi-product plants. I trimmed the 3rd and 97th percentiles in the distribution of each residualized variable analyzed

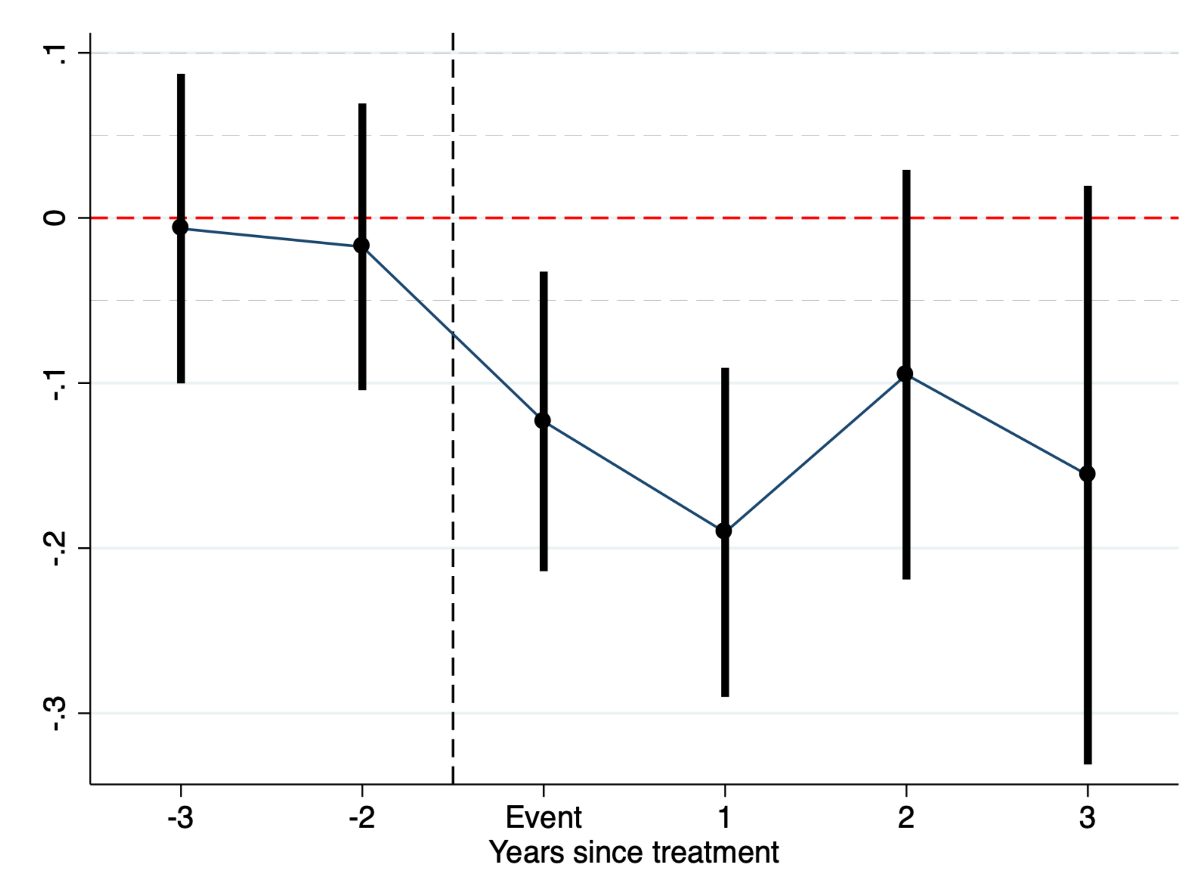
4.4.1 General Tax Reform Effects on Marginal Costs

Before analyzing the heterogeneous effect, I provide evidence of the general impacts of tax reform on marginal costs at the product level and the TFPQ effects at the plant level (see equation 7). Particularly, I estimate the difference-in-differences-in-differences specification in section 3.3 using the product-level marginal cost and the plant-level TFPQ estimates as dependent variables.

Figure 2 presents the results for the dynamic effects of the tax reform on marginal cost using the dynamic DDD setting as in equation 11 and employing instrumental variables to control for the potential endogeneity commented in section 3.6 . Notice that the estimates reported in figure 2 correspond to γ_k's in equation 11, taking as the base period the year before the reform was implemented. The observed effect of the tax reform's positive impact on marginal costs is implied by a generalized reduction in marginal costs during the reform year and the year after the reform (see figure 1). This result suggests that the reform exerted positive efficiency effects for the subset of plants analyzed. Multiple mechanisms explain this result. First, a direct impact of the VAT scheme caused a general tax decline, as shown by (Hoseini and Briand (2020)). Secondly, there are indirect effects caused by VAT, such as lower input-market distortions and a more straightforward tax system. Identifying the exact mechanism through which marginal costs are impacted is beyond the scope of this paper. However, analyzing the effects on TFPQ helps to gain intuition about the mechanisms.

