Taxation and Supplier Networks: Evidence from India*

Lucie Gadenne
Tushar K. Nandi
Roland Rathelot

May 2020

Abstract

Do tax systems distort firm-to-firm trade? This paper considers the effect of tax policy on supply chains in a large developing economy, the state of West Bengal in India. Using administrative panel data on firms, including transaction data for 4.8 million supplier-client pairs, we first document substantial segmentation of supply chains between firms paying Value-Added Taxes (VAT) and non-VAT-paying firms. We then develop a model of firms' sourcing and tax decisions within supply chains to understand the mechanisms through which tax policy interacts with supply networks. The model predicts partial segmentation in equilibrium because of both supply-chain distortions (taxes affect how much firms trade with each other) and strategic complementarities in firms' decision to pay VAT. Finally, we test the model's predictions using variations over time within firm and within supplier-client pairs. We find that the tax system distorts firms' sourcing decisions, and evidence of strategic complementarities in firms' tax choices within supplier networks. A hypothetical reform exempting all firm-to-firm transactions from the VAT would lead to growth of small- and medium-sized firms at the cost of a small decrease in tax revenues.

JEL: O23, H25, L14.

*Gadenne: University of Warwick, Institute for Fiscal Studies and CEPR. Nandi: CTRPFP, Center for Studies in Social Sciences, Calcutta. Rathelot: University of Warwick and CEPR. We would like to thank Pol Antràs, Andrew Bernard, Michael Best, Luis Candelaria, Michael Devereux, Swati Dhingra, Irem Guceri, Jyotsna Jalan, Sugata Marjit, Isabelle Méjean, Dennis Novy, Aureo de Paula, Ben Olken, Carlo Perroni, Imran Rasul, Debraj Ray, Camilla Roncoroni, Juan Carlos Suárez Serrato, Chris Woodruff, and numerous seminar participants for helpful comments. We are particularly thankful to the Directorate of Commercial Taxes of West Bengal, India for the permission to use their data.We gratefully acknowledge financial support from CAGE, the ESRC (grant reference ES/M010147/1), and UKAID through IFS's TAXDEV and the IGC. All errors are our own.

1 Introduction

Intra-national trade costs are large in developing countries; gains from better integration of supply chains within these countries are potentially substantial (WTO, 2004). Whilst the role of geography as a determinant of such costs is well established (see for example Atkin and Donaldson, 2015), tax policy can also play a role because most tax systems alter the incentives agents have to trade with each other (Fajgelbaum et al., 2019). This is true in particular of the Value-Added Tax (VAT), one of the largest source of revenues in developing countries (Bird and Gendron, 2007), as only VAT-paying firms can deduct VAT paid on their purchases from their tax liabilities. The tax system thus potentially distorts firm-to-firm trade in markets where VAT and non-VAT-paying firms co-exist – a pervasive feature of developing countries – leading to segmentation of supplier networks between firms that pay VAT and firms that don't.

This paper considers two related questions. First, how does the tax system affect firms' sourcing decisions, and therefore firm-to-firm trade, in a developing economy? Second, how do supplier networks affect firms' tax decisions, in particular the decision of whether or not to pay VAT? We build a model of firms' sourcing and tax decisions to understand the mechanisms through which tax policy interacts with supplier networks. We then test the predictions of our model using a rich panel dataset including firm-to-firm transactions for the state of West Bengal in India. We find both that the tax system distorts trade and that there are strategic complementarities in firms' decision to pay VAT within supplier networks.

Our first contribution is to document the segmentation of supplier networks between firms that pay VAT and firms that don't in a large developing economy. A key constraint faced by the literature on intra-national trade is that domestic trade flows are hard to characterize because firm-to-firm trade is rarely observed.[1] Similarly, administrative tax data typically does not contain information on firms that do not pay VAT. We overcome both these observational challenges by using administrative tax data on the universe of the 180,000 firms paying taxes in West Bengal, India, for the period 2010-2016, and two particularities of our context. First, we observe both VAT-paying and non-VAT-paying firms because firms below a size threshold can opt for a non-VAT 'simplified' tax scheme, under which they pay a small tax on their sales but cannot deduct VAT paid on their purchases from their tax liabilities. The existence of a size threshold below which firms can opt-out of paying

[1] Exceptions include Alfaro-Urea et al. (2018) who use administrative data on firm-to-firm transactions for Costa Rica to consider the effect of joining multinational supply chains on firm productivity, and Almunia et al. (2019) who use VAT declarations on firm to firm sales to study under-reporting in Uganda.

VAT is a nearly ubiquitous characteristic of tax systems around the globe (see Keen and Mintz, 2004, for a review). In West Bengal two-thirds of firms fall below this threshold, and amongst those a majority choose to pay VAT, so we observe VAT- and non-VAT-paying firms over a wide range of the firm size distribution. Second, we observe trade between 4.8 million annual client-supplier pairs, because VAT-paying firms report transactions with other tax-registered firms. This allows us to map supplier networks by matching clients' and suppliers' tax identifiers.

We find that VAT-paying firms are substantially more likely than non-VAT-paying firms to trade with other VAT-paying firms. The correlation between firms' tax scheme and how much they buy from, or sell to, VAT-paying firms is large and robust to controlling for firm size and detailed location and industry characteristics. VAT-paying firms on average sell 14 percentage points more to VAT clients, and buy 8 percentage points more from VAT suppliers, all else equal, than non-VAT-paying firms, in line with the idea that the VAT system leads to partial segmentation of supplier networks by tax scheme.

Our second contribution lies in a model that clarifies the mechanisms leading to supply chain segmentation. The model is a bi-partite application of models of supplier networks (see for example Tintelnot et al., 2018), augmented to include a tax scheme decision: firms choose whether or not to pay VAT. Our set-up is one in which firms at two stages in supply chains simultaneously make tax and sourcing decisions under monopolistic competition. Our main result is that under a VAT system there is partial segmentation of supply chains by tax scheme in equilibrium, for two reasons. First, the VAT's incentive structure leads to *supply-chain distortions*: all else equal a VAT-paying firm buys a higher share of its inputs from VAT-paying suppliers than a non-VAT-paying one does. This mechanism implies that the VAT decreases trade between firms in different tax schemes, even in a world where firms' tax schemes are exogenously given. Endogenising firms' tax choices introduces a second mechanism, *strategic complementarities in tax decisions*: firms are more likely to choose to pay VAT the more they trade with VAT-paying suppliers and clients.

Finally, our third contribution is to provide empirical evidence on the mechanisms outlined by our model and use our parameter estimates to quantify the size of the distortions created by the tax system. We first estimate the causal effect of taxes on firm-to-firm trade (supply chain distortions). We identify this effect by leveraging our transaction level data and within supplier-client pairs variations in trade over time, allowing for unobserved supplier productivity shocks and controlling for firm size. We find that firms buy 12% more on average from VAT-paying suppliers when they themselves choose to pay VAT. Our estimates enable us to identify the elasticity of substitution in production from transaction data; we

find values in the 4.5-6.5 range, in line with estimates obtained from firm or industry level data in the literature (see for example Bas et al., 2017; Broda et al., 2017).

To identify strategic complementarities in firms' choice of tax scheme we use within-firm changes over time in the share of sales (inputs) that firms can sell to VAT-paying clients (purchase from VAT-paying suppliers) generated by the entry and exit of their VAT-paying trading partners. Our estimates imply that forcing all of a firm's trading partners to pay the VAT would increase that firm's propensity to pay the VAT by up to 14 percentage points compared to a situation where none of its trading partners pay VAT. All our results are robust to controlling for location- and product-specific shocks that could explain both firm entry and changes in tax scheme, and to controlling for firm size.

The magnitude of the distortions implied by our estimates are economically meaningful. Using our model and parameter estimates, we simulate the effect of a hypothetical reform that would exempt *all* firm-to-firm transactions from the VAT, thus removing both supply chain distortions and most sources of complementarities. Our findings show that the design of tax policy can constrain firm size: the reform would lead to growth of small- and medium-sized firms, those not paying VAT in particular would grow by 7%. This growth effect would come at a small tax revenue cost due to a small number of firms leaving the VAT scheme. We also find that this reform would decrease segmentation 'upstream' in supply chains by roughly 50%. This suggests our two mechanisms explain a substantial share of the supply chain segmentation we observe.

Our results have several main implications. First, we find that VAT systems do distort firm-to-firm trade, suggesting that tax systems indeed contribute to the low levels of market integration observed in developing countries. Second, the same mechanisms that distort firm-to-firm trade also constrain the growth of small- and medium-sized firms that do not pay VAT. This result highlights how the design of tax policy can constrain firm growth, in line with the idea that government regulations contribute to keeping firms small in the developing world (see Hsieh and Olken, 2014, for a discussion of this literature). Third, our results enable us to revisit the trade-offs associated with differents types of consumption taxes. The literature on the VAT typically argues this tax is better suited than Retail Sales Taxes to contexts in which compliance is low (Pomeranz, 2015; Naritomi, 2018). Our findings suggest that this compliance advantage must be weighted against the efficiency cost due to supply chain distortions, we provide estimates of the magnitude of these distortions in our context.[2] Finally, we find that firms' decisions to pay VAT is influenced not just by

[2]See also Emran and Stiglitz (2005), Keen (2008) for theoretical work regarding the optimal tax policy mix in the presence of informal sectors.

the parameters of the tax system, but also by their position in supplier networks. This implies that tax interventions that incentivize some firms to pay VAT have spillover effects on these firms trading partners, as some of them will also start paying VAT. Studies that only measure the direct revenue effects of these interventions therefore risk under-estimating their benefits. We return to these implications when discussing what our results imply for tax policy.

A large literature has considered how taxes affect international trade flows (see Goldberg and Pavcnik, 2016, for a review on the role of tariffs), but there is limited evidence regarding how taxes determine intra-national trade – one exception is Fajgelbaum et al. (2019) who show that state taxes affect the spatial allocation of economic activity in the US.[3] To the best of our knowledge this paper is the first to show how the tax system shapes intra-national firm-to-firm trade. Our results more generally contribute to the recent literature that considers the role of intra-national trade costs (Agnosteva et al., 2014; Atkin and Donaldson, 2015; Cosar and Fajgelbaum, 2016; Fajgelbaum and Redding, 2018) by showing that the tax system affects these costs and therefore firms' sourcing decisions. This paper also speaks to the large literature on firms in developing countries that studies the role of market frictions in the formation of client-supplier relationships, and finds that enforcement and information constraints loom large in this context (McMillan and Woodruff, 1999; Banerjee and Duflo, 2000; Allen, 2014; Macchiavello and Morjaria, 2015). Unlike much of this literature, which studies relationships between multinational companies and their suppliers in developing countries, we focus on within-country trading relationships. We find that whilst frictions may also be substantial in our context they do not lead to a low willingness of firms to substitute across suppliers in response to changes in relative input costs.

Our results also contribute to the literature on public finance in developing countries that asks how the particular context of these countries changes tax policy trade-offs (Boadway and Sato, 2009; Gordon and Li, 2009; Best et al., 2015; Brockmeyer and Hernandez, 2016; Bachas and Soto, 2017; Carrillo et al., 2017; Jensen, 2019; Gadenne, forthcoming). We focus on how the VAT affects supply chains when VAT-paying and non-VAT-paying firms co-exist within markets, a pervasive characteristic of the developing country context. The idea of strategic complementarities in tax choices under a VAT was first introduced by De Paula and Scheinkman (2010); we build on their work by incorporating tax decisions in a supplier network model and providing causal evidence of the existence of these complementarities.[4]

[3] See also Benzarti et al. (2018) for evidence regarding the role of domestic taxes in international trade.

[4] Evidence consistent with the existence of complementarities in tax choices is also found in Almunia et al. (2017) who show that higher input use increases the probability that firms choose to voluntarily register to pay VAT in the UK. In an on-going project Gerard et al. (2020) study a similar question in the context of the

Finally, this paper also speaks to the growing literature on supplier networks that leverages new datasets on firm-to-firm transactions in developed economies to characterize the determinants of supplier networks and the propagation of shocks within these networks.[5] Using data for India we contribute to this literature in two ways. First, we consider theoretically, and provide empirical evidence on, the role of the tax system in shaping supplier networks. Second, we show how supplier networks affect public policy via strategic complementarities in firms' tax choices (see also Liu, 2019, for an analysis of how production networks affect public policy, in his case industrial policy).

The paper is organized as follows. Section 2 describes our context of study and data and provides descriptive evidence on the segmentation of supplier networks between VAT-paying and non-VAT-paying firms. Section 3 develops a model of firms' sourcing and tax scheme decisions and Section 4 discusses the empirical strategy used to provide causal evidence on the model's mechanisms. Section 5 presents our results and Section 6 discusses the magnitudes of our estimates and their implications for policy.

2 Context and data

2.1 Institutional background

Our context of study is West Bengal, a large state in India with 90 million inhabitants and a GDP per capita of 6000 ppp USD in 2018. Our period of study is 2010-2016. The main source of tax revenues at the state level is the value-added-tax (VAT). All firms with a turnover of more than 500,000 INR (7,100 USD) are required to remit taxes to the state. Amongst those, firms with a turnover of less than 5 million INR (70% of tax-registered firms) can opt to remit taxes under a 'simplified' tax scheme under which they pay a 0.25% tax on their total sales. Importantly for the purpose of this paper, firms in the simplified scheme cannot deduct taxes paid by their suppliers from their tax liabilities. All other firms must remit the VAT, and can deduct VAT paid on their inputs by their suppliers from their tax liabilities (see Ghosh and Nandi, 2017, for more details on West Bengal's tax system).[6]

state of Sao Paulo, Brazil.

[5] See Atalay et al. (2011); Bernard et al. (2015); Acemoglu et al. (2016); Carvalho et al. (2016); Bernard and Moxnes (2018); Tintelnot et al. (2018); Di Giovanni et al. (2018); Boehm and Oberfield (2018); Bernard et al. (2019); Dhyne et al. (2019).

[6] In addition most firms are liable to pay Corporate Income Tax (CIT) to the federal government. The CIT liability is not affected by the VAT liability, does not change at the 5 million INR threshold, and the state and federal governments did not share information about taxpayers during our period of study.

Firms face different VAT rates depending on the products they sell: 75% of them sell products belonging to the 'reduced' tax schedule and taxed at 4%, 21% sell products in the 'main' tax schedule taxed at 12.5%, the remainder of firms face 'super-reduced' rates of 0% or 1%. In fiscal year 2014 the VAT rates of the main and reduced schedules increased by 1 percentage point. India's VAT system changed substantially with the introduction of the General Sales Tax reform in 2017, after our period of study. We discuss what our results imply for the potential impact of this wide-ranging reform below.

2.2 Data

2.2.1 Firm data

We use administrative data on firm-level tax returns and tax registration information obtained from the West Bengal Directorate for Commercial Taxes for the fiscal years 2010-2011 to 2015-2016. This dataset contains the annual tax returns of the nearly 180,000 firms paying taxes to the state over the period, whether in the VAT or the simplified scheme.[7] Firms paying taxes under the VAT scheme report their total sales, which we use to proxy for turnover and firm size, total intermediate input purchases (purchases from all suppliers, including those that are not tax-registered), and VAT paid on these inputs, if any. The latter gives rise to an 'input tax credit' which is deducted from the total taxes due on sales. Firms paying taxes under the simplified scheme report their total sales and total intermediate input purchases. In addition to the variables used to compute their tax liabilities firm must report the main product they sell, we use this information to allocate firms to one of 170 product categories and a VAT tax schedule.[8] We obtain information on firms' location from their postcode in the tax registration data. Our sample contains 818,865 observations at the firm-year level for 178,011 firms over 6 years .

2.2.2 Trade data

Firms in the VAT scheme are required to report to the tax authorities all transactions with other firms registered to pay taxes in West Bengal, regardless of whether the trading partner is in the VAT or the simplified scheme.[9] They report the annual transaction amount

[7] We exclude from our sample the few firms that are registered to pay taxes, file taxes but report a null tax liability (no reported sales).

[8] 70% of firms report selling only one product. When a firm reports several products we keep the product that represents the largest share of its sales.

[9] A firm does not have to report a trading partner if its annual trade with this partner is less than 50,000 INR (710 USD).

as well as the tax identification number of the client or supplier involved in the transaction in the 'Annexure B' part of their tax returns. Firms in the simplified scheme do not report transactions to the tax authorities, so we do not observe trade between firms in the simplified scheme.

Transactions between VAT-paying firms must be reported by both parties in the transaction. These two parties have no incentive to collude (a transaction increases the tax liability of the supplier, but decreases the tax liability of the client) and the tax authorities systematically cross-check amounts reported by the two parties involved. Transactions between VAT-paying firms and firms in the simplified scheme however are only reported by VAT-paying firms, so they cannot be cross-checked against third-party information. VAT-paying firms have an incentive to report purchases from non-VAT-paying suppliers truthfully: these do not affect their tax liabilities, but all types of mis-reporting lead to fines if detected through a tax audit regardless of their impact on tax liabilities. Firms can similarly expect to be penalized if they mis-report sales to non-VAT-paying clients, but these sales can potentially increase their tax liabilities. If firms only report sales that the tax authorities have third-party reported information on, under-reporting of sales to clients in the simplified scheme is a potential concern.

Several pieces of evidence suggest that in practice under-reporting by VAT-paying firms of sales to firms in the simplified scheme is unlikely to be a major concern; this evidence is presented in Appendix B.1. Most importantly, we find that most firms have no incentive to under-report sales to clients in the simplified scheme because they report (and pay taxes on) total sales that are much larger than their total sales to VAT-paying clients: third-party reported sales represent only 30% of total reported sales on average. Reporting sales to clients in the simplified scheme truthfully will thus not increase the tax liability of the average firm. This is also true amongst the sample of small firms that are expected to evade taxes more – though note that our analysis below uses mostly information on transactions reported by large firms, so higher under-reporting by smaller firms would not bias our results.[10] We discuss in what follows robustness checks excluding the 10% of firms that mostly pay taxes on third-party reported sales, and whose total tax liabilities could therefore increase if they reported one more non-VAT paying client.

[10] Appendix B.1 presents several additional pieces of evidence indicating that asymmetric under-reporting is unlikely to be a concern. First, we find that firms are more likely to report sales to firms in the simplified scheme than purchases from these firms, contrary to what a simple model of evasion would suggest and in line with the evidence that firms in the simplified scheme are more likely to be downstream - see Table B.1. Second, reporting a client in the simplified scheme is positively correlated with the share of a firm's sales that are third-party reported, contrary to what an evasion model would suggest. See Table B.2, Figures B.3 and B.4 and the discussion in Appendix B.1

Our data contains information for 4.8 million annual supplier-client pairs. The fact that we cannot observe transactions between firms in the simplified tax scheme limits the set of theoretical predictions we can take to the data; we concentrate in what follows on deriving and estimating theoretical predictions that pertain to the relationship between a firm's choice of tax scheme (VAT or simplified) and its transactions with VAT-paying firms, which are documented for all firms. Combining the trade data with the firm data allows us to observe, for each firm in each year, its VAT-paying clients and suppliers, the share of its sales that are purchased by VAT-paying clients and the share of its intermediate inputs sourced from VAT-paying suppliers.

2.3 Descriptive statistics

Table 1 presents the key characteristics of firms in our data. The first column includes all firms in the simplified scheme, the second column all firms in the VAT scheme but with a turnover under 5 million INR (and therefore eligible to choose the simplified scheme) and the last all remaining VAT firms. Less than one-third are in the capital region Kolkata, though this share increases amongst larger firms. Appendix Figure B.1 plots the location of the firms in our data on a map; there are firms paying taxes under both tax schemes in all regions in the state. Looking at firm size (turnover) we see that most firms (70%) have a turnover of less than 5 million and are therefore eligible to choose between the VAT and simplified schemes. Among those the majority (85%) choose to pay taxes under the VAT scheme. The detailed distribution of firm size by tax scheme, presented in Appendix Figure B.2, however shows a substantial amount of bunching below the 5 million threshold for firms in the simplified scheme, suggesting some firms have a high preference for this scheme that leads them to produce less (or report less) sales to avoid paying VAT.

The remaining lines of Table 1 show that firms in the simplified scheme are substantially less likely than similar-sized firms in the VAT scheme to trade with VAT-paying firms. They sell a much smaller share of their sales to VAT-paying clients (1% vs 28%) and buy a smaller share of their inputs from VAT-paying suppliers (48% vs 59%). Part of the difference comes from decisions at the extensive margin: firms in the simplified scheme are less likely than those in the VAT scheme to trade with any VAT-paying firm, and, even when they do, they have less VAT-paying clients and suppliers. Figure 1 plots trade with VAT-paying firms as a function of firm size separately for firms in different tax schemes. We see that VAT-paying firms trade more with other VAT-paying firms than firms in the simplified scheme at all points of the size distribution below the 5 million threshold.

Table 1 and Figure 1 provide evidence of partial segmentation of supply chains by tax

scheme. This could be due to different characteristics of VAT- and non-VAT-paying firms, unrelated to their tax scheme, that lead them to choose not to trade with each other. Table 2 assesses whether this is the case by considering the correlations between a firm's own tax scheme (a variable equal to 1 if the firm is in the VAT scheme, 0 otherwise) and the share of its sales (intermediate inputs) that is sold to (purchased from) VAT firms, controlling flexibly for firm characteristics that affect their position in supply chains: the products they sell, their location and their size. We restrict the sample to firms which are eligible to choose their tax scheme; to allow for the possibility that firms could choose to produce more than 5 million whilst in the VAT scheme and bunch at the 5 million threshold whilst in the simplified scheme we consider all firms with a minimum turnover over the period of less than 7 million INR. We find that 40% of the correlation between firms' choice of tax scheme and how much they sell to VAT clients can be explained by firms in different tax schemes selling different products (columns 2 to 5) and/or being in different locations (columns 3 to 5) and of different size (column 5), though the correlation between tax scheme and purchases from VAT suppliers is unaffected by controls. The correlations remains large and statistically significant when controlling for all firm characteristics. Overall we find that, all else equal, VAT-paying firms sell 14 percentage points more to, and buy 8 percentage points more from, other VAT-paying firms than non-VAT-paying firms.

The types of products sold by firms are presented in Appendix Table B.1. We see that the share of VAT-paying firms among firms eligible to choose the simplified scheme is highest for products often used as intermediate inputs (machines, metal product and mining) and lowest for products more commonly sold to households (household goods, textiles and food). This is in line with the idea that firms selling to non-VAT-paying clients are less likely to choose to be in the VAT scheme, and explains why product fixed effects decrease the correlations in Table 1.

3 Model

We model an economy in partial equilibrium in which two different types of firms, upstream and downstream, take sourcing and tax decisions to maximize their profits. Upstream firms produce using only labor and sell to downstream firms and final consumers, whilst downstream firms produce using inputs purchased from upstream firms and sell only to final consumers.

We assume monopolistic competition, as is standard in the literature, and do not allow firms to under-report their tax liabilities to the tax authorities. We also assume a simple

supply chain structure with only two layers of firms, unlike recent papers in macroeconomics which use very general production networks.[11] We discuss the role played by these three simplifying assumptions at the end of the section.

3.1 Preferences and demand

The final consumer is endowed with exogenous income E and has CES preferences over a fixed and finite set of goods i:

$$U = \left(\sum_i (\beta_i q_{iF})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{1}$$

where q_{iF} is the quantity of good i consumed by the final consumer. Writing p_{iF} the consumer price of good i, utility maximization yields the following demand for good i:

$$q_{iF} = \left(\frac{\beta_i}{p_{iF}} \right)^{\sigma} P_F^{\sigma-1} E \tag{2}$$

where $P_F = \left(\sum_i \beta_i^{\sigma} p_{iF}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ is the consumer price index. We assume that final goods are substitutes ($\sigma > 1$).

3.2 Production and market structure

There is a fixed and finite set K of downstream firms that produce goods k using a CES input bundle of goods j with elasticity of substitution $\rho > 1$.
The production function of the firm producing good k, which we call firm k, is:

$$q_k = \phi_k \left(\sum_{j \in J} \alpha_{jk} q_{jk}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \tag{3}$$

where q_{jk} are the quantities of good j purchased by firm k, J is the set of upstream firms described below (J disjoint from K), ϕ_k is a productivity parameter and the α_{jk} terms are technology parameters. Writing p_{jk} the price paid by k for good j, we can write demand of firm k for good j as:

$$q_{jk} = \frac{q_{kF}}{\phi_k} \left(\frac{\alpha_{jk} P_k}{p_{jk}} \right)^{\rho} \tag{4}$$

[11] See for example Tintelnot et al. (2018); Lim (2018); Liu (2019); Huneeus (2019); Baqaee and Farhi (2020).

and firm k's cost function as:

$$c_k = \frac{P_k}{\phi_k} \text{ with } P_k = \left(\sum_{j \in J} \alpha_{jk}^{\rho} p_{jk}^{1-\rho} \right)^{\frac{1}{1-\rho}} \tag{5}$$

where P_k is firm k's input price index.

There is a fixed and finite set J of upstream firms that produce goods j using only labor as an input and sell to downstream firms and final consumers. The production function of firm j is: $q_j = \phi_j q_{\ell j}$ and its cost function is $c_j = \frac{P_j}{\phi_j}$ with $P_j = w$ where w is the exogenous cost of labor.

We assume the market structure is monopolistic competition so that firms sell to consumers at a mark-up $\mu = \frac{\sigma}{\sigma - 1}$ and to other firms at a mark-up $\nu = \frac{\rho}{\rho - 1}$.

3.3 Taxes and tax scheme choice

Downstream and upstream firms choose whether to pay taxes under the VAT scheme or under the simplified tax scheme. Under the VAT scheme firm i pays a tax t_i on its sales and deducts the VAT paid on its input purchases from its tax liabilities. Under the simplified scheme it pays a tax τ on its total sales and is constrained to sell less than a fixed amount: $x_i \leq \bar{x}$ where x_i is the firm's total sales in value. In what follows we assume that τ approximates to zero and all the VAT rates t_i are small compared to one. These assumptions are in line with our empirical context and simplify the expressions.[12]

We write v_i the tax scheme of firm $i \in K, J$, with $v_i = 1$ if i chooses to pay taxes under the VAT scheme, zero otherwise. Defining the tax wedges $\gamma_{iF} = 1 - \tau - v_i(t_i - \tau)$ on sales from firm $i \in K, J$ to the final consumer and $\gamma_{jk} = (1 - \tau - v_j(t_j - \tau) + v_j v_k t_j)$ on sales from firm $j \in J$ to firm $k \in K$, we can write the prices to final consumers and to intermediate firms as:

$$p_{iF} = \frac{P_i \mu}{\phi_i \gamma_{iF}}, \forall i \in K, J \tag{6}$$

$$p_{jk} = \frac{P_j \nu}{\phi_j \gamma_{jk}} \forall j \in J, k \in K \tag{7}$$

Firm i's profits when its sales are unconstrained can be written as:

[12] The trade-offs associated with choosing a tax scheme in our setting are similar to the trade-offs associated with formality choice in De Paula and Scheinkman (2010).

$$\Pi_i(v_i) = q_{iF}(\gamma_{iF} p_{iF} - c_i), \ \forall i \in K \tag{8}$$

$$= q_{iF}(\gamma_{iF} p_{iF} - c_i) + \sum_{k \in K} q_{ik}(\gamma_{ik} p_{ik} - c_i), \ \forall i \in J \tag{9}$$

We assume that firms choose the tax scheme v_i that maximizes their profit Π_i taking all other firms' tax scheme as given. Some firms choose the VAT scheme regardless of their size, governed by ϕ_i. Other firms choose the simplified scheme for small values of ϕ_i, the VAT scheme for large values of ϕ_i. Among those, some choose to sell exactly $\bar{x}$ and remain in the simplified scheme for intermediate values of ϕ_i.

3.4 Equilibrium

Market clearing implies that $q_k = q_{kF}, \forall k \in K$ and $\sum_k q_{jk} + q_{jF} = q_j, \forall j \in J$. An equilibrium is characterized by the tax scheme of all firms, $\{v_i\}$, which in turn determines prices and production through equations (2), (4), (6), and (7).[13] Using firm-level prices and production we obtain each firm's position in the supply chain, which we characterize using the terms s_{jk}, the share of firm k's purchases from firm j in its total input costs, and λ_{jk}, the share of firm j's sales to firm k in its total sales. Below we also refer to equilibrium values in an world in which there are no taxes (the 'no-tax' world in which $t_i = \tau = 0, \forall i$). In this world the supply chain parameters denoted by $\tilde{s}_{jk}$ and $\tilde{\lambda}_{jk}$ are a function of technology, productivity and mark-up parameters only.

Our first proposition considers the impact of a change in a downstream firm k's tax scheme on its trade with upstream firms j, keeping the tax scheme of all other firms' constant.

Proposition 1. ***Impact of tax system on trade (supply chain distortions).*** *The effect of a change in downstream firm k's tax scheme on its trade with upstream firm j can be expressed as:*

$$\log(s_{jk}(v_k = 1)) - \log(s_{jk}(v_k = 0)) = (\rho - 1)(t_j v_j - \bar{t}_k s_{Vk0}) \tag{10}$$

where s_{Vk0} is the share of k's inputs purchased from VAT-paying suppliers when k is in the simplified scheme and $\bar{t}_k$ is a weighted average of the VAT rates of the suppliers of firm k, defined by $(1 - \bar{t}_k)^{\rho-1} = s_{Vk0}^{-1} \sum_{j \in J} \tilde{s}_{jk} v_j (1 - t_j)^{\rho-1}$.
Proof: see Appendix.

[13] The propositions below consider small changes along the equilibrium and do not impose the restriction of a unique equilibrium.

This proposition states that the tax system causes supply chain distortions: downstream firms will, all else equal, buy more from VAT-paying upstream firms when they themselves pay VAT. Consider the impact of a change in downstream firm k's tax scheme on how much it buys from an upstream firm j in the VAT scheme, the expression we take to the data below. Assuming for simplicity that all of firm k's suppliers pay the same VAT rate t_j, we obtain:

$$\log(s_{jk}(v_k = 1)) - \log(s_{jk}(v_k = 0)) = (\rho - 1)(1 - s_{Vk0})t_j \tag{11}$$

This supply chain distortions mechanism leads to partial market segmentation between VAT- and non-VAT-paying firms, even in a world in which firms' tax schemes are exogenously given. This is because firms pay a tax on their purchases from VAT-paying suppliers only when they themselves do not pay VAT: inputs purchased from suppliers in the VAT scheme are cheaper for firms in the VAT scheme than for firms in the simplified scheme. The effect of downstream firm k's tax scheme on its purchases from a VAT-paying supplier j is moreover decreasing in s_{Vk0}, how much the firm buys from VAT-paying suppliers when it is in the simplified scheme. This effect goes through firm k's input price index P_k: the more k buys from VAT-paying suppliers the more P_k decreases when it becomes VAT, mitigating the effect of the decrease in input cost of any particular VAT-paying supplier. At the limit when firm k already buys all its inputs from VAT-paying suppliers when in the simplified scheme ($s_{Vk0} = 1$) the relative price of its inputs is unaffected by its choice of tax scheme, so a change in its tax scheme does not affect its input mix.[14]

Our second proposition characterizes the impact of firms' supplier networks on their propensity to choose to pay taxes under the VAT scheme. We define a firm's propensity to choose the VAT scheme as the difference between the profit it obtains when in the VAT scheme and the profit it obtains in the simplified tax scheme. Firms' position in supplier networks are by definition endogenous to their choice of tax scheme so we write firms' tax scheme choice as a function of their position in the no-tax network, which is exogenous to their tax choices.

Proposition 2. ***Strategic complementarities in firms' tax scheme choice.*** *Firm i's propensity to choose the VAT scheme is:*

1. *Decreasing in the firm's own VAT rate t_i.*
2. *Increasing in $t_i\tilde{\lambda}_{iV}$ where $\tilde{\lambda}_{iV} = \sum_{k \in K} v_k \tilde{\lambda}_{ik}$ is the share of i's sales sold to VAT-paying clients, where each client is weighted by its sales share in the no-tax world.*

[14] This will be the case if $\alpha_{jk} = 0$ for suppliers j for which $v_j = 0$.

3. *Increasing in $\tilde{s}_{Vi}\bar{t}_i$, where $\tilde{s}_{Vi} = \sum_j v_j \tilde{s}_{ji}$ is the share of i's inputs purchased from VAT-paying suppliers, where each supplier is weighted by its input share in the no-tax world, and $\bar{t}_i$ is a weighted average of the VAT rates of i's suppliers: $(1-\bar{t}_i)^{\rho-1} = \tilde{s}_{Vi}^{-1} \sum_{j \in J} \tilde{s}_{ji} v_j (1-t_j)^{\rho-1}$.*

In addition, firm's propensity to bunch – produce exactly $\bar{x}$ to remain in the simplified tax scheme – is also decreasing in t_i and increasing in $t_i \tilde{\lambda}_{iV}$ and $\tilde{s}_{Vi}\bar{t}_i$.
Proof: see Appendix.

This proposition states that there are strategic complementarities in firms' tax decisions within supply chains: the more a firm buys from, and sells to, VAT-paying firms, the more likely it is to itself choose to pay VAT. Intuitively, firms with many potential VAT-paying suppliers will face a lower input price index if they choose to be in the VAT scheme rather than in the simplified scheme. Similarly firms with many potential VAT-paying clients will face more demand for their products if they choose to be in the VAT scheme.

3.5 Discussion

Three of our assumptions warrant discussion. First, our assumption of monopolistic competition implies that firms fully pass taxes through to their clients: the full incidence of taxes is paid by the buyer at all stages of production (equivalently, suppliers' mark-ups are not affected by taxes). This implies in particular that firms cannot charge different prices to clients in different tax schemes by setting different mark-ups on sales to clients in different schemes.[15] This assumption is not key to deriving our propositions. Intuitively both our supply chain distortions and strategic complementarities mechanisms stem from the fact that the tax system introduces a wedge between the price paid by the buyer and that received by the seller for only some transactions. VAT is paid on transactions between VAT-paying suppliers and clients in the simplified scheme (or final consumers), not on transactions between two firms that pay VAT. This wedge increases the relative cost of trade between firms in different tax schemes regardless of which of the trading partners effectively bears the burden of the tax. One could think of a different model in which suppliers bear part of the incidence. In this model VAT-paying suppliers would earn lower mark-ups the more they sell to clients in the simplified scheme. This would also lead to less trade, all else equal, between VAT-paying suppliers and clients in the simplified scheme (proposition 1). Proposition 2 would similarly be unaffected. This assumption does however affect the extent to which some of our empirical estimates can be used to identify model parameters of interest. We clarify when this is the case when discussing our empirical results below.

[15] Firms do charge different prices to firms and consumers (see expressions (6) and (7)) but this is because they have different preferences, leading to different mark-ups (μ and ν), not because of tax scheme choices.

Second, we assume firms pay taxes on their total tax liabilities: they cannot hide part of their real activity from the tax authorities. In doing so we abstract from the possibility that the reporting of transactions between VAT-paying firms by both parties involved in the transaction could affect firms' compliance decisions, and in particular their decision to under-report part of their sales. This assumption is motivated both by a limitation of our data (we only observe reported sales and have no information on evasion) and the empirical evidence, discussed above, showing that firms report substantially more sales to the tax authorities than the total of their third-party-reported sales (see also Appendix B.1).[16] We return to the possibility that some transactions may be under-reported when discussing potential sources of bias in our empirical estimates below.

Third, we assume a simple supply-chain structure, with only two layers of firms. One way to relax this assumption and more closely match our data would be to allow all firms to use both labor and intermediate products as production inputs, and to sell to firms as well as the final consumer (as in for example Tintelnot et al., 2018; Liu, 2019; Baqaee and Farhi, 2020). We would obtain a prediction similar to Proposition 1 with this set-up under the assumption that firms' tax schemes are exogenously given. We would however not be able to derive simple predictions regarding the determinants of firms' tax scheme choices, one of the key aims of this model. We do however relax this assumption when considering the effect of a hypothetical policy reform below, by allowing all firms to buy from other firms.

Overall the model predicts that there will be partial segmentation of supply chains between VAT- and non-VAT- paying firms in equilibrium, because of two mechanisms. The supply chain distortions mechanism, detailed in Proposition 1, states that the tax system distorts firms' choice of input mix and leads to more trade, all else equal, between firms in the same tax scheme than between firms in different schemes. The strategic complementarities mechanism, detailed in Proposition 2, states that firms with many VAT-paying trading partners are more likely to choose to pay VAT, re-enforcing market segmentation. The following sections provide evidence regarding both mechanisms, then combine our model and estimates to consider the impact of a hypothetical policy reform that removes the source of both complementarities and supply chain distortions.

[16] The compliance characteristics of the VAT are moreover well studied in the existing literature

4 Empirical strategy

4.1 Graphical evidence

Our model predicts that VAT-paying firms trade more with other VAT-paying firms than firms in the simplified scheme. In particular we predict that firms that trade more with VAT-paying partners are less likely to constrain their sales to be just under the 5 million threshold in order to qualify for the simplified tax scheme. Figure 2 shows this prediction is borne out by the data. We plot the distribution of firms by turnover around the threshold separately for firms with below and above median shares of sales sold to VAT-paying clients and shares of intermediate inputs purchased from VAT-paying suppliers.[17] As predicted, we see more bunching among firms that sell less to, and purchase less from, VAT-paying firms. The model also predicts that firms facing a higher VAT rate will be more likely to bunch when their share of sales to VAT-paying clients decreases, relative to firms facing a low VAT rate. Appendix Figure C.6 tests this prediction: we see substantially more bunching among firms with a low share of VAT sales in the group facing a higher VAT rate.

This graphical evidence is in line with the model's predictions, but unobserved firm characteristics may be driving part of the cross-sectional correlation between firms' tax scheme choice and trade with VAT-paying firms. In what follows we address this concern by using within-firm and within-trading-relationship changes over time to estimate both the causal effect of the tax system on firms' sourcing decisions and strategic complementarities in firms' tax choices. Figure 3 presents graphical evidence regarding the within-firm correlation over time between tax scheme and trade. It plots the average share of intermediate inputs purchased from VAT-paying firms before and after firms change tax scheme as well as the average in each year for firms that never change tax scheme. We observe 7,648 firms changing tax scheme over time, 60% of them go from the simplified scheme to the VAT scheme.

We see a clear positive correlation between a firm's decision to switch to a new tax scheme and its sourcing decisions, with firms buying 10-20 percentage points more of their inputs from VAT-paying firms when they enter the VAT scheme. This correlation suggests at least one of the mechanisms outlined by our model is at play: firms may be entering the VAT scheme because their suppliers enter the VAT scheme (strategic complementarities) and/or buying more from VAT suppliers because they've chosen to enter the VAT scheme (supply chain distortions). This section presents the empirical strategy that enables us to separately

[17] All the extra mass just below the threshold comes from the sample of firms in the simplified scheme, see Appendix Figure C.5.

estimate the magnitude of each mechanism.