The results on the effects of tax reform on TFPQ are summarized in figure 3. Generally speaking, there is no evidence of tax reform causing a generalized improvement in TFPQ. This suggests that the tax reform is not a technical foster at the plant level. However, my results may hide within plant efficiency dynamics that I cannot capture with my production function estimation technique. The lack of effects on TFPQ suggests that the observed effects on marginal costs come through either scale effects or a mechanical effect implied by lower input prices caused by lower tax rates under the VAT scheme. To empirically contrast what could be the dominant mechanism, I repeat the above exercise using as dependent variables the intermediate inputs usage, capital usage and a plant-specific input price index. Results in the first row of table 8 suggest that the effect on marginal costs is explained mainly by increases in intermediate input usages and a mechanical effect through lower input prices. These can be interpreted as VAT creating positive effects in reducing input-market distortions, instead favouring within-plant TFPQ.

Figure 2: The dynamic effect of tax reform on product-level marginal costs

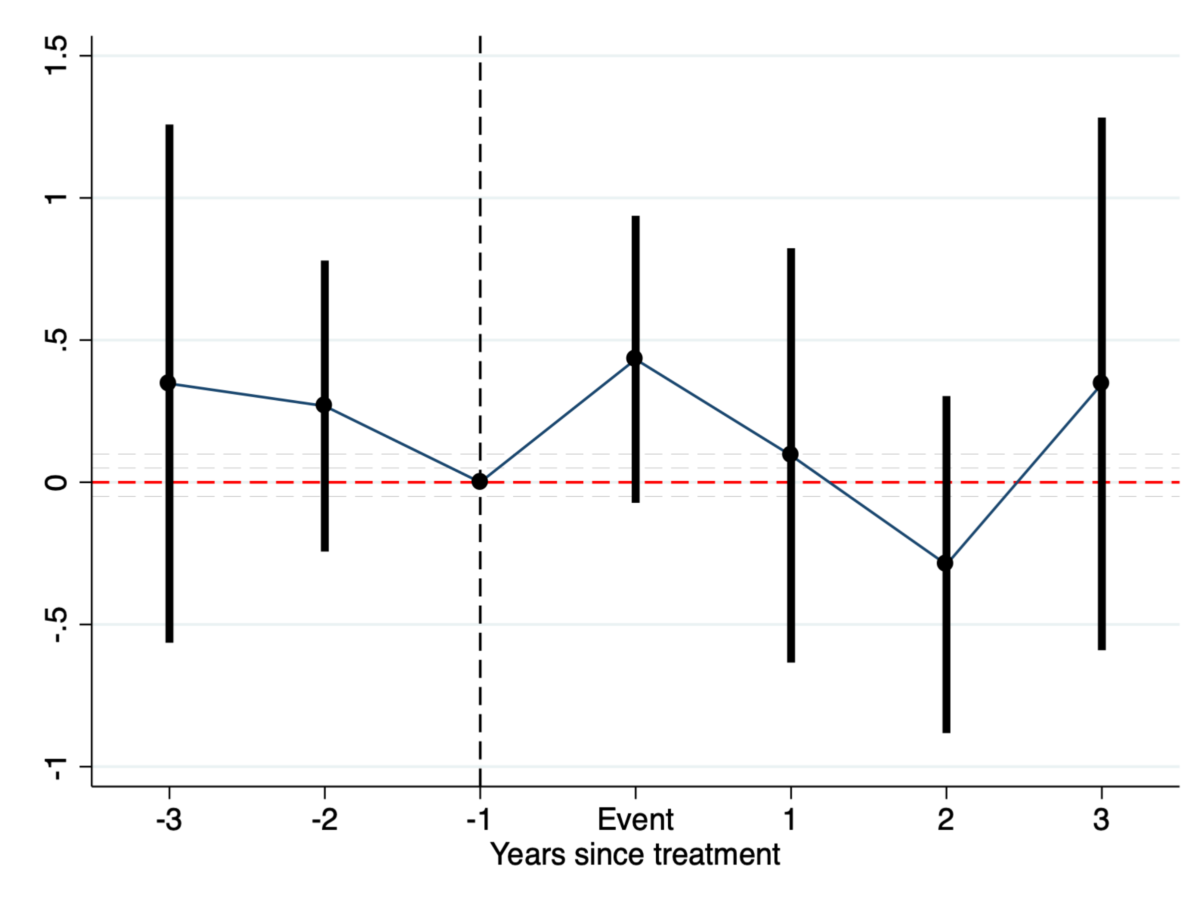


Notes: The estimates presented in this figure correspond to γ_k coefficients in equation 11. I include plant, product and year fixed effects (saturated by cohort), state fixed effects, industry fixed effects and control for firm size (approximated by the number of workers), number of products produced, firm age and rurality. Standard errors are clustered at the product level. The period before the reform is implemented (-1) serves as the base period

4.4.2 Heterogeneous Effects

In this section, I explore the differential effects that the reform has on vertical and non-vertically integrated plants. My rationale indicates that productivity gains should be lower for vertically integrated plants after the tax reform. A key challenge emerges when this analysis is executed: both my dependent productivity measures and the decision of plants to vertically integrate before and after the reform are intrinsically correlated, which can generate biased estimates because of simultaneity. To avoid this, I estimate equation 12 using an instrumental variable approach. As discussed, I use the ratio of vertically integrated plants in the same industry but in different states than the one in

Figure 3: The dynamic effect of tax reform on plant-level TFPQ



Notes: The estimates presented in this figure correspond to γ_k coefficients in equation 11. This regression limits the analysis exclusively to multi-product firms. I include plant and year fixed effects (saturated by cohort), state and industry fixed effects and control for firm size (approximated by number of workers), firm age and rurality. Standard errors are clustered at the product level. The period before the reform is implemented (-1) serves as the base period