4.2 Supply chain distortions

We test Proposition 1 by considering the causal effect of a change in firms' tax scheme on their sourcing decisions. We start from expression (11) characterizing the impact of a change in client k's tax scheme on its purchases from VAT-paying supplier j, augmented to allow for pair jk specific fixed effects and supplier-year jt shocks.

$$\log(s_{jkt}) = (\rho - 1)(1 - s_{Vk0}))t_j v_{kt} + \gamma_{jk} + \gamma_{jt} \tag{12}$$

We therefore estimate the following equation:

$$\log(s_{jkt}) = \beta_1 v_{kt} + \beta_2 v_{kt} s_{Vk0} + \gamma_t + \gamma_{jk} + \gamma_{jt} + \epsilon_{jkt} \tag{13}$$

where s_{jkt} is the share of the transaction between client k and supplier j in k's total intermediate input purchases in year t, $v_{kt} = 1$ if the client k is in the VAT scheme in year t, 0 otherwise, s_{Vk0} is the average share of k's inputs purchased from VAT-paying suppliers when k is in the simplified scheme, γ_{jk} is a pair jk fixed effect and γ_{jt} different year fixed effects for each supplier.[18] We allow for potential changes in input mix as firms grow by controlling for the client firm k's turnover; this ensures that our estimates of β_1 and β_2 are not capturing effects going through firm size. Because the client's turnover can be determined jointly with its VAT choice when firms are close to the 5 million threshold our preferred specification considers only pairs in which the client has a turnover of less than 4 million INR. We allow for correlation in error terms both within location (postcode) and within types of product sold by firm k.

Our model predicts $\beta_1 = (\rho - 1)t_j = -\beta_2$ and therefore yields three testable predictions. First, we expect $\beta_1 > 0$ and $\beta_2 < 0$. Second, the model predicts that a change in tax scheme will have no impact on trade if the client purchases all its inputs from VAT suppliers, so we test for $\beta_2 + \beta_1 = 0$. Third, the magnitudes of β_1 and β_2 are expected to increase in the VAT rate t_j paid by the supplier. We split our sample between pairs in which the supplier pays the high and the low VAT rate to test this prediction. Finally, under our assumptions of monopolistic competition and CES production our estimate of β_1 can be used to identify the elasticity of substitution in production ρ using the expression $\rho = 1 + \frac{\beta_1}{t_j}$.

Several identification challenges must be addressed for specification (13) to identify the

[18] For firms that are never observed in the simplified scheme we use the average share of inputs from VAT suppliers over the period to proxy for s_{Vk0}. Note that these firms are not used to identify β_1 or β_2.

causal effect of within firm changes in tax scheme over time on their trade with VAT firms. First, strategic complementarities (or correlated shocks) and the fact that we do not observe trade between firms in the simplified tax scheme may lead to reverse causality: if two firms j and k trade whilst in the simplified scheme, and firm j enters the VAT scheme in the same year t as firm k we will observe no trade between j and k before t and positive trade after t, even if the real trade between both firms does not change in year t. To circumvent this issue we restrict our sample to pairs (j,k) in which the supplier j is always much bigger than the eligibility cut-off size (our baseline sample is restricted to suppliers with a minimum turnover over the period larger than 7 million INR). This ensures that the supplier is constrained to remit VAT throughout the period so that we always observe the transaction when the pair trades.

Second, some pairs may be more likely to trade for reasons we do not model but are correlated with their choice of tax scheme. Firms whose owners belong to the same community for example may be more likely to both trade with each other and share information on the tax system. We allow for such unobserved determinants of trade by including pair fixed effects (γ_{jk}) in all specifications. Third, reverse causality may be a cause for concern even when we restrict the sample to pairs in which the potential supplier is always in the VAT scheme. Shocks to VAT firms' productivity may make them more attractive to all potential clients and induce some non-VAT firms to buy from them. Strategic complementarities imply that some of these firms may choose to enter the VAT scheme because they have acquired a new VAT supplier. We include supplier$\times$year fixed effects (γ_{jt}) to allow for such unobserved changes in suppliers' productivity over time. Our preferred specification thus identifies the causal effect of interest by comparing the relative changes over time in trade between a large VAT-paying supplier and its clients that change tax scheme and those that do not. Our identifying assumption is that there are no unobserved pair-specific productivity shocks that vary over time and lead some clients to both start trading with large VAT suppliers and change tax scheme.

Finally, our estimate of β_1 could be biased upwards if large VAT-paying suppliers are less likely to report a client when this client leaves the VAT scheme. As explained above, and detailed in Appendix B.1, the evidence suggests most firms have no incentive to undertake such asymmetric under-reporting. We present results excluding the suppliers that may have an incentive not to report transactions with clients in the simplified tax scheme as a robustness check.

We restrict our attention to pairs (j,k) that trade at least once over the period in years during which both firms k and j file tax returns. We impute a value equal to the minimum

reporting threshold of 50,000 INR to transactions between firms that do not trade in a given year. Our baseline sample consists of 2.6 million observations and 506,787 pairs; in 32,144 pairs the client changes tax scheme over the period.

Table 3 presents descriptive statistics on the sample used to estimate (13), separately for pairs in which the supplier faces the reduced VAT rate and those in which it faces the normal VAT rate.[19] We see that the firms are only observed trading 55-61% of the time, so there is substantial variation on the extensive margin of trade. The average transaction represents a very small share of suppliers' sales (0.5%) but a non trivial share of client's intermediate inputs, as expected in a sample of pairs in which the client is small and the supplier large. Our identification strategy relies on comparing clients of the same supplier over time, it is therefore reassuring to see that the average supplier has 83 clients in the reduced tax group and 133 clients in the normal tax group (the median number of clients are, respectively, 43 and 44).

4.3 Strategic complementarities in tax choices

Our model assumes that when firms choose whether or not to pay VAT they take into account not only the parameters of the tax system (such as the VAT rate) but also the behavior of others in their supply chains. We test our Proposition 2, which stems from this assumption, by estimating the causal effect of changes in how much firms trade with VAT-paying firms on their choice of tax scheme. Following our predictions for firms' tax scheme choice we estimate the following equation on the sample of firms that can choose between the simplified and the VAT scheme:

$$v_{it} = \delta_1 t_i \sum_k \lambda_{ikt} v_{kt} + \delta_2 \sum_j t_j s_{jit} v_{jt} + \gamma_i + \gamma_{gt} + \epsilon_{it} \tag{14}$$

where v_{it} is equal to 1 if firm i is in the VAT scheme in year t, λ_{ikt} is the share of i's sales sold to firm k in year t, s_{ji} is the share of i's intermediate inputs purchased from firm j, and γ_i and γ_{gt} are respectively firm and year×product fixed effects.[20] We allow for correlation in error terms both within location (postcode) and within types of product sold by firm i and include firm i's turnover to control for firm size in all specifications.

[19] There are very few pairs in which the supplier faces the super-reduced 0% and 1% rates (2% of our sample). Our model predicts $\beta_1 = 0$ for these pairs, but there are too few pairs such that the supplier has several clients and at least one of them changes tax scheme over time for us to test this prediction. We therefore exclude these pairs from the analysis.

[20] We include product×year fixed effects to control for potential effects of the small increase in VAT rates on most products during the period.

Our estimates of interest – δ_1 and δ_2 – are 'supplier network effects' similar to the social effects estimated in the social networks literature (see for example Giorgi et al., 2010). Proposition 2 predicts both $\delta_1 > 0$ and $\delta_2 > 0$. Several challenges arise when attempting to identify such network effects; to circumvent them we use instruments for the terms $t_i \sum_k \lambda_{ik} v_{kt}$ and $\sum_j t_j s_{ji} v_{jt}$ in which we shut the sources of variation that could bias our estimates.

First, Proposition 1 makes clear that the structure of the network (the existence and size of a link between two firms) is endogenous to firms' tax scheme decision. Using the observed network would therefore bias our estimates upwards. The model also offers a solution to this problem: the relevant network is not the realized network but the potential network in the absence of taxes, characterized by the $\tilde{\lambda}_{ik}$ and $\tilde{s}_{ji}$ terms which are by definition exogenous to firms' tax choices. To proxy for this potential network we use the fact that $\tilde{s}_{ji} = s_{ji}$ when firm i is in the VAT scheme, and $\tilde{\lambda}_{ik} = \lambda_{ik}$ when i is in the simplified scheme: we proxy for $\tilde{\lambda}_{ik}$ and $\tilde{s}_{ji}$ using the average sale shares and the average input shares observed when firms are in the VAT scheme.[21] Importantly this method yields time-invariant network variables; combined with firm fixed effects this specification ensures that changes in the network cannot be driving our estimates.

Second, network effects naturally give rise to a reflection problem, compounded by the possibility of unobserved correlated effects across firms driving tax choices (Manski, 1993). This implies that using variations in the right-hand-side variables that comes from changes in firms' trading partners' tax schemes over time could also bias the estimates. In practice however only a small share of the within-firm variation in both these variables comes from changes in firms' trading partners decision to pay VAT (less than 5%). This is because firms that change tax scheme over time are by definition small, and represent only a small fraction of their trading partners' sales and input purchases.[22] We shut down this source of variation by holding the tax scheme of firm i's trading partners fixed. To do this we set v_{kt} and v_{lt} equal to the tax scheme of firms k and l observed in their first year in the data, v_{k0} and v_{l0}. The variation that remains comes from the fact that firms' trading partners enter and exit the data over time. We observe on average 9% of firms entering the data, and 8% exiting the data, in each year.[23] When firms are not in our data they are either not

[21] See Appendix C.2 for a proof that these are the best proxies available. When a firm is never observed in the VAT scheme we use the average share over the period. This introduces some measurement error, but our parameters of interest are identified using firms that change tax scheme over time which by definition we observe at least once in the VAT scheme.

[22] Small VAT-paying firms only buy 17% of their intermediate inputs from other small VAT-paying firms (one-fourth of their total inputs from VAT-paying firms), and sell them 11% of their sales (one-third of their total sales to VAT-paying firms).

[23] Entry and exit rates are lower for firms with a turnover of more than 7 million (5% entry rate, 4% exit

operating, or operating in the informal sector and therefore not filing taxes. Under both these scenarios entry and exit of a firm's VAT-paying trading partner therefore changes how much this firm can potentially trade with VAT-paying firms.

Formally, we use as instruments the predicted share of a firm's sales it can potentially sell to VAT-paying clients, and the predicted share of its intermediate input purchases it can potentially buy from VAT-paying suppliers: the variables $t_i \sum_k \tilde{\lambda}_{ik} v_{k0} e_{kt}$ and $\sum_j t_j \tilde{s}_{ji} v_{j0} e_{jt}$, where e_{kt} (e_{jt}) is a variable equal to 1 if supplier k (client j) is in the data in year t, 0 otherwise. The only source of variation over time in our instruments comes from the fact that firms' potential VAT-paying trading partners enter and exit from the data over time. These instruments enable us to identify our estimates of interest under the assumption that entry and exit of firms' VAT-paying trading partners only affect their tax scheme decision through their propensity to sell to, or buy from, VAT-paying firms. Because entry and exit of trading partners likely affect firms' size this identification assumption is potentially violated for firms close to the 5 million INR threshold for which size is co-determined with the tax scheme choice. We therefore consider results on a sample consisting of all firms eligible to choose their tax scheme (those with a turnover below 7 million INR) and on a sample consisting of only firms whose turnover is less than 4 million INR, for which we control for firm size. Note that this size restriction is the same as the one we apply to the sample used to estimate supply chain distortions. Both our mechanism estimates are thus identified from the same sample of firms.

Appendix Table C.3 presents descriptive statistics on the sample used to estimate equation (14). We see that over 90% of firms have at least one VAT-paying trading partner over the period, and 60% have at least one VAT-paying trading partner that exits or enters over the period. These trading partners are much larger than the firms whose choice of tax scheme specification (14) seeks to explain, their entry or exit from the data therefore represent substantial changes to how much these firms can trade with VAT-paying firms.

5 Results

5.1 Supply chain distortions

Table 4 presents results obtained by running specification (13), where the outcome variable is the log of the ratio of the transaction between two firms to the total intermediate input purchases of the client. In the first three columns the sample consists of all pairs in which

rate), but comparable across tax scheme among firms with a turnover of less than 7 million: entry (exit) rates are 8% (8%) for firms in the simplified tax scheme, 11% (10%) for firms in the VAT scheme.

the supplier is in the medium tax schedule (facing a VAT rate of 4-5%), and in the last three columns the supplier is in the high tax schedule (facing a VAT rate of 12.5-13.5%). Columns 2-3 and 5-6 present results obtained using our preferred specification with supplier×year fixed effects to allow for unobserved shocks to suppliers' productivity.

Results indicate that the average firm trades more with VAT suppliers when it enters the VAT scheme: $\beta_1 > 0$, as predicted by the model. Effects are smaller, as expected, when we include supplier×year fixed effects. On average firms that enter the VAT scheme buy 11% more from their VAT-paying suppliers in the medium tax scheme (column 2), 17% more from those in the high tax scheme (column 5); 12% on average across the whole sample. The relative magnitude of the effects in the two sub-samples is also in line with the model's prediction that the causal effect of a client entering the VAT on trade with VAT-paying suppliers is increasing in the tax rate paid by the suppliers ($\beta_1 = (\rho - 1)t_j$).

Columns 3 and 6 show that effects vary with how much the client trades with VAT-paying suppliers when in the simplified scheme (captured by the variable labelled 'VAT input share'): the effect of joining the VAT scheme is smaller the higher this share. The magnitude of the estimates imply that there is no effect of joining the VAT scheme on trade with VAT-paying suppliers for a firm that only buys its inputs from VAT suppliers regardless of its tax scheme ('VAT input share' equal to 1). This is in line with the model's prediction that $\beta_1 = -\beta_2$. Appendix Figure D.7 plots the estimates of the effect obtained for each quintile of the distribution of the VAT input share variable. We see that the effect decreases as the VAT input share increases, as predicted.

We present several robustness checks in the Appendix. First, we consider the possibility that VAT-paying suppliers under-report transactions with clients in the simplified tax scheme; this would bias our results upwards. To do so, we remove from the sample all pairs in which the supplier has a share of third-party-reported sales to total sales of more than 90%. As explained in Section 2 these suppliers are the only ones that may have an incentive not to report clients in the simplified scheme to lower their tax liabilities. This reduces our sample size by 25% but estimates are very similar, suggesting under-reporting of transactions is not driving our results (see Appendix Table D.4, panel A).[24] Results are also unaffected when we allow for location specific shocks (Appendix Table D.4, panel B) or consider alternative sub-samples of potential pairs (Appendix Table D.5). Overall, firms trade 10-12% more with VAT-paying suppliers in the medium tax scheme and 15-20% more with those in the high tax scheme when they enter the VAT scheme. We obtain similar

[24]Only 10% of the firms in our data sell more than 90% of their sales to other tax-registered firms but these firms are over-represented as suppliers in this sample because they are involved in more transactions than the average firm.

results when using an indicator equal to one if the pair trades as an outcome variable, suggesting most of the effects are driven by the extensive margin of firms deciding whether or not to trade (see Appendix Table D.4, panel C).

Our estimates can be compared to the existing literature in two ways. The value of the elasticity of substitution in production ρ that our results imply (under our assumptions of constant markups) is relatively stable across samples and takes a value between 4.5 and 6.5. This is within the range of estimates reported in the literature, though previous work typically does not use firm-level transaction data to identify this parameter (see for example Bas et al., 2017; Broda et al., 2017). In our model $1-\rho$ is also the trade elasticity (elasticity of trade shares with respect to trade costs). Our estimates imply that this elasticity is in the -3.5 to -5.5 range in this sample, again well within the set of estimates obtained in the existing literature (Bartelme et al., 2018, see).

5.2 Strategic complementarities in tax choices

Table 5 presents results obtained by running specification (14) which models a firm's tax scheme choice as a function of the share of its intermediate inputs purchased from VAT-paying suppliers and the share of its sales sold to VAT-paying clients. The dependent variable is an indicator equal to 1 if the firm pays taxes under the VAT scheme, 0 otherwise. All columns include firm and year fixed effects. The first three columns are obtained on the sample of all firms with a minimum turnover of less than 7 million INR, and the last three columns on the sample of all firms with a minimum turnover of less than 4 million INR. In the last three columns our specifications control for the firm's turnover.

We first present OLS results in columns 1 and 4. We find that changes over time in how much firms trade with VAT-paying firms are positively correlated with changes in firms' decisions to pay VAT. Variation in these trade shares over time come from three sources: i) changes in trade between firms ii) changes in the VAT decisions of the firms' clients and suppliers iii) entry and exit of clients and suppliers. The first two sources likely bias the estimates upwards relative to the causal effect of interest; in columns 2 and 5 we present reduced form results in which the right-hand-side variables are our predicted trade instruments, which only vary over time because of entry and exit of the firms' VAT-paying trading partners. This reduces the magnitude of the estimates, as expected, but the estimated effect of changes in predicted trade with VAT firms on firms' choice of tax scheme are large and statistically significant. These instruments are strong predictors of how much firms trade with VAT-paying firms. Appendix Table D.7 presents first stage results, we find that a 10% increase in predicted trade with VAT-paying firms leads to a 6-8% increase in

observed trade with VAT-paying firms.[25]

In columns 3 and 6 we present results obtained using our preferred specification in which the predicted trade shares are used as instruments for the observed trade shares with VAT-paying firms. Estimates are similar when we restrict the sample to firms far away from the threshold above which they have to be in the VAT scheme and control for firm size in column 6, suggesting the effect isn't driven by firms changing both size and tax scheme when they acquire or lose a trading partner. Appendix Table D.8 shows that the magnitudes of the effects are also unaffected when we change the thresholds used to define our sample or include year×location fixed effect to allow for local shocks which could potentially affect both entry dynamics and firms' choice of tax scheme.

Overall, we find that firms are strategic in their decision to pay VAT: they take the VAT status of their trading partners into account. Our results indicate that a 10% increase in the share of a firm's intermediate inputs purchased from VAT-paying suppliers increases the probability that this firm pays the VAT by 8-9 percentage points, and a 10% increase in the share of firm's sales that are sold to VAT-paying clients increases this probability by 1.4-1.5 percentage points. These estimates imply that a firm facing a VAT rate of 13.5% and whose suppliers also face that rate will increase its propensity to pay the VAT by 13-14 percentage points if all its trading partners simultaneously switch from the simplified to the VAT scheme.[26]

6 Discussion

6.1 Policy counterfactual

In this section we illustrate the magnitude and implications of our results by considering the effect of a simple hypothetical policy reform. This reform enables firms in the simplified scheme to deduct VAT paid by their suppliers from their tax liabilities, but leaves all other parameters of the tax system unchanged. This leads to a mechanical transfer of revenues from the government – which no longer raises taxes on transactions between VAT-paying firms and their clients in the simplified scheme – to firms, which make slightly higher

[25] We compute conditional F-statistics testing for weak instruments with multiple endogenous variables (Sanderson and Windmeijer, 2016). For all specifications, the values of these F statistics are much higher than the 5% Stock-Yogo critical values for a maximal 10% size of Wald tests for the two-endogenous-regressor, two-instrument case, equal to 7.03.

[26] These numbers are obtained by multiplying our estimates of strategic complementarities by the VAT rate, here 0.135. Using the estimates from column 6 in Table 5 we find $0.135 \times (0.887 + 0.154) = 0.141$.

profit.[27] Firms' behavioral responses to the reform yield two additional effects. This first is a growth in firms' output. By ensuring that VAT is never paid on transactions between any tax-registered firms, the reform removes tax-induced distortions in trade and leads to firm growth. The second is a loss of tax revenues due to weaker complementarities in firms' choice of tax scheme. The reform implies that firms with a high share of inputs purchased from VAT-paying suppliers no longer pay taxes on these inputs regardless of their tax scheme, so they no longer have an incentive to choose the VAT scheme.

We use our model to obtain expressions for these behavioral effects of the reform and combine our estimates obtained in section 5 with our data on supplier networks to calibrate these expressions. Appendix E details our procedure. We relax our model's assumption that firms are either downstream or upstream and allow firms to both buy from and sell to other firms. This enables us to consider effects spreading through supply chains. We find that the removal of supply chain distortions leads to a small increase in output of 0.8% for the average firm. The increase in output is much larger (7.5%) amongst firms in the simplified tax scheme than amongst firms in the VAT scheme. This is because the reform decreases the input costs of firms in this scheme, allowing them to charge lower prices and therefore sell and produce more.[28] Firms in the VAT scheme are affected negatively as competitors to firms in the simplified scheme (though effects are very small), and positively as suppliers to these firms, or as suppliers to firms upstream of firms in the simplified scheme. Positive effects going through supply chains dominate, because 47% of firms in the VAT scheme are in the same supply chain as at least one firm in the simplified scheme.

Figure 4 presents our estimate of firm growth due to this hypothetical policy reform for each decile of the firm size distribution prior to the reform. Effects are negligible (below 0.2%) amongst firms in the top four deciles of the distribution. We find non-trivial effects however in the bottom and middle of the distribution, which is where firms in the simplified scheme are located.[29] Overall, firm growth effects are concentrated amongst small-

[27] This transfer corresponds to less than 0.7% of the government's total VAT revenues.

[28] Restricting the effect of the reform so that no firm in the simplified scheme ever reaches an output level above 5 million INR hardly affects this number (from 7.55 to 7.54). This is because the effect of the reform on these firms' output is a function of how much they buy from VAT-paying suppliers, and we have seen that firms in the simplified scheme close to the 5 million threshold buy little from these suppliers, as predicted by our model (see Figure 2). An approximation for the effect on the output of firms in the simplified scheme is $\partial q_i \bar{t}_i \tilde{s}_{vi} \sigma$, where $\bar{t}_i \tilde{s}_{vi}$ captures the change in the firm's input cost P_i, itself a function of how much the firm buys from VAT suppliers ($\tilde{s}_{vi}$) and the VAT rate paid by these suppliers $\bar{t}_i$. For the average firm we have $\tilde{s}_{vi}$ close to 0.5, $\bar{t}_i$ equal to 0.06, using a value of 2.5 for σ (from DellaVigna and Gentzkow (2019)) we obtain $\partial q_i = 0.075$.

[29] The much smaller effect amongst firms in the first decile can be explained by the fact that 94% of firms in this group are in the VAT scheme. This is because a large share of these firms are 'voluntarily tax-registered' firms: they are not required to pay taxes because their turnover is below 500,000 INR, but chose to pay taxes

and medium-sized firms.

We then compute the amount of VAT revenues lost due to the fact that the reform removes some firms' incentives to choose the VAT scheme. Using our estimates of the strength of strategic complementarities in Table 5, we find that 3% of firms eligible to the simplified scheme would leave the VAT because of the reform. This only has a small effect on government revenues (equal to 0.8% of VAT revenues), because the firms that leave the VAT scheme are by definition small and pay only a very small share of the total VAT paid.[30]

We finally consider how much of the existing segmentation in supply chains would remain after the implementation of this reform. It removes the supply chain distortions mechanisms and (part of) the complementarities mechanism, so this exercise tells us how much of the observed segmentation can be explained by the incentives created by the VAT system. Appendix Table E.11 presents the correlation between firms' propensity to pay VAT and how much they trade with VAT-paying firms before and after the reform. We find that the reform would remove 50% of the 'upstream' supply chain segmentation (the correlation between firms' choice of tax scheme and the share of their inputs they purchase from VAT-paying suppliers) and 15% of the 'downstream' segmentation (the correlation with the share of sales sold to VAT-paying clients) observed at baseline in Table 2. The difference between the upstream and downstream effects can largely be explained by the fact that the reform does not affect the incentives of final consumers (and downstream non-tax registered firms) who still buy more, all else equal, from firms in the simplified scheme.[31]

Whether this hypothetical reform is optimal depends on the relative weight put by the government on small- and medium-size firm growth compared to the (small) fall in VAT revenues. This quantification exercise highlights how the design of tax policy can constrain firm growth, and in particular small firms' growth: we find that a small change in the VAT system has a non-trivial effect on these firms' output. This is in line with a large literature on firm size in the developing world which argues that tax-induced distortions may be one

regardless. These firms are all in the VAT scheme, as expected.

[30] This number is moreover an upper bound on the amount of tax revenues lost by the government due to the reform for several reasons, discussed in Appendix E. In particular the reform would also increase some firms' incentives to pay VAT via complementarities: the reform removes the disincentive to pay VAT faced by firms in the simplified scheme at baseline selling to other firms in this scheme. We cannot estimate this effect in the absence of data on transactions between firms in the simplified scheme, but note that this effect is likely to be small, both because complementarities effect with respect to clients are small (see Table 5 and because these firms are unlikely to be found upstream in supply chains (see Appendix Table B.1). See the Appendix for a more complete discussion of how this omission affects our results.

[31] In addition, whilst there is an increase in trade between clients in the simplified scheme and VAT-paying suppliers, it has a much smaller effect on these suppliers' sales than on these clients' inputs. This is because firms in the simplified scheme are substantially smaller than their VAT-paying suppliers. See Appendix E for a more complete discussion.

of the reasons firms remain small (see Hsieh and Olken, 2014, for a discussion). Note that this result would hold even in the absence of size-contingent tax regulations (here, the fact that firms must remain below the 5 million INR threshold to avoid paying VAT) which further restrict firm size.

6.2 Policy implications

Our results have more general implications for tax policy. First, they allow us to revisit debates regarding the relative efficiency of Value-Added-Taxes with respect to Retail Sales Taxes (RST) – taxes paid only by retailers. In a world with perfect tax compliance these two taxes are equivalent (Kopczuk and Slemrod, 2006). In contexts with imperfect compliance, previous literature has pointed out that the VAT may be more revenue-efficient because third-party-reporting on firm-to-firm transactions increases compliance (see for example Pomeranz, 2015). Our supply chain distortions mechanism implies however that there is also a production-efficiency disadvantage of the VAT with respect to the RST when some firms in the economy do not pay the tax: these firms' sourcing decisions are distorted by the VAT, but not by a RST. The above results suggest these distortions have a non-trivial effect on their production level.

Our evidence regarding the existence of strategic complementarities in firms' tax choices points to another difference between the two tax systems. Strategic complementarities imply that tax interventions that incentivize some firms to start paying VAT will have spillover effects on these firms' supply chains, as some of their suppliers and clients will also start paying VAT. These spillover – or multiplier – effects are intrinsically neither good nor bad from a policy perspective. However, in contexts in which many firms do not pay taxes, strategic complementarities will increase the tax returns of such interventions.

Our results also speak directly to the potential impact of India's recent large-scale VAT reform, known as the GST (Goods and Services Tax). Prior to the introduction of GST in 2017, each one of India's 27 states had its own VAT system, and firm-to-firm transactions across state borders were taxed. The GST reform created a centralized VAT system which harmonizes taxes across all Indian states, and in particular allows for VAT paid on their inputs to be deducted by VAT-paying firms even when the buyer and the seller are in different states. Our results suggest this reform has boosted inter-state trade by removing distortions in VAT-paying firms' sourcing decisions across suppliers in different states, thus increasing market integration in India. The strategic complementarities mechanism moreover implies that some firms may have joined the VAT as a consequence of the GST reform: the introduction of cross-state VAT-deductability increased the relative returns from paying VAT for

firms that trade with clients or suppliers in another state.

Our findings more generally have implications for tax policy in developing countries, most of which use a form of VAT (Ebrill, 2001). A key characteristic of these countries is that tax-paying firms co-exist in markets, and potentially trade with, a large number of non-tax paying firms in the informal sector. Our results regarding how the tax system affects trade between VAT and non-VAT-paying firms naturally extend to trade between VAT-paying firms and those in the informal sector, which we cannot observe in our data. Informal firms, like the firms in the simplified tax scheme we consider, pay taxes on purchases from VAT-paying suppliers and are therefore less likely to source inputs from them than from other informal firms. The magnitude of the effects we estimate suggests these distortions in input mix could be substantial. We similarly expect strategic complementarities in firms' decisions of whether to enter the formal sector under a VAT system. Our results imply that a compliance shock causing some informal firms to start paying VAT will have spillover effects on these firms' supply chains: their informal trading partners may start paying VAT themselves.[32]

7 Conclusion

In this paper we set out to understand how tax policy affects firm-to-firm trade and how firms' tax decisions are linked within supply chains, by looking at the role of VAT in a large developing economy. We use novel panel data from the state of West Bengal in India in which we observe both VAT- and non-VAT- paying firms and firm-to-firm transactions. This enables us to document the segmentation of supplier networks between firms in different tax schemes (VAT-paying and non-VAT-paying firms). We find evidence that VAT-paying firms trade more with other VAT-paying firms, all else equal, than non-VAT-paying firms.

To help us understand the mechanisms leading to market segmentation we then build a model of firms' sourcing and tax decisions within supply chains. Our key prediction is that under a VAT system there is partial market segmentation by tax scheme in equilibrium for two reasons. First, the VAT's incentive structure leads to supply-chain distortions: all else equal a VAT-paying firm buys a higher share of its inputs from VAT-paying suppliers than a non-VAT-paying one does. Second, there are strategic complementarities in firms' tax decisions: firms are more likely to choose to pay VAT the more VAT-paying suppliers and clients they have.

[32]See also Emran and Stiglitz (2005) for a discussion of the welfare properties of the VAT in the presence of an informal sector.

We provide empirical evidence on the mechanisms defined by our model using within-firm and within supplier-client pairs variations over time. We find that firms buy 12% more from VAT-paying suppliers on average when they themselves choose to pay VAT. Our estimates imply a trade elasticity and an elasticity of substitution in production that are within the range of estimates obtained in the international trade literature. We also find evidence of strategic complementarities in firms' tax choices. Our estimates imply that forcing all of a firm's trading partners to pay the VAT would increase that firm's propensity to pay the VAT by up to 14 percentage points compared to a situation where none of that firm's trading partners pay VAT.

Our results suggest that tax-induced distortions have economically meaningful consequences. In particular, we find that a hypothetical reform exempting transactions between all firms from the VAT would have non-trivial effects on the output of some small and medium-sized firms. Finally, we note that our analysis cannot consider how these distortions affect firms in the informal sector, which we do not observe in our data. We expect however similar mechanisms to also affect the production decisions of the many firms operating informally in developing countries.

References

Acemoglu, D., U. Akcigit, and W. Kerr (2016): "Networks and the Macroeconomy: An Empirical Exploration," *NBER Macroeconomics Annual*, 30, 273–335.

Agnosteva, D. E., J. E. Anderson, and Y. V. Yotov (2014): "Intra-national Trade Costs: Measurement and Aggregation," NBER Working Papers 19872, National Bureau of Economic Research, Inc.

Alfaro-Urea, A., I. Manelici, and J. Vasquez (2018): "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," Mimeo, University of California, Berkeley.

Allen, T. (2014): "Information Frictions in Trade," *Econometrica*, 82, 2041–2083.

Almunia, M., J. Hjor, J. Knebelmann, and L. Tian (2019): "Strategic or Confused Firms? Evidence from Missing Transactions in Uganda," Tech. rep., Mimeo, Columbia University.

Almunia, M., L. Liu, and B. Lo (2017): "VAT notches, voluntary registration and bunching: Theory and UK," Tech. rep., Oxford University Centre for Business Taxation Working Paper No.16/10.

Atalay, E., A. Hortaçsu, J. Roberts, and C. Syverson (2011): "Network structure of production," *Proceedings of the National Academy of Sciences*, 108, 5199–5202.

Atkin, D. and D. Donaldson (2015): "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," NBER Working Papers 21439, National Bureau of Economic Research.

Bachas, P. and M. Soto (2017): "Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement," Mimeo, World Bank.

Banerjee, A. V. and E. Duflo (2000): "Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry*," *The Quarterly Journal of Economics*, 115, 989–1017.

Baqaee, D. R. and E. Farhi (2020): "Productivity and Misallocation in General Equilibrium*," *The Quarterly Journal of Economics*, 135, 105–163.

Bartelme, D., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2018): "Economies of Scale and Industrial Policy: A View from Trade," Tech. rep., Mimeo, MIT.

Bas, M., T. Mayer, and M. Thoenig (2017): "From micro to macro: Demand, supply, and heterogeneity in the trade elasticity," *Journal of International Economics*, 108, 1–19.

Benzarti, Y., A. Tazhitdinova, and L. Bar-El (2018): "Do Value-Added Taxes Affect International Trade Flows? Evidence from 30 Years of Tax Reforms," Tech. rep., Mimeo, UCSB.

Bernard, A. B., E. Dhyne, G. Magerman, K. Manova, and A. Moxnes (2019): "The Origins of Firm Heterogeneity: A Production Network Approach," Working Paper 25441, National Bureau of Economic Research.

Bernard, A. B. and A. Moxnes (2018): "Networks and Trade," *Annual Review of Economics*, 10, 65–85.

Bernard, A. B., A. Moxnes, and Y. U. Saito (2015): "Production Networks, Geography and Firm Performance," Tech. rep., NBER Working Paper 201082.

Best, M., A. Brockmeyer, H. J. Kleven, J. Spinnewijn, and M. Waseem (2015): "Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan," *Journal of Political Economy*, 123.

Bird, R. and P.-P. Gendron (2007): *The VAT in Developing and Transitional Countries*, Cambridge University Press.

Boadway, R. and M. Sato (2009): "Optimal Tax Design and Enforcement with an Informal Sector," *American Economic Journal: Economic Policy*, 1, 1–27.

Boehm, J. and E. Oberfield (2018): "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," Working Paper 24937, National Bureau of Economic Research.

Brockmeyer, A. and M. Hernandez (2016): "Taxation, information, and withholding :

evidence from Costa Rica," Policy Research Working Paper Series 7600, The World Bank.

Broda, C., J. Greenfield, and D. E. Weinstein (2017): "From groundnuts to globalization: A structural estimate of trade and growth," *Research in Economics*, 71, 759–783.

Carrillo, P., D. Pomeranz, and M. Singhal (2017): "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement," *American Economic Journal: Applied Economics*, 9, 144–164.

Carvalho, V. M., M. Nirei, Y. Saito, and A. Tahbaz-Salehi (2016): "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake," Tech. rep., Becker Friedman Institute for Research in Economics Working Paper No. 2017-01.

Cosar, A. K. and P. D. Fajgelbaum (2016): "Internal Geography, International Trade, and Regional Specialization," *American Economic Journal: Microeconomics*, 8, 24–56.

De Paula, A. and J. A. Scheinkman (2010): "Value-Added Taxes, Chain Effects, and Informality," *American Economic Journal: Macroeconomics*, 2, 195–221.

DellaVigna, S. and M. Gentzkow (2019): "Uniform Pricing in U.S. Retail Chains*," *The Quarterly Journal of Economics*, 134, 2011–2084.

Dhyne, E., G. Magerman, and A. K. kikkawa (2019): "Imperfect Competition in Firm-to-Firm Trade," Working Papers ECARES 2019-05, ULB – Universite Libre de Bruxelles.

Di Giovanni, J., A. A. Levchenko, and I. Mejean (2018): "The Micro Origins of International Business-Cycle Comovement," *American Economic Review*, 108, 82–108.

Ebrill, L. P. (2001): *The Modern VAT*, Washington, D.C.: International Monetary Fund.

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

Fajgelbaum, P. and S. Redding (2018): "Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914," NBER Working Papers 20217, National Bureau of Economic Research.

Fajgelbaum, P. D., E. Morales, J. C. S. Serrato, and O. Zidar (2019): "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 86, 333–376.

Gadenne, L. (forthcoming): "Can Rationing Increase Welfare? Theory and an Application to India's Ration Shop System," *American Economic Journal: Economic Policy*.

Gerard, F., J. Naritomi, and A. Seibold (2020): "Tax systems and inter-firm trade: evidence from the VAT in Brazil," Mimeo London School of Economics.

Ghosh, M. and T. Nandi (2017): "From Sales Tax to GST : The journey of West Bengal in 75 years," CTRPFP Working Paper No. 2017-0.

Giorgi, G. D., M. Pellizzari, and S. Redaelli (2010): "Identification of Social Interactions through Partially Overlapping Peer Groups," *American Economic Journal: Applied Economics*, 2, 241–275.

Goldberg, P. K. and N. Pavcnik (2016): "The Effects of Trade Policy," NBER Working Papers 21957, National Bureau of Economic Research, Inc.

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

Hsieh, C.-T. and B. A. Olken (2014): "The Missing Missing Middle," *Journal of Economic Perspectives*, 28, 89–108.

Huneeus, F. (2019): "Production Network Dynamics and the Propagation of Schocks," Mimeo, Yale University.

Jensen, A. (2019): "Employment Structure and the Rise of the Modern Tax System," NBER Working Papers 25502, National Bureau of Economic Research, Inc.

Keen, M. (2008): "VAT, tariffs, and withholding: Border taxes and informality in developing countries," *Journal of Public Economics*, 92, 1892–1906.

Keen, M. and J. Mintz (2004): "The optimal threshold for a value-added tax," *Journal of Public Economics*, 88, 559 – 576.

Kopczuk, W. and J. Slemrod (2006): "Putting Firms into Optimal Tax Theory," *American Economic Review*, 96, 130–134.

Lim, K. (2018): "Endogenous Production Networks and the Business Cycle," Mimeo, Princeton Universtiy.

Liu, E. (2019): "Industrial Policies in Production Networks," *The Quarterly Journal of Economics*, 134, 1883–1948.

Macchiavello, R. and A. Morjaria (2015): "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports," *American Economic Review*, 105, 2911–45.

Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60, 531–542.

McMillan, J. and C. Woodruff (1999): "Interfirm Relationships and Informal Credit in Vietnam," *The Quarterly Journal of Economics*, 114, 1285–1320.

Naritomi, J. (2018): "Consumers as Tax Auditors," Mimeo, London School of Economics.

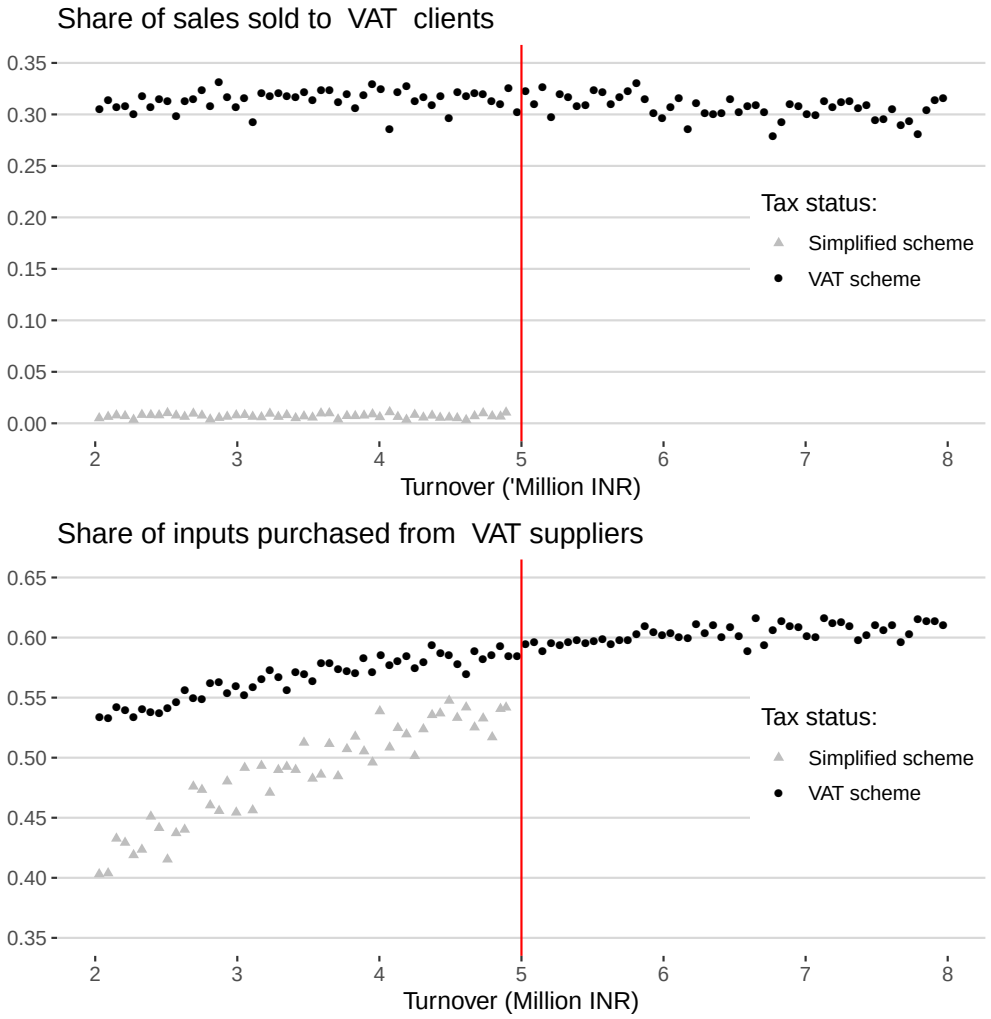
Pomeranz, D. (2015): "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 105, 2539–2569.