which the plant is located to instrument the vertical integration decision. This instrument is based on the idea that industry-specific dynamics lead plants to vertical integration. For example, technical advances within industries lead to more complex products and force plants to produce specialized inputs within the organization. These dynamics have a market nature and should not be related to individual plants' marginal costs. Hence, *a priori* I am allowed to identify the effect of vertical integration on marginal cost. It is important to note that the results in this section should be cautiously interpreted as the instrument vector proposed is relatively weak. I reported the Sanderson-Windmeijer multivariate F test, commonly reported when multiple endogenous variables are involved (see appendix C). Results in table C.1 suggest that the power of the IV to instrument the vertical integration indicator is limited as the F-statistic is below the rule of thumb of 10. The interactions seem to have better instrumentation power because of the mechanical relation through the *post* and *treated* variables.

Table 6 shows the equation 12 estimates using a traditional OLS approach and the IV proposed above (The dynamic version is presented in B. The main result is related to the coefficient β_5. The IV estimator shows that the vertically integrated plants exhibit a marginal cost increase of around 30% relative to non-vertically integrated plants after the reform (row 4). This supports the idea of vertical integration being a source of inefficiencies under a VAT scheme through higher marginal costs. Furthermore, the result associated with the coefficient β_4 evidences a general reduction in marginal costs after the reform of around 11% that applies to all plants, which aligns with the dynamic effects presented above in figure 2.

Table 7 explores the same relations using the TFPQ as the dependent variable. In this case, the results exhibit a positive differential gain in TFPQ for vertically integrated relative to non-vertically integrated plants, after the tax reform. These results suggest that the gains of non-vertically integrated plants relative to vertically integrated plants after the reform may be explained again either by scale effects (vertically integrated plants exhibit a relatively lower input use than their non-vertically integrated counterparts) and/or by input-price effects (vertically integrated plants exhibit relative higher input prices than their non-vertically integrated counterparts). Results in table 7 potentially suffer from selection issues. Here, I confine my attention to multi-product plants, which may self-select into being multi-product conditional on the observables. I am limited by the method I used to obtain the production function estimates, which prevents me from recovering the productivity expression for single-product plants.

Combining the results in table 7 table 6 findings suggest that opposite to the initial hypothesis, vertically integrated plants exhibit TFPQ gains relative to non-vertically integrated plants after the reform. Hence, the results suggest that the positive differential in marginal costs between vertical and non-vertically integrated plants after the reform is explained mainly by scale and input-price differences instead of within-plant non-optimal goods portfolio decisions.