Sanderson, E. and F. Windmeijer (2016): "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 190, 212 – 221, endogeneity Problems in Econometrics.

Tintelnot, F., A. K. Kikkawa, M. Mogstad, and E. Dhyne (2018): "Trade and Domestic Production Networks," Working Paper 25120, National Bureau of Economic Research.

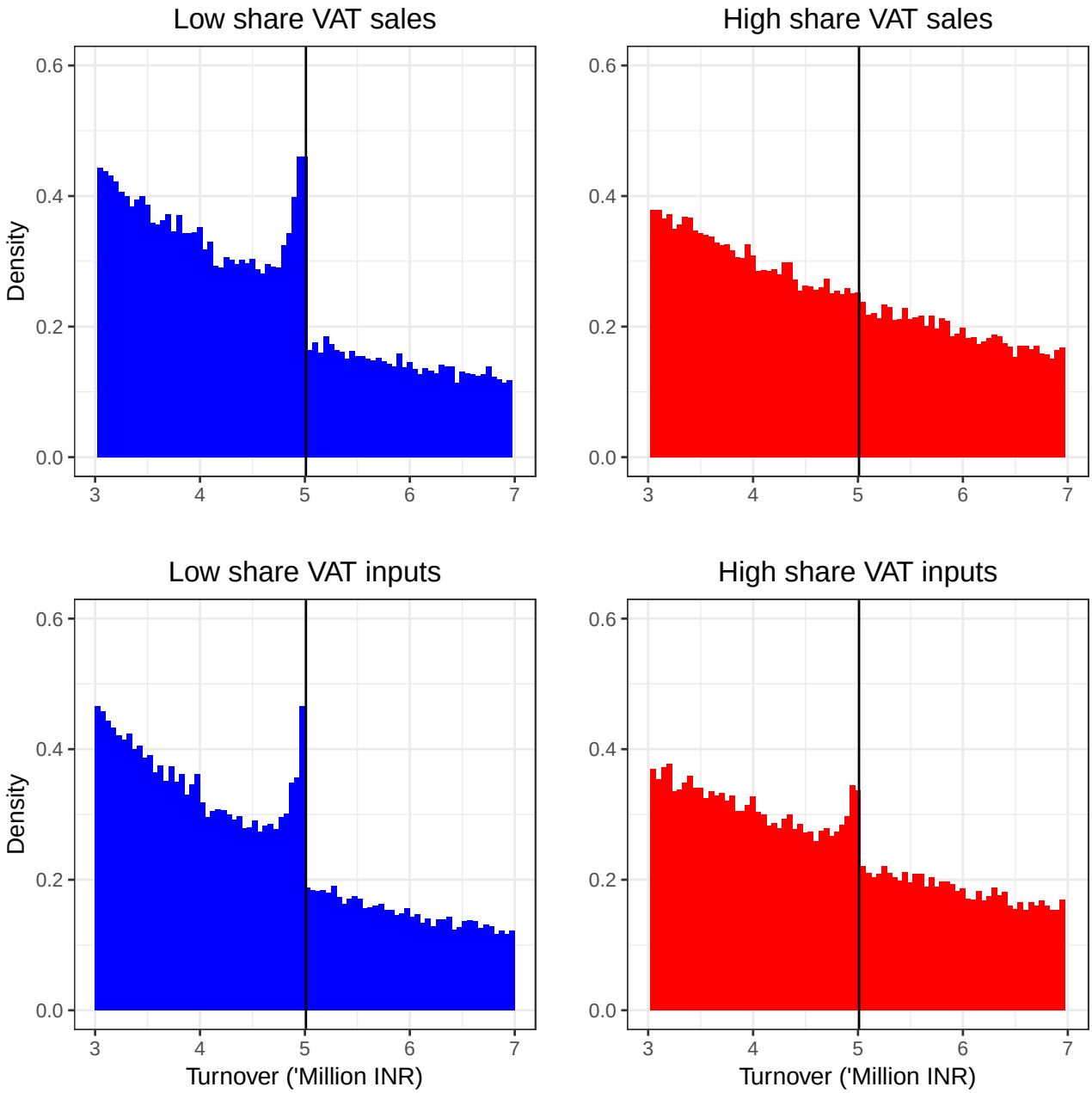
WTO (2004): "World Trade Report," Tech. rep., World Trade Organization.

Figure 1: Trade with VAT-paying firms and tax scheme choice



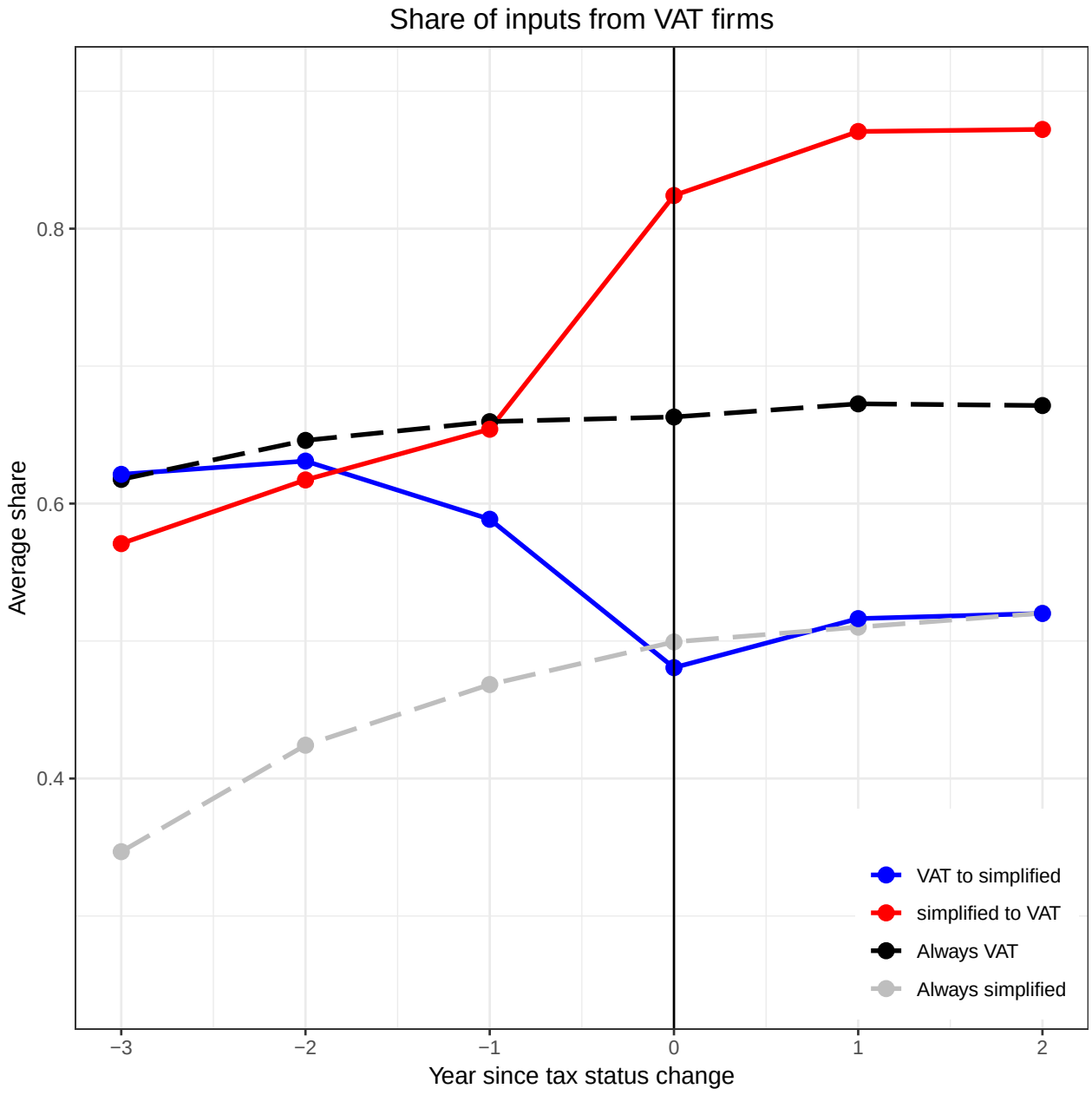
These graph plot the firm-level share of sales that are sold to VAT-paying clients (top graph) or the share of intermediate inputs purchased from VAT-paying suppliers (bottom graph) as a function of firm size. The black dots indicate VAT-paying firms, the grey triangles firms paying taxes under the simplified scheme. The vertical line indicates the size threshold above which firms have to pay VAT. We restrict the sample to firms with a turnover between 2 and 8 million INR, which represent 30% of the total sample.

Figure 2: Distribution of firm size and trade with VAT-paying firms



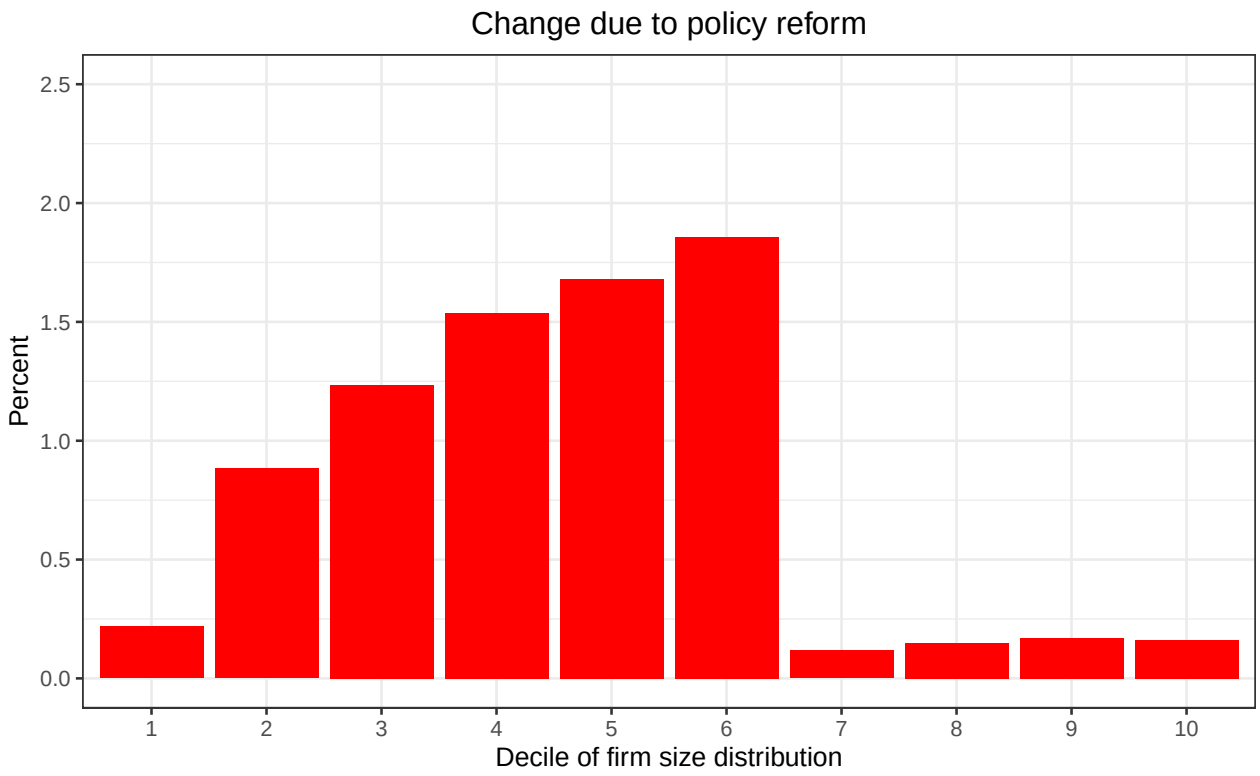
These graph plot the density distribution of firms by firm size for different sub-samples of firms with a turnover between 3 and 7 million INR. The top left (right) graph considers firms whose share of sales sold to VAT-paying clients is below (above) the sample median; the bottom left (right) graph considers firms whose share of inputs purchased from VAT-paying suppliers is below (above) the sample median. Bin sizes are 50,000 INR in all graphs.

Figure 3: Share of inputs purchased from VAT suppliers and tax scheme choice



This graph plots the average share of inputs purchased from VAT suppliers over time for four groups of firms with a turnover of less than 7 million: firms that switch from the simplified to the VAT scheme (red line, 3,861 firms), firms that switch from the VAT to the simplified scheme (blue line, 2,490 firms), firms that remain in the VAT scheme throughout the period (black line, 82,635 firms) and firms that remain in the simplified scheme throughout the period (grey line, 10,810 firms). Each point represents an annual average, where year 0 is the year of the tax scheme change for firms that change tax scheme, and 2014 for those that do not.

Figure 4: Effect of the hypothetical reform on firm growth



This graph plots the average effect of the hypothetical policy reform described in section 6 on firm growth for each decile of the firm size distribution. Effects are presented in percentage points and obtained using 2016 data. See Appendix E for a description of the method used.

Table 1: Firm-level descriptive statistics

	Simplified scheme	VAT scheme (small)	VAT scheme (large)
Turnover	1843	1591	117,133
	(1393)	(1367)	(1,321,115)
In Kolkata	0.18	0.27	0.38
Share VAT sales	0.01	0.28	0.33
	(0.07)	(0.38)	(0.36)
Share VAT inputs	0.48	0.59	0.81
	(0.42)	(0.44)	(0.33)
Has a VAT client	0.03	0.45	0.76
Has a VAT supplier	0.65	0.69	0.92
Number VAT clients	1.20	3.07	17.31
	(0.72)	(3.17)	(36.30)
Number VAT suppliers	2.73	3.63	12.56
	(2.26)	(3.61)	(19.89)
Number of firms	18,176	106,447	53,388
Observations	86,708	417,660	314,497

Mean (standard deviation). Column 1 includes all firms in the simplified tax scheme, column 2 all firms in the VAT scheme with a turnover under 5 million INR, column 3 all firms in the VAT scheme with a turnover over 5 million INR. The last two rows (number of VAT trading partners) are conditional on the firm having at least one VAT client or supplier. The variable "share VAT sales" is the ratio of total sales to VAT firms reported in the transaction data to total sales reported by the firm in the firm data, the variable "share VAT inputs" is the ratio of total purchases from VAT firms in the transaction data to total intermediate input purchases reported by the firm in the firm data. "Number of VAT clients" and "Number of VAT suppliers" are conditional on being greater than zero. Period: fiscal year 2010-2011. Turnover is in 1000 INR.

Table 2: Correlation between a firm's tax scheme and its use of VAT trading partners

	Dependent variable: In VAT scheme				
	(1)	(2)	(3)	(4)	(5)
Share VAT sales	0.229***	0.200***	0.168***	0.133***	0.140***
	(0.026)	(0.027)	(0.021)	(0.019)	(0.020)
Share VAT inputs	0.078***	0.089***	0.084***	0.088***	0.079***
	(0.017)	(0.014)	(0.014)	(0.014)	(0.013)
Product FE		x	x		
Location FE			x		
Product x Location FE				x	x
Firm size					x
Observations	640,634	640,634	640,634	640,634	571,498

Standard errors in parentheses are clustered at the level of the product sold by the firm and the location of the firm. The dependent variable is an indicator equal to 1 if firm i is in the VAT scheme in year t, 0 if it is in the simplified tax scheme. Each column presents estimates from a regression of this indicator variable on the share of firm i's sales that are sold to VAT clients and the share of firm i's intermediate inputs purchased from VAT suppliers in year t, as well as product fixed effects (columns 2 and 3), location fixed effects (column 3), product $\times$ location fixed effects (column 4 and 5) and firm size (turnover, column 5). The sample includes all firms with a turnover of less than 7 million INR over the fiscal years 2010-2011 to 2015-2016, all specifications include year fixed effects.Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 3: Descriptive statistics on the sample of pairs

VAT rate on transaction:	Reduced tax rate	Normal tax rate
Pair characteristics		
Positive trade	0.611	0.551
Share of trade in client's inputs	0.122 (0.207)	0.155 (0.251)
Share of trade in supplier's sales	0.005 (0.020)	0.005 (0.018)
Client characteristics		
Number of suppliers	5.7 (9.22)	5.31 (8.57)
Turnover	3,096 (1,570)	3,077(1,041)
In VAT scheme	0.919	0.799
Number of firms	79,286	39,123
Supplier characteristics		
Number of clients	83.85 (117.03)	133.86 (226.10)
Turnover	637,238 (4,500,625)	965,276 (2,711,895)
Number of firms	21,105	6,430
Number of pairs	408,551	98,236
Number of observations	2,062,269	505,504

Mean (standard deviation). This table presents descriptive statistics on the sample of annual transactions used to estimate supply chain distortions following specification (13). There is one observation per pair$*$year as long as both firms are in our data. Turnover is in thousand INR, the variable 'Positive trade' is an indicator equal to 1 when we observe a transaction between the two firms.

Table 4: Results on supply chain distortions

	(1)	(2)	(3)	(4)	(5)	(6)
Supplier tax rate t_j:	Reduced tax			Normal tax		
Outcome variable: Log input share s_{jkt}						
VAT client	0.138***	0.108***	0.288***	0.196***	0.167***	0.541***
	(0.015)	(0.014)	(0.038)	(0.029)	(0.028)	(0.062)
VAT client $*$ VAT input share			−0.304***			−0.589***
			(0.061)			(0.065)
p. value of $\beta_1 + \beta_2 = 0$			0.584			0.126
$\rho \approx 1 + \beta/t_j$			6.5			4.8
			(0.722)			(0.434)
Supplier$*$Year FE		x	x		x	x
Observations	2,062,269	2,062,269	2,062,269	505,504	505,504	505,504

Standard errors in parentheses are two-way clustered at the level of the product sold by the client and the location of the client, standard errors for ρ are obtained using the Delta method.The sample includes all pairs that trade at least once during the period in which the supplier is never eligible to the simplified scheme (minimum turnover greater than 7 million), and the client is always eligible to the scheme (minimum turnover lower than 4 million). In columns 1 to 3 we consider pairs in which the supplier pays the medium VAT rate, in columns 4 to 6 pairs in which the supplier pays the high VAT rate. The variable 'VAT client' is equal to 1 if the client is in the VAT scheme, 0 otherwise. The variable 'VAT input share' is equal to the share of the client's intermediate inputs purchased from VAT suppliers when the client is in the simplified scheme. All specifications include pair fixed effects and control for the size of the client, in columns 1 and 4 we include year fixed effects and in columns 2,3, 5 and 6 year x supplier fixed effects. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5: Results on strategic complementarities in tax choices

	Outcome: In VAT Scheme					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV	OLS	OLS	IV
Sample : turnover less than	7 million	7 million	7 million	4 million	4 million	4 million
Weighted share VAT inputs	0.616***		0.887***	0.576***		0.814***
	(0.078)		(0.198)	(0.072)		(0.192)
Weighted share VAT sales	0.207***		0.154**	0.215***		0.144**
	(0.027)		(0.064)	(0.029)		(0.067)
Predicted share VAT inputs		0.542***			0.503***	
		(0.122)			(0.119)	
Predicted share VAT sales		0.121**			0.111**	
		(0.051)			(0.054)	
Turnover control				x	x	x
Observations	640,634	640,634	640,634	571,498	571,498	571,498

Standard errors in parentheses are two-way clustered at the product and location level. The sample includes all firms with a minimum turnover over the period of less than 7 million INR in columns 1 to 3, and all firms with a minimum turnover of less than 4 million INR in columns 4 to 6. The dependent variable is an indicator equal to 1 if the firm is in the VAT scheme, 0 otherwise. The right-hand side variables 'weighted share VAT inputs' and 'weighted share VAT sales' are the variables $t_i \sum_k \lambda_{ikt} v_{kt}$ and $\sum_j t_j s_{jit} v_{jt}$: the share of the firm's intermediate inputs purchased from VAT-paying suppliers, weighted by each supplier's VAT rate, and the share of the firm's sales sold to VAT-paying clients multiplied by the firms' own VAT rates. In columns 3 and 6 these variables are instrumented for using the instruments 'predicted share VAT sales' ($t_i \sum_k \tilde{\lambda}_{ik} v_{k0} e_{kt}$) and 'predicted share VAT inputs' ($\sum_j t_j \tilde{s}_{ji} v_{j0} e_{jt}$). Sanderson and Windmeijer (2016) conditional F stats are 850.0 and 251.8 in column 3 and 930.5 and 247.7 in column 6. See the text for a description of these variables. All specifications include firm and year$x$product fixed effects, and in columns 3 to 6 a control for the firm size. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

For Online Publication: Appendix to "Taxation and Supplier Networks"

A Theory Appendix

A.1 Proof of Proposition 1

Step 1: Getting an expression of s_{jk} as a function of the no-tax-world $\tilde{s}_{jk}$
We start by: $p_{jk}q_{jk} = \alpha_{jk}^{\rho} p_{jk}^{1-\rho} P_k^{\rho} q_{kF} \phi_k^{-1}$ and by:

$$s_{jk} = \frac{p_{jk}q_{jk}}{\sum_j p_{jk}q_{jk}}$$

Note that all the terms in k that do not depend on j are going to be eliminated in the ratio. Opening up P_j and eliminating the terms that do not depend on j, we get an expression of s_{jk} that depends only on structural parameters and tax rates/statuses:

$$s_{jk} = \frac{\gamma_{jk}^{\rho-1}\phi_j^{\rho-1}\alpha_{jk}^{\rho}\alpha_{lj}^{1-\rho}}{\sum_j \gamma_{jk}^{\rho-1}\phi_j^{\rho-1}\alpha_{jk}^{\rho}\alpha_{lj}^{1-\rho}}$$

We introduce the input share $\tilde{s}_{jk}$ that prevails in the absence of taxes:

$$\tilde{s}_{jk} = \frac{\phi_j^{\rho-1}\alpha_{jk}^{\rho}\alpha_{lj}^{1-\rho}}{\sum_j \phi_j^{\rho-1}\alpha_{jk}^{\rho}\alpha_{lj}^{1-\rho}}$$

and note that:

$$s_{jk} = \frac{\gamma_{jk}^{\rho-1}\tilde{s}_{jk}}{\sum_j \gamma_{jk}^{\rho-1}\tilde{s}_{jk}}$$

Step 2: Taking the differences between $v_k = 1$ and $v_k = 0$
Under the assumption that the tax rate on simplified firms is negligible, we have $s_{jk}(1) = s_{jk}(v_k = 1) = \tilde{s}_{jk}$. The input share of a simplified client is:

$$s_{jk}(0) = \frac{\tilde{s}_{jk}(v_j(1-t_j)^{\rho-1} + (1-v_j))}{\sum_j \tilde{s}_{jk}(v_j(1-t_j)^{\rho-1} + (1-v_j))}$$

Define the share of VAT suppliers in the no-tax world when k is in the VAT regime as

$s_{Vk}(1) = \sum_j \tilde{s}_{jk} v_j$. When k is in the simplified regime, this is:

$$\begin{aligned} s_{Vk}(0) &= \frac{\sum_j \tilde{s}_{jk} v_j (1-t_j)^{\rho-1}}{\sum_j \tilde{s}_{jk}(v_j(1-t_j)^{\rho-1} + (1-v_j))} \\ &= \frac{(1-\bar{t}_k)^{\rho-1} s_{Vk}(1)}{(1-\bar{t}_k)^{\rho-1} s_{Vk}(1) + 1 - s_{Vk}(1)} \end{aligned}$$

Inverting this relationship, we can get an expression of $s_{Vk}(1)$ as a function of $s_{Vk}(0)$. Then, we get:

$$\frac{s_{Vk}(0)}{s_{Vk}(1)} = (1-\bar{t}_k)^{\rho-1} + s_{Vk}(0)[1-(1-\bar{t}_k)^{\rho-1}]$$

Now, going back to the expression of $s_{jk}(0)$, and dividing it by $s_{jk}(1) = \tilde{s}_{jk}$:

$$\frac{s_{jk}(0)}{s_{jk}(1)} = \frac{v_j(1-t_j)^{\rho-1} + (1-v_j)}{\sum_j \tilde{s}_{jk}(v_j(1-t_j)^{\rho-1} + (1-v_j))} \tag{15}$$

From the expression of $s_{Vk}(0)$ above, the denominator of this expression is equal to $\frac{(1-\bar{t}_k)^{\rho-1} s_{Vk}(1)}{s_{Vk}(0)}$, which can be replaced using the expression of the ratio above:

$$\frac{s_{jk}(0)}{s_{jk}(1)} = \left(v_j(1-t_j)^{\rho-1} + (1-v_j)\right)\left(1 + s_{Vk}(0)\left[\frac{1}{(1-\bar{t}_k)^{\rho-1}} - 1\right]\right)$$

Using a linear approximation valid for $t_j \ll 1$, we have:

$$\frac{s_{jk}(0)}{s_{jk}(1)} = \left(1-(\rho-1)t_j v_j\right)\left(1 + s_{Vk}(0)(\rho-1)\bar{t}_k\right)$$

Developing, subtracting 1, eliminating the second-order term, and taking the opposite:

$$\log(s_{jk}(1)) - \log(s_{jk}(0)) = (\rho-1)(t_j v_j - \bar{t}_k s_{Vk}(0)).$$

A.2 Proof of Proposition 2 – tax scheme choice with no size restriction

Start by assuming that firms can choose to pay taxes under the simplified scheme regardless of their total sales. The profits of upstream firm j and downstream firm k are given by:

$$\Pi_j = \kappa_{jF}\gamma_{jF}^{\sigma} + \sum_k \kappa_{jk} r_k^{\rho-\sigma}\gamma_{jk}^{\rho}\gamma_{kF}^{\sigma}$$
$$\Pi_k = \kappa_k \frac{\gamma_{kF}^{\sigma}}{r_k^{\sigma-1}}$$

with:

$$\kappa_{jF} = E\beta_j^{\sigma}\left(\frac{P_F}{\alpha_{\ell j} w\mu}\right)^{\sigma-1}\sigma^{-1}\phi_j^{\sigma-1}$$
$$\kappa_{jk} = P_F^{\sigma-1}E\left(\frac{\beta_k}{\mu}\right)^{\sigma}\alpha_{jk}^{\rho}(\nu P_j)^{1-\rho}\tilde{P}_k^{\rho-\sigma}\rho^{-1}\phi_j^{\rho-1}$$
$$\kappa_k = \sigma^{-1}P_F^{\sigma-1}E\beta_k^{\sigma}\mu^{1-\sigma}\tilde{P}_k^{1-\sigma}\phi^{\sigma-1}$$
$$r_k = \frac{P_k}{\tilde{P}_k} \approx 1 + s_{Vk}\bar{t}_k.$$

For downstream firms, we have:

$$\frac{\Pi_k^V}{\Pi_k^S} = \frac{(1-t_k)^{\sigma}}{(1-s_{Vk}\bar{t}_k)^{\sigma-1}}$$

For upstream firms, we define $\kappa_{jS} = \sum_k \kappa_{jk}(1-v_k)$, and $\kappa_{jV} = \sum_k \kappa_{jk}v_k$.

$$\Pi_j^V = \kappa_{jF}(1-t_j)^{\sigma} + \kappa_{jV}(1-\bar{t}_j)^{\sigma-1} + \kappa_{jS}\bar{r}_j^{\rho-\sigma}(1-t_j)^{\rho}$$
$$\Pi_j^S = \kappa_{jF} + \kappa_{jV}(1-\bar{t}_j)^{\sigma-1} + \kappa_{jS}\bar{r}_j^{\rho-\sigma}$$

Therefore:

$$\Pi_j^V - \Pi_j^S = -\kappa_{jF}[1-(1-t_j)^{\sigma}] - \kappa_{jS}\bar{r}_j^{\rho-\sigma}[1-(1-t_j)^{\rho}]$$

The first two results of the Proposition come directly out of these expressions. For the third one, we need to link the expression for the upstream firms with the share of VAT clients in

the no-tax world. We have $x_{jk} = \kappa_{jk}\rho r_k^{\rho-\sigma}\gamma_{jk}^{\rho-1}\gamma_{kF}^{\sigma}$, so that the share of VAT clients is:

$$\lambda_{jV} = \frac{\sum_k x_{jk} v_k}{x_{jF} + \sum_k x_{jk}} = \frac{\kappa_{jV}(1-\bar{t}_j)^{\sigma}}{\frac{\sigma}{\rho}\kappa_{jF}(1-t_j)^{\sigma-1}\kappa_{jV}(1-\bar{t}_j)^{\sigma} + \kappa_{jS}\bar{r}_j^{\rho-\sigma}(1-v_j t_j)^{\rho-1}}$$

And the share of potential VAT clients is:

$$\tilde{\lambda}_{jV} = \frac{\kappa_{jV}}{\frac{\sigma}{\rho}\kappa_{jF} + \kappa_{jV} + \kappa_{jS}}$$

From there:

$$\Pi_j^V - \Pi_j^S = (\tilde{x}_{jF} + \sum_k \tilde{x}_{jk}) \left[-\tilde{\lambda}_{jF}\sigma^{-1}[1-(1-t_j)^{\sigma}] - \tilde{\lambda}_{jS}\rho^{-1}\bar{r}_j^{\rho-\sigma}[1-(1-t_j)^{\rho}]\right]$$

Take the first approximation of $1-(1-t_j)^{\rho}$ and $1-(1-t_j)^{\sigma}$:

$$\Pi_j^V - \Pi_j^S = (\tilde{x}_{jF} + \sum_k \tilde{x}_{jk}) \left[-\tilde{\lambda}_{jF} t_j - \tilde{\lambda}_{jS} t_j \bar{r}_j^{\rho-\sigma}\right]$$

Consider an increase in $\tilde{\lambda}_{jV}$ that reduces either $\tilde{\lambda}_{jS}$, $\tilde{\lambda}_{jF}$, or both. For instance, consider $\frac{\partial \lambda_{jS}}{\partial \lambda_{jV}} = -\theta$ and $\frac{\partial \lambda_{jF}}{\partial \lambda_{jV}} = \theta - 1$, where $\theta \in (0,1)$. Then:

$$\frac{\partial(\Pi_j^V - \Pi_j^S)}{\partial \lambda_{jV}} = (\tilde{x}_{jF} + \sum_k \tilde{x}_{jk}) \left[(1-\theta) t_j + \theta t_j \bar{r}_j^{\rho-\sigma}\right]$$

A.3 Proof of Proposition 2 – tax scheme choice with a size restriction

We now introduce a constraint on firms paying taxes in the simplified scheme: these firms total sales must be lower than a threshold $\bar{x}$. Bunching occurs when the firm chooses to keep its total sales constant at $\bar{x}$ while its productivity increases, to remain in the simplified regime. We write ϕ_j^b the lowest productivity level at which firm j can obtain a level of sales $\bar{x}$. We write Π_j^b the profit of a bunching upstream firm j:

$$\Pi_j^b = \bar{x}_{jF}(1 - \frac{P_j}{\phi_j p_{jF}}) + \sum_k \bar{x}_{jk}(1 - \frac{P_j}{\phi_j p_{jk}})$$

From the expressions (6) and (7), taken for the case where j is in the simplified regime, we have $p_{jk} = p_{jF}\mu\nu^{-1}$ and:

$$\Pi_j^b = \bar{x}_{jF}(1 - \frac{P_j}{\phi_j p_{jF}}) + \sum_k \bar{x}_{jk}(1 - \frac{P_j \mu}{\phi_j p_{jF} \nu})$$

Where the $\tilde{}$ terms indicate values in the no-tax world. Defining $\tilde{\lambda}_{jV}, \tilde{\lambda}_{jS}, \tilde{\lambda}_{jF}$ the share of sales that firm j makes to VAT, simplified, and final clients in the no-tax world and $\tilde{x}_j(\phi_j)$ the total sales of firm j with productivity ϕ_j in a no-tax world, we can write:

$$\begin{aligned}
\sum_k \bar{x}_{jk} 1\{v_k = 1\} &= \tilde{\lambda}_{jV}(1 - \bar{t}_j)^\sigma \tilde{x}_j(\phi_j^b) \\
\sum_k \bar{x}_{jk} 1\{v_k = 0\} &= \tilde{\lambda}_{jS} \bar{r}_j^{\rho-\sigma} \tilde{x}_j(\phi_j^b) \\
\bar{x}_{jF} &= \tilde{\lambda}_{jF} \tilde{x}_j(\phi_j^b)
\end{aligned}$$

So that:

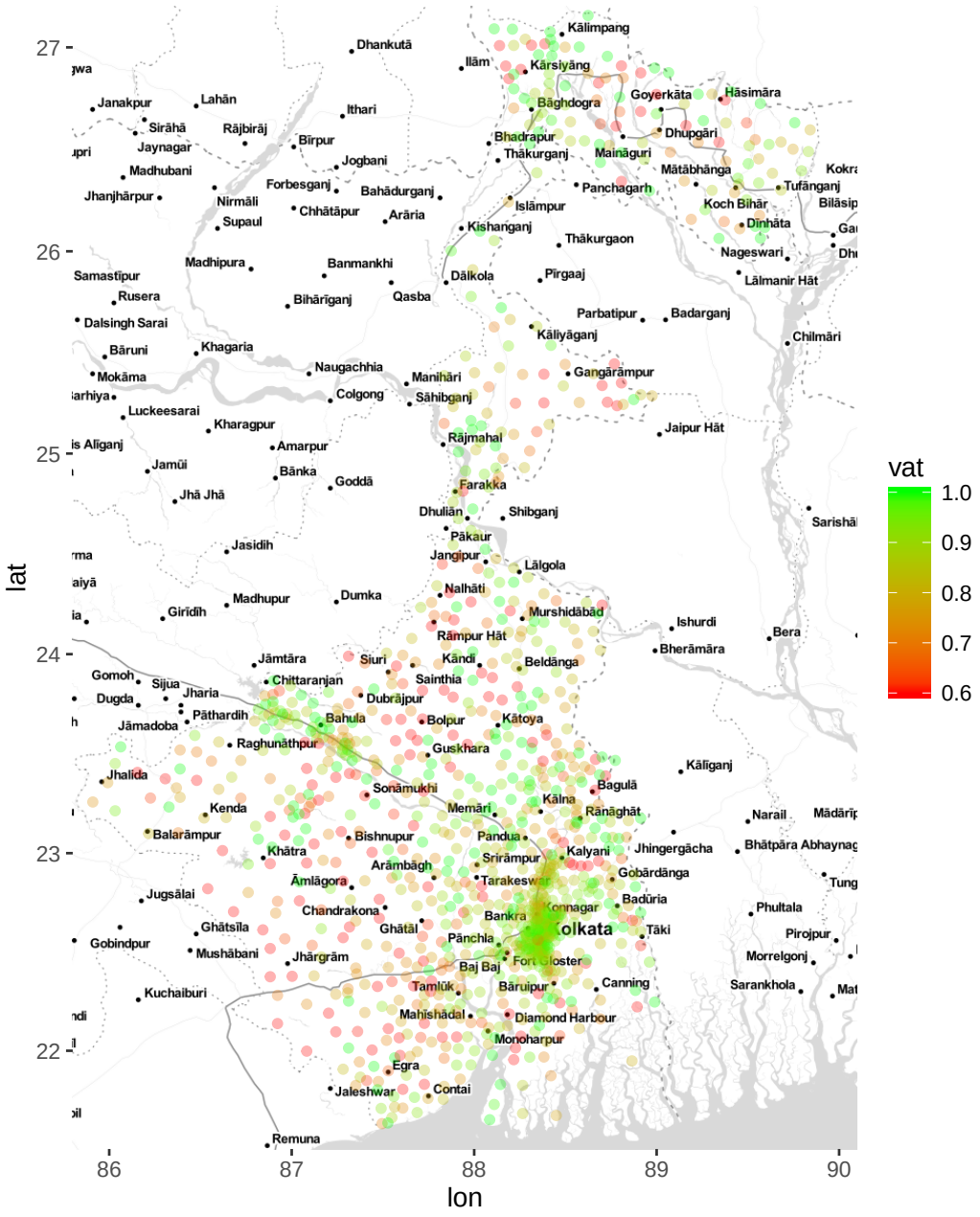
$$\Pi_j^b = \tilde{x}_j(\phi_j^b)[\tilde{\lambda}_F(1 - \frac{P_j}{\phi_j p_{jF}}) + \tilde{\lambda}_V(1 - \frac{P_j}{\phi_j p_{jF}}\frac{\mu}{\nu})(1 - \bar{t}_j)^\sigma + \tilde{\lambda}_S(1 - \frac{P_j}{\phi_j p_{jF}}\frac{\mu}{\nu})\bar{r}_j^{\rho-\sigma}]$$

Note that: $\bar{x}_{jF} = \tilde{\lambda}_{jF}\tilde{x}_j(\phi_j^b)$, and $p_{jF} = P_F E^{\mu-1} \beta_j^\mu x_{jF}^{1-\mu}$, so that $p_{jF} = P_F E^{\mu-1} \beta_j^\mu (\tilde{\lambda}_{jF}\tilde{x}_j(\phi_j^b))^{1-\mu}$. This expression does not depend on ϕ_j or any parameter of the tax system, from there it is easy to show that p_{jF} will decrease when $\tilde{\lambda}_{jF}$ increases.