Table 6: Heterogenoeus Effects on product-level (Log) Marginal Costs

	OLS	IV
	(1)	(2)
$Post \times Treatment\,(\beta_4)$	-0.0346*	-0.113***
	(-1.82)	(-3.18)
$VI\,(\beta_1)$	0.0885	2.464
	(1.33)	(1.14)
$Post \times VI\,(\beta_3)$	-0.0419	-0.318***
	(-1.26)	(-2.77)
$Post \times TreatmentXVI\,(\beta_5)$	0.0474	0.252***
	(1.53)	(3.07)
Obs	574890	554595
Year FE	Yes	Yes
Product FE	Yes	Yes
Plant FE	Yes	Yes
State FE	Yes	Yes
Industry FE	Yes	Yes
Plant Age Control	Yes	Yes
Plant Size Control	Yes	Yes
Rurality Control	Yes	Yes
Number of Goods Control	Yes	Yes

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The estimates presented in this table are associated with the coefficients in equation 12

I show evidence related to the argument above in table 8. This table shows the estimates of equation 12 using the intermediate input usage and the plant-specific input-price index as a dependent variable. After the reform, row 4 shows that vertically integrated plants exhibit relatively lower intermediate input usages and higher input prices than non-vertically integrated plants. One possible interpretation for these results is related to the initial hypothesis: vertically integrated plants do not take full advantage of the improved input market after VAT, which leads them to face higher input prices and, consequently, reduce their input usage relative to their non-vertically integrated counterparts, after the reform.

If plants sacrifice marginal costs in their production process when they keep the ver-

Table 7: Heterogenoeus Effects on Plant-level (Log) TFPQ

	OLS (1)	IV (2)
$Post \times Treatment\,(\beta_4)$	-0.0204	-0.183*
	(-0.60)	(-1.89)
$VI\,(\beta_1)$	0.0752***	0.0440
	(2.79)	(0.06)
$Post \times VI\,(\beta_3)$	-0.0130	-0.110
	(-0.44)	(-0.67)
$Post \times Treatment \times VI\,(\beta_5)$	0.0689	0.467**
	(1.50)	(2.12)
Obs	85372	85310
Year FE	Yes	Yes
Plant FE	Yes	Yes
State FE	Yes	Yes
Industry FE	Yes	Yes
Plant Age Control	Yes	Yes
Plant Size Control	Yes	Yes
Rurality Control	Yes	Yes
Number of Goods Control	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The estimates presented in this table are associated with the coefficients in equation 12

tical integration status, an essential question remains: What subset of goods experiences a higher decline in productivity after the reform in plants that are vertically integrated? To evaluate this, I take advantage of my product-level marginal cost estimates, and I use a difference-in-difference-in-differences setting to compare marginal costs' differences between products produced by vertically integrated plants and non-vertically integrated plants located in states that adopted the VAT and the differences between products produced by vertically integrated and non-vertically integrated firms located in states that have not adopted the VAT yet. Since ASI provides information on each plant's ten most important products, I can estimate equation 12 by-product-importance level. I do this to understand better which products suffer the most from the productivity trade-off implied by vertically integrated firms producing inputs instead of focusing production efforts on producing their outputs. I group products into three categories: the three most important

Table 8: Heterogenoeus Effects on plant-level (Log) Input Usage and (Log) Input-Prices

	(Log) Interm. Input Usage (1)	(Log) Input-Price Index (2)
$Post \times Treatment\,(\beta_4)$	0.123***	-0.0941***
	(3.53)	(-2.79)
$VI\,(\beta_1)$	0.0564	0.0533
	(1.32)	(1.34)
$Post \times VI\,(\beta_3)$	-0.0227	-0.0587
	(-0.54)	(-1.50)
$Post \times Treatment \times VI\,(\beta_5)$	-0.213***	0.232***
	(-3.19)	(3.70)
Obs	229019	229019
Year FE	Yes	Yes
Plant FE	Yes	Yes
State FE	Yes	Yes
Industry FE	Yes	Yes
Plant Age Control	Yes	Yes
Plant Size Control	Yes	Yes
Rurality Control	Yes	Yes
Number of Goods Control	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The estimates presented in this table are associated with the coefficients in equation 12 using OLS

products (ordered by sales) constitute the first or ”core products,” the second three most important are in the second category, and the rest in the last category.

To understand the tax reform’s effects on product-level marginal costs and their interaction with vertical integration status, my main approach consists of estimating equation 12 using the instrumental variable described in section 3.6 in order to avoid the simultaneity problem described above. Table 9 summarizes the main results. Column 1 presents the estimates, confining attention exclusively to the three most important products within the plant. For this subset of products, results in the fourth-row evidence that vertically integrated plants exhibit higher marginal costs after the reform than non-vertically integrated plants.

Moreover, the fifth row suggests a generalized decrease in marginal costs for this subset of products after the reform. Interestingly, those patterns are absent when second and 3rd product tiers are analyzed (columns 2 and 3). The results in table 9 suggest that under a VAT scheme, internal input production generates a comparative disadvantage only in the core goods production compared to plants that rely on the external input market. This pattern could be explained by non-vertically integrated plants taking advantage of the more efficient input market that the tax reform intends to generate, allowing them to produce their core products more efficiently. This result is in line with findings in the literature on multi-product firms that show that core products react the most when structural market changes occur in the economy (Dhyne et al. (2017)).

Table 10 shows the differential effects on prices. The results suggest that the reform did not impact the price dimension. Focusing on all plants, table 10 results suggest that lower marginal costs implied by the reform for the core products were not passed through to the final consumers via lower prices, which mechanically generates a markup increase after the VAT reform 11. However, after the reform, vertically integrated firms exhibit lower markups than non-vertically integrated plants, a mechanical result of the absence of price effects. All in all, my results suggest that the tax reform caused an increase in market power approximated by higher markups in the core products of the observed plants and also caused market power increases in non-vertically integrated plants relative to non-vertically integrated plants.

Table 9: Effects of Tax Reform and VI status on (Log) Product Marginal Costs by Product Type

	1st Product Tier	2nd Product Tier	3rd Product Tier
	(1)	(2)	(3)
$VI\,(\beta_1)$	5.135	1.887	-1.474
	(1.25)	(0.30)	(-0.22)
$Post \times VI\,(\beta_3)$	-0.235*	-0.869	-0.308
	(-1.93)	(-0.88)	(-0.37)
$Post \times Treatment \times VI\,(\beta_5)$	0.249***	0.372	0.236
	(2.80)	(0.52)	(0.30)
$Post \times Treatment\,(\beta_4)$	-0.138***	-0.128	-0.172
	(-3.09)	(-0.34)	(-0.61)
Obs	452134	55527	13538
Year FE	Yes	Yes	Yes
Plant FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Unit FE	Yes	Yes	Yes
Plant Age Control	Yes	Yes	Yes
Plant Size Control	Yes	Yes	Yes
Rurality Control	Yes	Yes	Yes
Number of Goods Control	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The estimates presented in this table are associated with the coefficients in equation 12 using the IV

Table 10: Effects of Tax Reform and VI status on (Log) Product Marginal Costs by Product Type

	1st Product Tier	2nd Product Tier	3rd Product Tier
	(1)	(2)	(3)
$VI\,(\beta_1)$	4.596	7.338	6.596
	(1.26)	(0.85)	(0.65)
$Post \times VI\,(\beta_3)$	0.150	-0.609	1.708
	(1.38)	(-0.54)	(1.60)
$Post \times Treatment \times VI\,(\beta_5)$	-0.0155	0.544	-1.347
	(-0.24)	(0.66)	(-1.21)
$Post \times Treatment\,(\beta_4)$	-0.0478	-0.391	0.448
	(-1.29)	(-0.81)	(1.17)
Obs	452134	55527	13538
Year FE	Yes	Yes	Yes
Plant FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Unit FE	Yes	Yes	Yes
Plant Age Control	Yes	Yes	Yes
Plant Size Control	Yes	Yes	Yes
Rurality Control	Yes	Yes	Yes
Number of Goods Control	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The estimates presented in this table are associated with the coefficients in equation 12 using the IV

Table 11: Effects of Tax Reform and VI status on (Log) Product Marginal Costs by Product Type

	1st Product Tier	2nd Product Tier	3rd Product Tier
	(1)	(2)	(3)
$VI\,(\beta_1)$	-2.530	1.291	1.634
	(-0.89)	(0.32)	(0.54)
$Post \times VI\,(\beta_3)$	0.340***	0.404	0.624
	(3.22)	(0.58)	(1.19)
$Post \times Treatment \times VI\,(\beta_5)$	-0.204***	0.0292	-0.115
	(-2.92)	(0.06)	(-0.19)
$Post \times Treatment\,(\beta_4)$	0.0878***	-0.104	0.0523
	(2.65)	(-0.40)	(0.16)
Obs	448871	55094	13298
Year FE	Yes	Yes	Yes
Plant FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Unit FE	Yes	Yes	Yes
Plant Age Control	Yes	Yes	Yes
Plant Size Control	Yes	Yes	Yes
Rurality Control	Yes	Yes	Yes
Number of Goods Control	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The estimates presented in this table are associated with the coefficients in equation 12 using the IV