B Context and data

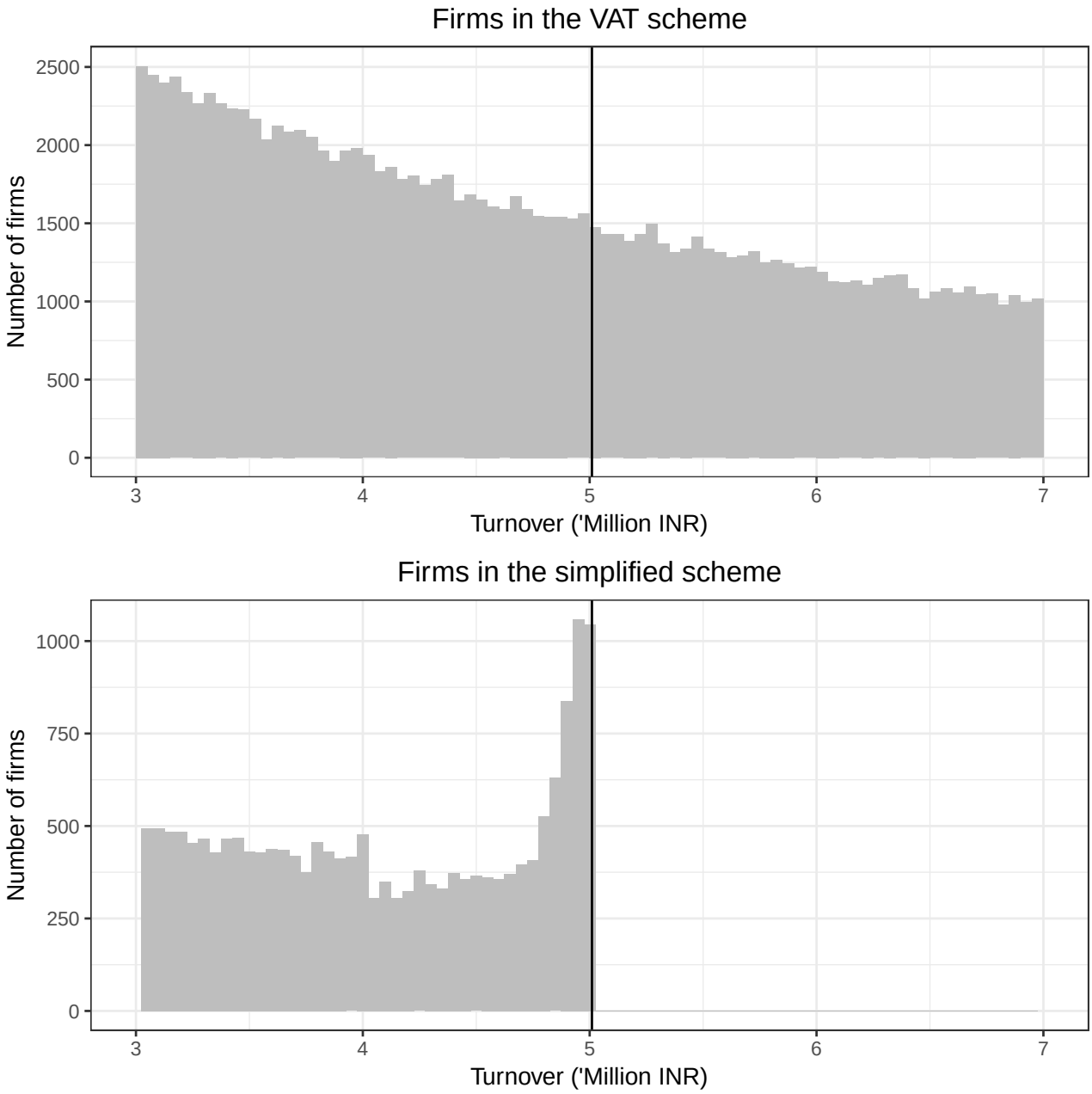
Figure B.2 plots the distribution of firms around the 5 million INR sales threshold above which firms have to pay taxes under the VAT scheme. Figure B.1 plots the location of firms in our data on a map of West Bengal. Each dot represents a postcode, the color of the dot indicates the share of firms in that postcode that are in the VAT scheme. State frontiers are indicated by the grey dotted lines. Table B.1 presents the types of products sold by firms as well as the average size and the share of VAT-paying firms among firms selling each product.

Figure B.1: Share of firms in the VAT scheme by location



Each dot represents a postcode in which firms in our data are located, the color of the dot represents the share of firms in the postcode that pay taxes under the VAT scheme. The dotted lines are state borders.

Figure B.2: Distribution of firms around the simplified tax threshold



This graph plots the density distribution of firms by turnover size around the 5 million INR threshold above which firms can no longer opt to be in the simplified tax scheme. This threshold is indicated by the black vertical lines. The top graph plots the distribution for firms paying taxes under the VAT scheme, the bottom graph the distribution for firms paying taxes under the simplified scheme.

Table B.1: Products sold and firm tax scheme in 2010-2011

Commodity type	Turnover	% Large firms	Amongst small, % VAT	Nb firms
Machines & equipment	28,285 (422,356)	32.05	91.32	19510
Construction materials	12,137 (153,641)	24.69	79.19	16911
Electrical & electronic goods	33,474 (833,588)	33.62	81.21	15560
Food, drink & tobacco	40,277 (531,480)	40.82	73.61	14828
Chemical products	41,336 (977,889)	37.05	76.09	11107
Textiles	24,235 (170,384)	31.61	72.03	10969
Metal products	109,361 (781,319)	54.46	94.07	10739
Wood & paper	20,826 (140,983)	29.00	90.6	9417
Other commodities	60,963 (1,097,142)	27.57	88.78	8479
Rubber & plastic	44,919 (1,095,713)	34.48	87.42	4672
Household goods	9,656 (90,727)	17.86	77.06	3444
Mining & energy	72,134 (1,042,568)	52.29	89.95	2689
All	38,376 (657,094)	34.17	82.72	128325

This table presents descriptive statistics by type of main product sold by firms, where we have classified over 170 different product types into 13 large product categories. The first column presents the share of firms with a turnover of over 5 million INR, the second column the share of VAT firms amongst firms with a turnover of less than 5 million INR and the third column the total number of firms in that category in 2010-2011. Categories are ranked from the one with the largest number of firms (Machines & equipment) to the one with the lowest number of firms (Mining & energy) in 2010-2011.

B.1 Additional evidence on asymmetric under-reporting

This sub-section presents several pieces of evidence that suggest that under-reporting by VAT-paying firms of sales to firms in the simplified scheme is unlikely to be a major concern. First, we find no evidence that firms are less likely to report sales to non-VAT clients than purchases from non-VAT suppliers, despite the fact that only sales can potentially increase their tax liabilities. On the contrary, firms are a lot more likely to report a client in the simplified scheme than to report a supplier in the simplified scheme: we find that 6% of VAT firms report at least one non-VAT client, less than 1% report a non-VAT supplier. See also Table 1 in the paper. We hypothesize that this is because firms in the simplified scheme are often located downstream in supply chains and therefore substantially more likely to buy from other firms than to sell to other firms – see Table B.1.

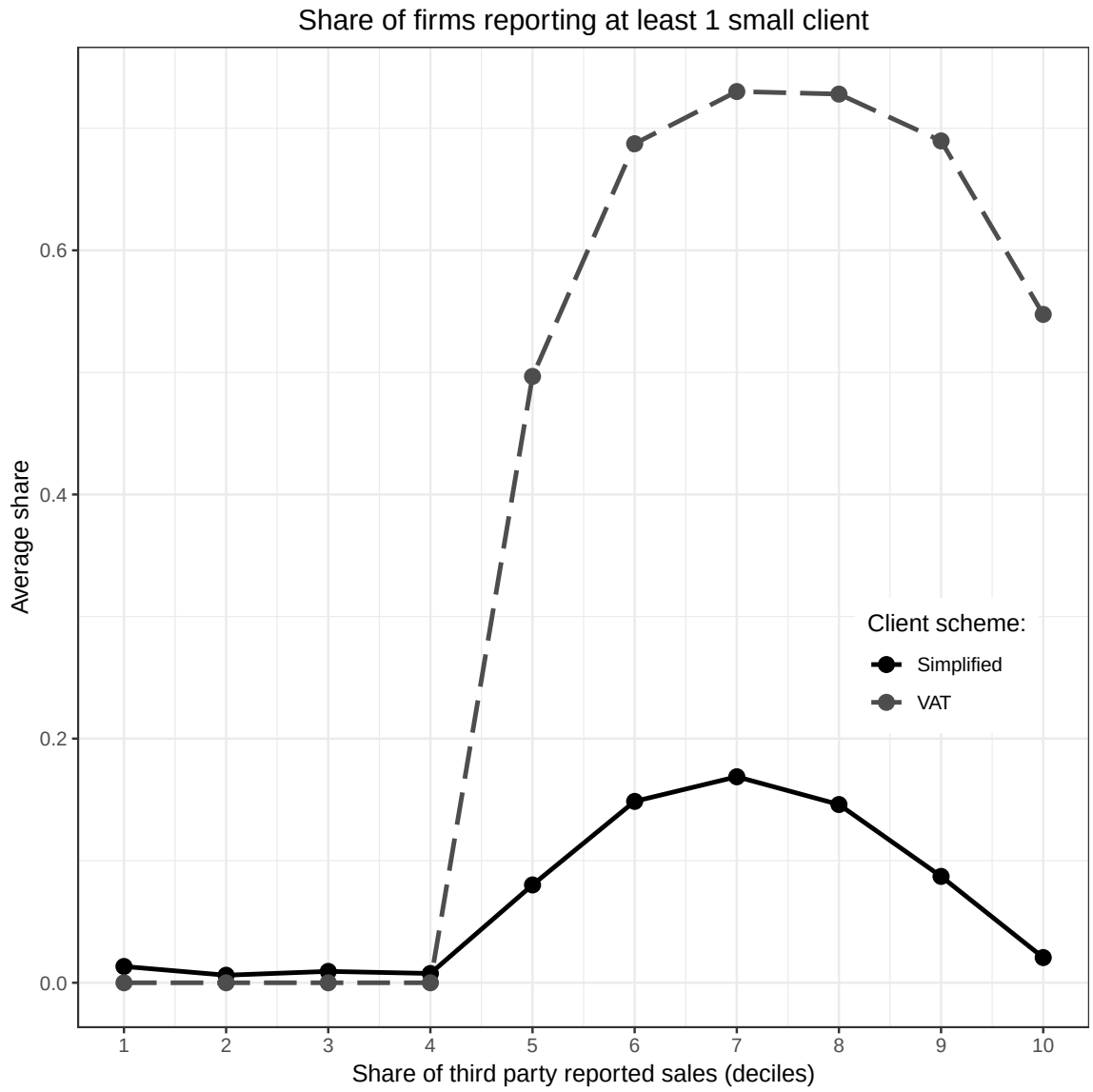
Second, we decompose firms' sales into 'third-party-reported sales' and 'voluntarily reported sales'. We define a sale as 'third-party-reported' if both the seller and the buyer involved in the transaction report the sale, i.e., both pay VAT. Declaring sales to clients in the simplified scheme only increases firms' tax liabilities if these sales are larger than the 'voluntarily reported sales' that firms report regardless. Table B.2 shows that firms report total sales that are on average three times larger than their total third-party-reported sales. This is true both for the entire sample of firms in the VAT scheme (panel A) and when we restrict the sample to firms with a turnover of less than 7 million (panel B), the sample of firms we focus on in the paper. Reporting non-VAT clients truthfully will therefore not increase the liabilities of the average firm.

Third, we find that the share of third-party reported sales is not negatively correlated with the share of sales to clients in the simplified scheme. If firms with a smaller 'voluntarily reported sales' share were less willing to truthfully report sales to these clients because it will increase their tax liabilities, we would see such a negative correlation. In Table B.2 we see that, on the contrary, the share of third-party-reported sales among total sales is slightly larger for firms that report at least one transaction with a client in the simplified scheme than for those that do not. Again, this is also true when we focus on firms with a turnover of less than 7 million.

To investigate this further, we plot the average share of firms reporting at least one client with a turnover of less than 7 million INR for each decile of the share of third-party-reported sales. Figure B.3 plots this for all firms in the VAT scheme, and Figure B.4 for firms in the VAT scheme with a turnover of less than 7 million. The black full line plots the share of firms reporting at least one client in the simplified scheme, the grey dashed line plots the share of firms reporting at least one small client in the VAT scheme.

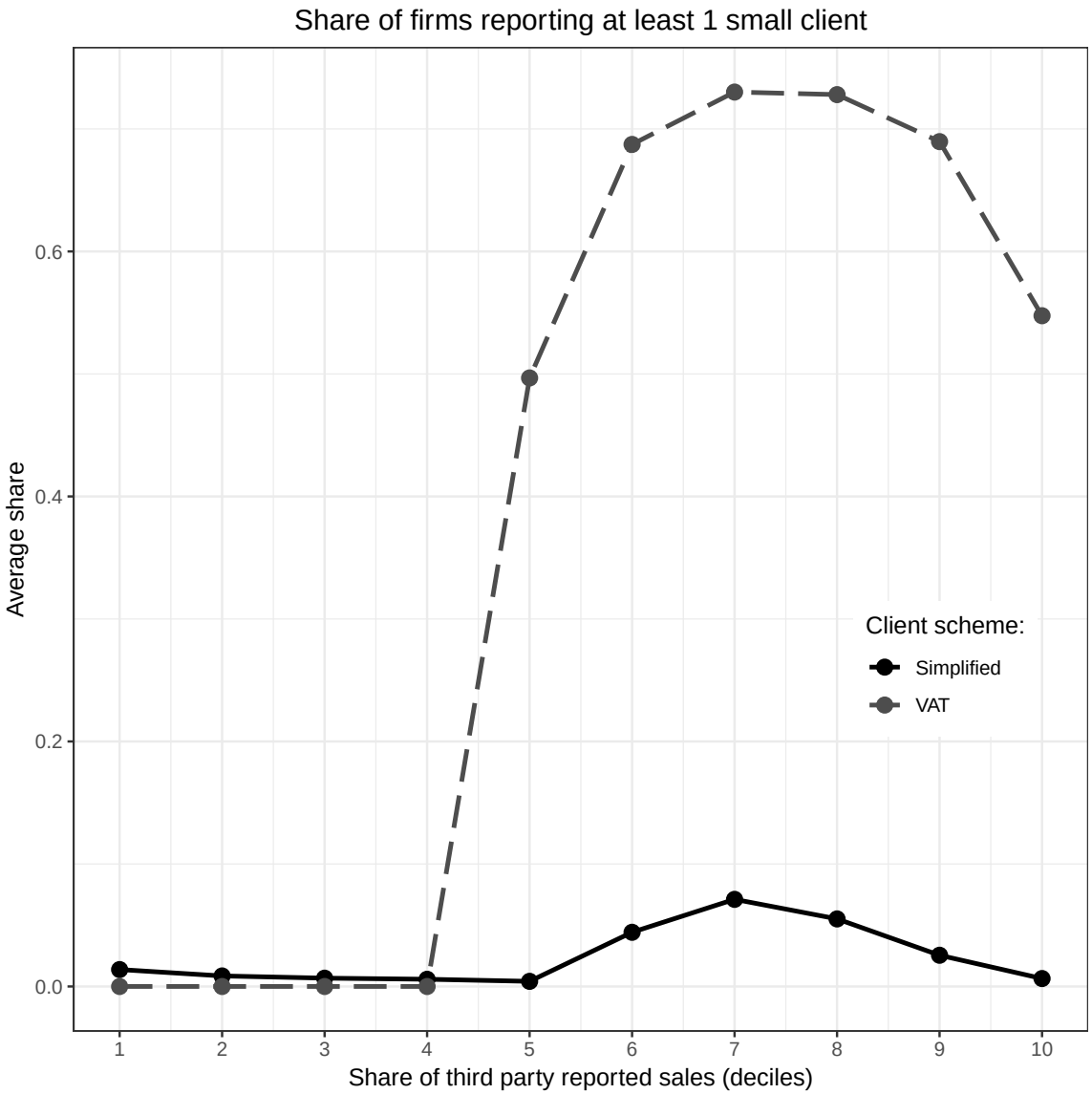
If firms were only willing to report sales to clients in the simplified scheme when reporting this transaction has no impact on their total sales, we would see a steep decline in the share of firms reporting clients in the simplified scheme as the share of third-party-reported sales increases. We would not however expect to see such a decline in the share of firms reporting a small VAT-paying client, as transactions with small VAT-paying clients are third-party reported. We see that this is not the case: firms with high shares of third-party-reported sales are less likely to report trading with small clients, regardless of the tax status of the client.

Figure B.3: Sales to simplified scheme clients and of third party reported sales: all firms



This graph plots the share of firms trading with at least one client with a turnover of less than 5 million INR for each decile of the distribution of the share of firms' sales that are third party reported. The sample of firms considered contains only firms in the VAT scheme. The dashed line plots the share of firms trading with at least one client in the VAT scheme, the unbroken line the share of firms trading with at least one client in the simplified scheme. 40% of firms have no third-party reported sales so the share of firms with a VAT client is zero by definition for the first four deciles.

Figure B.4: Sales to simplified scheme clients and third party reported sales: small firms only



This graph plots the share of firms trading with at least one client with a turnover of less than 5 million INR for each decile of the distribution of the share of firms' sales that are third party reported. The sample of firms considered contains only firms in the VAT scheme with a turnover of less than 7 million INR. The dashed line plots the share of firms trading with at least one client in the VAT scheme, the unbroken line the share of firms trading with at least one client in the simplified scheme. 50% of firms have no third-party reported sales so the share of firms with a VAT client is zero by definition for the first four deciles.

Table B.2: Share of third-party-reported sales and sales to simplified scheme clients

	Share of third-party-reported sales	Observations
A. Entire sample		
All firms	0.30 (0.38)	732,114
Firms with > 0 sales to simplified scheme clients	0.38 (0.30)	50,358
Firms with no sales to simplified scheme clients	0.29 (0.28)	681,756
B. Firms with a turnover of less than 7 million INR		
All firms	0.28 (0.38)	466,177
Firms with > 0 sales to simplified scheme clients	0.34 (0.30)	11,275
Firms with no sales to simplified scheme clients	0.28 (0.28)	454,902

Mean values, standard errors in parentheses. Panel A includes all firms in the VAT scheme, panel B all firms in the VAT scheme with a turnover of les than 7 million. In both panels the second line restricts the sample to firms with positive sales to clients in the simplified scheme in the second line and th third line to firms with no sales to clients in the simplified scheme. We define 'third-party-reported sales' as sales to clients in the VAT scheme, and repor the share of these sales in the total sales of the firm.

C Empirical strategy

C.1 Bunching evidence

Figure C.5 plots the distribution of firms by turnover around the threshold separately for firms in the VAT scheme (top graph) and firms in the simplified scheme (bottom graph).We see that all the extra mass just below the 5 million threshold comes from firms in the simplified scheme, there is no evidence of bunching among firms paying VAT. Figure C.6 plots the distribution of firms by turnover size for firms with high and low shares of sales to VAT clients separately for firms in the low and medium tax schedules, and firms in the high tax schedule. We see substantially more bunching among firms with a low share of VAT sales in the group facing a high VAT rate. Table C.3 presents descriptive statistics on the variables and the sample used to estimate strategic complementarities, see the main body of the paper for a discussion.

C.2 Proxying potential input and sales shares

We do not observe $\tilde{s}_{Vk}$, the share of input firm k buys from suppliers in the VAT regime in the no-tax counterfactual and $\tilde{\lambda}_{jV}$, the share of sales firm j sells to clients in the VAT regime in the no-tax counterfactual.

As showed in the Proof of Proposition 1, the share of input k buys from supplier j is equal to $\tilde{s}_{jk}$ with k in the VAT regime, and the the share of input firm k buys from suppliers in the VAT regime when k is in the VAT regime $s_{Vk}(1)$ is equal to $\tilde{s}_{Vk}$.