5 Discussion and Conclusion

In this paper, I evaluate the effects of transitioning from a sales tax to a VAT scheme on marginal costs and productivity. VAT is regularly considered an efficient and productivity-foster tax scheme. However, as I show in this paper, the relationship between this tax policy and economic efficiency depends significantly on the plant-level vertical integration decision. Mainly, I find that the transition from sales tax to VAT caused an increase in the economy's vertical integration prevalence, which is consistent with previous literature findings for the Indian economy (Singh (2019)).

Moreover, my estimates suggest that tax caused a generalized decline in marginal costs but did not affect TFPQ. Both results suggest that plants benefited from the reform through a more efficient input market that implied lower input prices and higher intermediate input use. Secondly, my estimates suggest that vertically integrated plants after the reform showed higher marginal costs than their non-vertically integrated counterparts, and simultaneously, they registered higher relative to non-vertically integrated plants. These results prevent me from arguing in favour of my original hypothesis that vertically integrated plants make sub-optimal decisions regarding their goods portfolio production under VAT. However, here, it is important to notice two things. First, I cannot observe product-level productivity estimates, which prevents me from identifying within plant TFPQ dynamics. Secondly, in the plant-level TFPQ estimates, I can observe only TFPQ for multi-product plants, which may cause selection issues. Hence, my results should be taken with caution.

The results in this paper suggest that VAT has unintended effects on vertical integration, which incentivizes plants to adopt this kind of production scheme. Moreover, this reform in India does not affect internal plant productivity. However, evidence suggests that plants took advantage of the lower distortions in the input market generated by the VAT scheme. Notably, multiple dynamics are involved in the vertical integration decision that is not discussed or empirically considered in this paper, which is out of the scope of this document. However, further research could focus on understanding the decision of vertical integration as a response to policies that may affect plant-level production dimensions through a structural model. This paper is intended to show empirical facts that may motivate the pursuit of this kind of research.

References

D. R. Agrawal and L. V. Zimmermann. The Effects of Adopting a Value Added Tax on Firms. *SSRN Electronic Journal*, 2022. ISSN 1556-5068. doi: 10.2139/ssrn.4116972. URL https://www.ssrn.com/abstract=4116972.

A. Aneja, N. Kulkarni, and S. K. Ritadhi. Consumption Tax Reform and the Real Economy: Evidence From India's Adoption of a Value-Added Tax. *Journal of Empirical Legal Studies*, 18(3):569–602, Sept. 2021. ISSN 1740-1453. doi: 10.1111/jels.12296. URL https://doi.org/10.1111/jels.12296. Publisher: John Wiley & Sons, Ltd.

B. S. Blum, S. Claro, I. Horstmann, and D. A. Rivers. The ABCs of Firm Heterogeneity when Firms Sort into Markets: The Case of Exporters. 2022.

S. Bond, A. Hashemi, G. Kaplan, and P. Zoch. Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics*, 121:1–14, July 2021. ISSN 0304-3932. doi: 10.1016/j.jmoneco.2021.05.004. URL https://www.sciencedirect.com/science/article/pii/S0304393221000544.

B. Callaway and P. H. Sant'Anna. Difference-in-Differences with multiple time periods. *Themed Issue: Treatment Effect 1*, 225(2):200–230, Dec. 2021. ISSN 0304-4076. doi: 10.1016/j.jeconom.2020.12.001. URL https://www.sciencedirect.com/science/article/pii/S0304407620303948.

D. Cengiz, A. Dube, A. Lindner, and B. Zipperer. The Effect of Minimum Wages on Low-Wage Jobs*. *The Quarterly Journal of Economics*, 134(3):1405–1454, Aug. 2019. ISSN 0033-5533. doi: 10.1093/qje/qjz014. URL https://doi.org/10.1093/qje/qjz014.

J. De Loecker and F. Warzynski. Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–71, 2012. doi: 10.1257/aer.102.6.2437. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.6.2437.

J. De Loecker, P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik. Prices, Markups, and Trade Reform. *Econometrica*, 84(2):445–510, Mar. 2016. ISSN 0012-9682. doi: 10.3982/ECTA11042. URL https://doi.org/10.3982/ECTA11042. Publisher: John Wiley & Sons, Ltd.

E. Dhyne, A. Petrin, V. Smeets, and F. Warzynski. Multi Product Firms, Import Competition, and the Evolution of Firm-product Technical Efficiencies. Technical Report w23637, National Bureau of Economic Research, Cambridge, MA, July 2017. URL http://www.nber.org/papers/w23637.pdf.

A. Gandhi, S. Navarro, and D. A. Rivers. On the Identification of Gross Output Production Functions. *Journal of Political Economy*, 128(8):2973–3016, Aug. 2020. ISSN 0022-3808. doi: 10.1086/707736. URL https://doi.org/10.1086/707736. Publisher: The University of Chicago Press.

J. Gardner. Two-stage differences in differences, July 2022. URL http://arxiv.org/abs/2207.05943. arXiv:2207.05943 [econ].

M. Hoseini and O. Briand. Production efficiency and self-enforcement in value-added tax: Evidence from state-level reform in India. *Journal of Development Economics*, 144:102462, May 2020. ISSN 0304-3878. doi: 10.1016/j.jdeveco.2020.102462. URL https://www.sciencedirect.com/science/article/pii/S0304387820300377.

W. Kopczuk and J. Slemrod. Putting Firms into Optimal Tax Theory. *American Economic Review*, 96(2):130–134, 2006. doi: 10.1257/000282806777212585. URL https://www.aeaweb.org/articles?id=10.1257/000282806777212585.

J. Levinsohn and A. Petrin. Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341, Apr. 2003. ISSN 0034-6527. doi: 10.1111/1467-937X.00246. URL https://doi.org/10.1111/1467-937X.00246.

T. Mayer, M. J. Melitz, and G. I. P. Ottaviano. Market Size, Competition, and the Product Mix of Exporters. *American Economic Review*, 104(2):495–536, 2014. doi: 10.1257/aer.104.2.495. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.2.495.

G. S. Olley and A. Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–1297, 1996. ISSN 00129682, 14680262. doi: 10.2307/2171831. URL http://www.jstor.org/stable/2171831. Publisher: [Wiley, Econometric Society].

S. Orr. Within-Firm Productivity Dispersion: Estimates and Implications. *Journal of Political Economy*, 130(11):2771–2828, Nov. 2022. ISSN 0022-3808. doi: 10.1086/720465. URL https://doi.org/10.1086/720465. Publisher: The University of Chicago Press.

P. ROY, A. RAYCHAUDHUR, and S. K. SINHA. Is Value Added Tax (VAT) Reform in India Poverty Improving? An Analysis Of Data From Five Major States. *Indian Economic Review*, 45(1):131–158, 2010. ISSN 00194670. URL http://www.jstor.org/stable/29793957. Publisher: Department of Economics, Delhi School of Economics, University of Delhi.

D. Singh. Merging to Dodge Taxes? Unexpected Consequences of VAT Adoption in India. *SSRN Electronic Journal*, 2019. ISSN 1556-5068. doi: 10.2139/ssrn.3505534. URL https://www.ssrn.com/abstract=3505534.

C. Syverson. What Determines Productivity? *Journal of Economic Literature*, 49(2): 326–65, 2011. doi: 10.1257/jel.49.2.326. URL https://www.aeaweb.org/articles?id=10.1257/jel.49.2.326.

J. M. Wooldridge. On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114, Sept. 2009. ISSN 0165-1765. doi: 10.1016/j.econlet.2009.04.026. URL https://www.sciencedirect.com/science/article/pii/S0165176509001487.

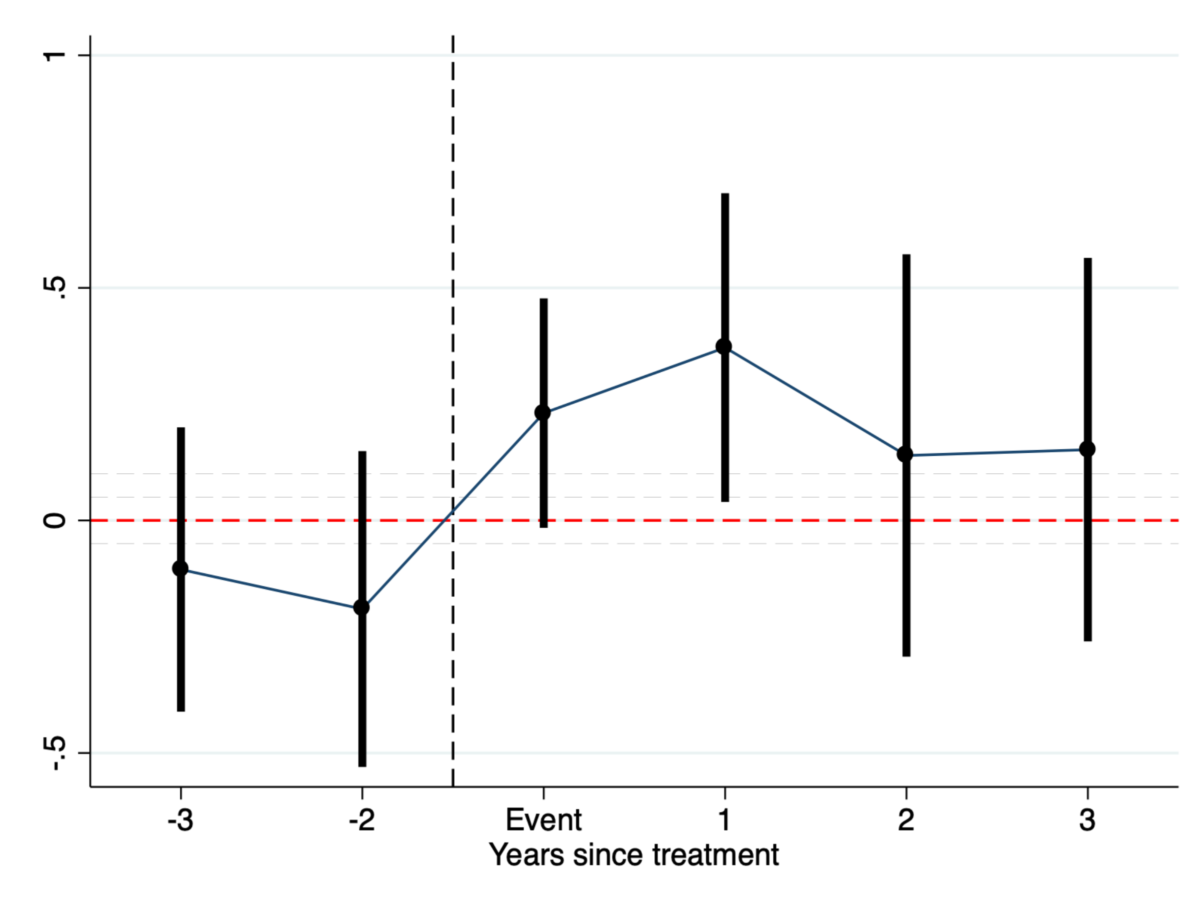
A Adoption Timing

Table A.1: VAT Adoption Years by State

State	VAT Adoption Year
Andhra Pradesh	2005
Arunachal Pradesh	2005
Assam	2005
Bihar	2005
Chandigarh	2005
Chhattisgarh	2006
Dadra and Nagar Haveli	2005
Daman and Diu	2005
Delhi	2005
Goa	2005
Gujarat	2006
Haryana	2003
Himachal Pradesh	2005
Jammu and Kashmir	2005
Jharkhand	2006
Karnataka	2005
Kerala	2005
Madhya Pradesh	2006
Maharashtra	2005
Manipur	2005
Meghalaya	2005
Mizoram	2005
Nagaland	2005
Orissa	2005
Puducherry	2007
Punjab	2005
Rajasthan	2006
Sikkim	2005
Tamil Nadu	2007
Telangana	2005
Tripura	2005
Uttar Pradesh	2008
Uttarakhand	2005
West Bengal	2005

B Dynamic Heterogeneous Effects on Marginal Costs

Figure B.1: Dynamic effect of tax reform on vertically integrated vs non-vertically integrated plants' product-level marginal costs



Notes: The estimates presented in this figure correspond to κ_k coefficients in equation 11. This regression limits the analysis exclusively to multi-product firms. I include plant and year fixed effects (saturated by cohort), state fixed effects and control for firm size (approximated by number of workers), firm age and rurality. Standard errors are clustered at the product level. The period before the reform is implemented (-1) serves as the base period

C Weak IV-Test

Table C.1: Sanderson-Windmeijer Test

Variable	Sanderson-Windmeijer multivariate F test
VI	8.21
$VI \times Treated$	61.95
$VI \times Post$	68.08
$VI \times Post \times Treated$	96.29