We don't have such equality for sales shares. From the proof of Proposition 2, the share of VAT clients is:

$$\lambda_{jV} = \frac{\kappa_{jV}(1-\bar{t}_j)^{\sigma}}{\frac{\sigma}{\rho}\kappa_{jF}(1-t_j)^{\sigma-1}\kappa_{jV}(1-\bar{t}_j)^{\sigma} + \kappa_{jS}\bar{r}_j^{\rho-\sigma}(1-v_j t_j)^{\rho-1}}$$

And the share of potential VAT clients is:

$$\tilde{\lambda}_{jV} = \frac{\kappa_{jV}}{\frac{\sigma}{\rho}\kappa_{jF} + \kappa_{jV} + \kappa_{jS}}$$

To simplify the rest of the reasoning, assume that there is only one VAT rate $t_j = t, \forall j$. We have:

$$\lambda_{jV}(0) = \frac{\kappa_{jV}}{\frac{\sigma}{\rho}\kappa_{jF}(1+t)^{\sigma} + \kappa_{jV} + \kappa_{jS}(1+s_V t)^{\rho-\sigma}(1+t)^{\sigma}}$$
$$\lambda_{jV}(1) = \frac{\kappa_{jV}}{\frac{\sigma}{\rho}\kappa_{jF} + \kappa_{jV} + \kappa_{jS}(1+s_V t)^{\rho-\sigma}}$$

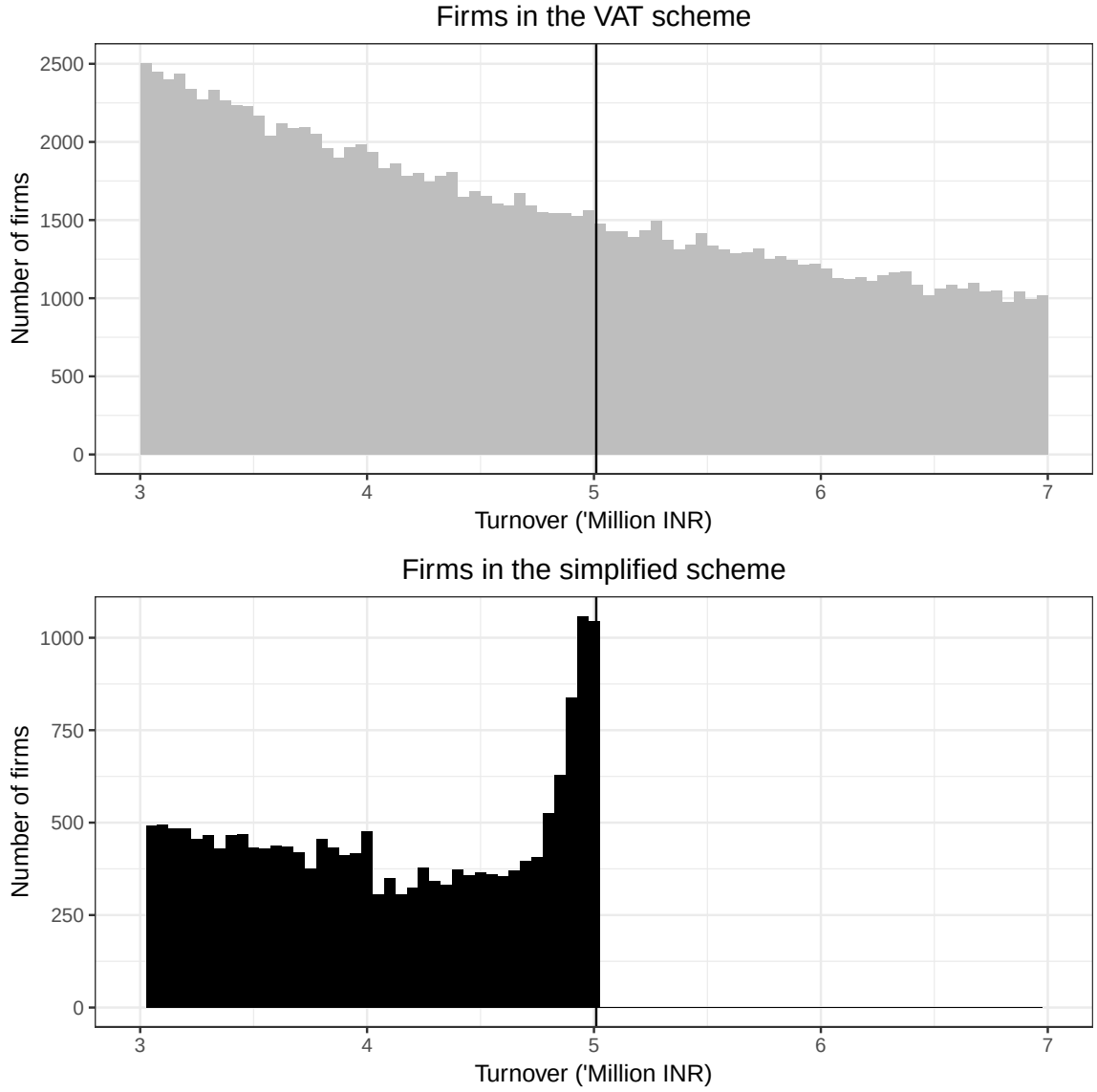
where $\lambda_{jV}(0)$ ($\lambda_{jV}(1)$) is the VAT sale share of firm j when $v_j = 0$ ($v_j = 1$). Assuming that $t << 1$, the first-order approximation of the previous expressions are:

$$\lambda_{jV}(0) \approx \tilde{\lambda}_{jV}\left(1 - t\frac{\frac{\sigma^2}{\rho}\kappa_{jF} + \sigma\kappa_{jS} + (\rho-\sigma)\kappa_{jS}s_V}{\frac{\sigma}{\rho}\kappa_{jF} + \kappa_{jV} + \kappa_{jS}}\right)$$
$$\lambda_{jV}(1) \approx \tilde{\lambda}_{jV}\left(1 - t\frac{(\rho-\sigma)\kappa_{jS}s_V}{\frac{\sigma}{\rho}\kappa_{jF} + \kappa_{jV} + \kappa_{jS}}\right)$$

From these expressions, we see that both $\lambda_{jV}(0)$ and $\lambda_{jV}(1)$ are below $\tilde{\lambda}_{jV}$ and that $\lambda_{jV}(1)$

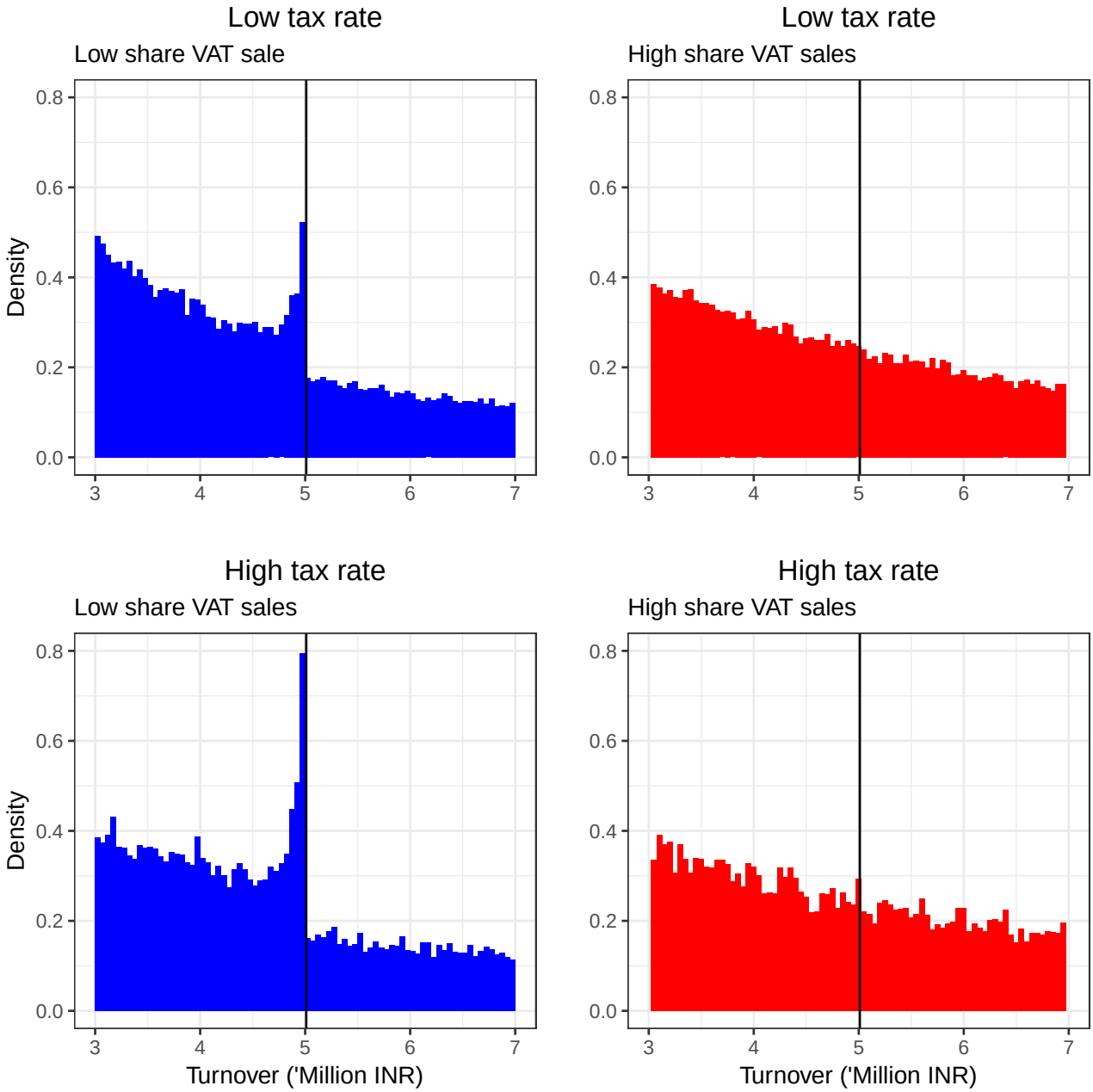
is a better proxy for $\tilde{\lambda}_{jV}$.

Figure C.5: Distribution of firm size by tax scheme



This graph plots the distribution of firms by firm size for firms in the VAT scheme (top graph) and firms in the simplified scheme (bottom graph). Bin sizes are 50,000 INR in all graphs.

Figure C.6: Distribution of firm size and tax rate



These graph plot the density distribution of firms by firm size for different sub-samples of firms with a turnover between 3 and 7 million INR. The top two graphs considers firms that sell products in the low and medium tax schedule, the bottom two firms that sell products in the high tax schedule. Graphs on the left (right) include firms whose share of sales sold to VAT-paying clients is below (above) the sample median. Bin sizes are 50,000 INR in all graphs.

Table C.3: Descriptive statistics, strategic complementarities sample

	Always VAT scheme	Always simpl. scheme	Switchers
Right-hand-side variables:			
Turnover	4077 (6337)	1745 (1322)	3564 (4844)
Weighted share VAT inputs	0.033 (0.031)	0.031 (0.037)	0.047 (0.041)
Weighted share VAT sales	0.013 (0.019)	0.001 (0.005)	0.002 (0.009)
Predicted share VAT inputs	0.036 (0.028)	0.035 (0.033)	0.051 (0.035)
Predicted share VAT sales	0.016 (0.019)	0.001 (0.004)	(0.002) (0.009)
Trading partners:			
Has a trading partner	0.92	0.80	0.92
Nb trading partners	16.9 (21.8)	5.3 (4.9)	9.2 (9.8)
Has an enter./exiting partner	0.67	0.34	0.54
Nb enter./exiting partners	4.6 (5.5)	1.8 (1.2)	2.7 (2.6)
Trading partner's turnover	400,289 (575,501)	481,510 (653,791)	563,548 (648,470)
Enter./exiting partner's turnover	129,825 (287,608)	158,574 (394,523)	200,645 (460,281)
Number of firms	121,690	13,576	7,648
Observations	535,198	66,779	38,657

Mean (standard deviation). This table presents descriptive statistics on the sample of firms used to estimate strategic complementarities in (14). Turnover is in thousand INR, the variables ' weighted share VAT inputs', 'weighted share VAT sales', 'predicted share VAT inputs' and 'predicted share VAT sales' are defined in the text. The variable 'has a trading partner' ('has an entering/exiting partner') are equal to 1 if the firm has at least one supplier or client over the period (at least one supplier or client that enters or exit over the period). The variables 'Nb trading partners' and 'Nb entering/exiting partners' are conditional on the firm having at least one partner. The last two lines present the average turnover of the firm's trading partners. The sample in the first column includes all firms that are always in the VAT scheme, in the second column all firms always in the simplified scheme, and in the third column all firms that change tax scheme over the period.

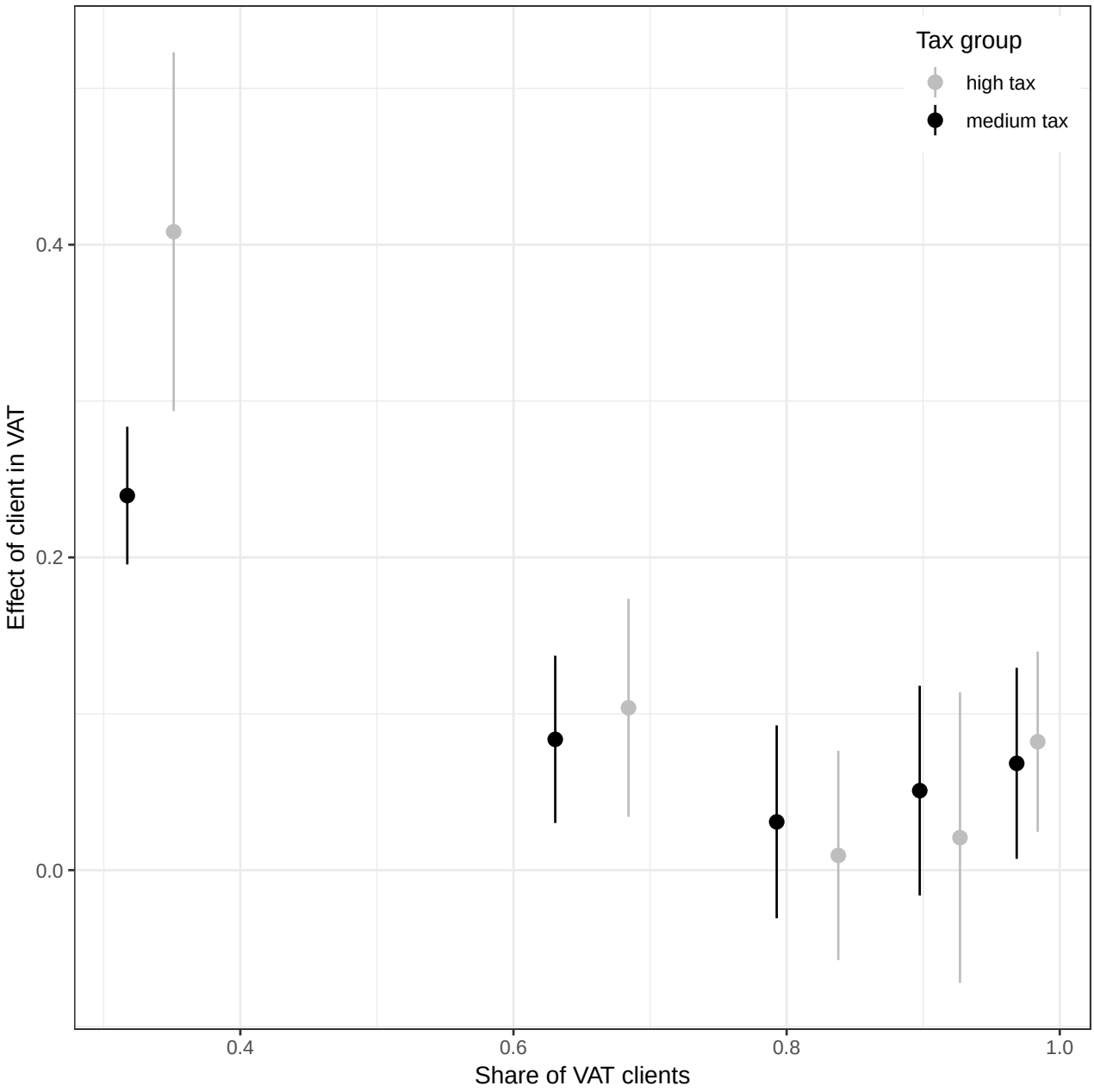
D Results

Tables D.4 and D.5 present estimates of supply chain distortions obtained on alternative samples or using alternative specifications. Figure D.7 plots the estimated coefficients obtained by running specification (13) on sub-samples of pairs. We split the sample by quintile of the distribution of s_{Vk}, the VAT share of firms' intermediate inputs. We see that the effect is decreasing with how much the client buys from VAT suppliers when in the simplified

scheme in both sub-samples. The estimates obtained on sub-samples in which the client's VAT input share is very close to one are not equal to zero, which may be due to the fact that we do not observe labor input costs. Labor costs are non VAT-deductible, so labor's cost increases relative to that of inputs purchased from VAT suppliers when clients enter the VAT scheme. Our model's logic implies that firms that buy a large share of their inputs from VAT-paying suppliers when they are in the simplified scheme (those to the left of the x-axis scale in Figure D.7) may therefore substitute away from labor when they enter the VAT scheme and buy even more inputs from VAT-paying suppliers. Our estimates are in line with this conjecture.

Table D.7 presents first stage results for our estimation of strategic complementarities and Table D.8 additional estimates of strategic complementarities obtained on alternative samples or using alternative specifications. Finally Table D.9 presents unconditional correlations between firms' propensity of being in the VAT scheme and their trade with VAT-paying firms under different counterfactual scenarios.

Figure D.7: Supply chain distortions, results by quintile of VAT input share



This graph plots the estimated coefficients obtained by running specification (13) on sub-samples of pairs. We split the sample by quintile of the distribution of s_{Vk}, the VAT share of firms' intermediate inputs. Each point represents an estimate obtained on a sub-sample among pairs in which the supplier pay the high tax rate in black, and pairs in which the supplier pay the medium tax rate in grey. The location on the x-axis corresponds to the average VAT share of inputs in the sub-sample. Bars represent the 95% confidence intervals.

Table D.4: Additional results on supply chain distortions - robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
Supplier tax rate t_j:	Medium tax			High tax		
Panel A. Excluding suppliers with large shares of third-party reported sales						
VAT client	0.138***	0.113***	0.289***	0.183***	0.147***	0.494***
	(0.015)	(0.013)	(0.038)	(0.028)	(0.025)	(0.062)
VAT client $*$ VAT input share			−0.300***			−0.554***
			(0.060)			(0.067)
p. value of $\beta_1 + \beta_2 = 0$			0.698			0.076
$\rho \approx 1 + \beta / t_j$			6.491			4.458
			(0.722)			(0.434)
Supplier$*$Year FE		x	x		x	x
Observations	1,558,605	1,558,605	1,558,605	438,607	438,607	438,607
Panel B. Adding year x location fixed effects						
VAT client	0.133***	0.107***	0.288***	0.188***	0.165***	0.542***
	(0.014)	(0.014)	(0.038)	(0.030)	(0.028)	(0.062)
VAT client $*$ VAT input share			−0.306***			−0.594***
			(0.058)			(0.067)
p. value of $\beta_1 + \beta_2 = 0$			0.492			0.096
$\rho \approx 1 + \beta / t_j$			6.472			4.794
			(0.722)			(0.434)
Supplier$*$Year FE		x	x		x	x
Observations	2,062,269	2,062,269	2,062,269	505,504	505,504	505,504
Panel C. Outcome variable: Trade > 0						
VAT client	0.138***	0.111***	0.264***	0.139***	0.126***	0.328***
	(0.007)	(0.006)	(0.010)	(0.009)	(0.008)	(0.015)
VAT client $*$ VAT input share			−0.258***			−0.318***
			(0.014)			(0.018)
p. value of $\beta_1 + \beta_2 = 0$			0.467			0.262
Supplier$*$Year FE		x	x		x	x
Observations	2,062,269	2,062,269	2,062,269	505,504	505,504	505,504

The dependent variable is the log input share s_{jkt} in Panels A and B and an indicator equal to 1 if the two firms trade in panel C. The sample in panel A excludes pairs in which the supplier sells more than 90% of its sales to other firms, the sample in panels B and C is the sample used in our baseline specification in Table 4. All specifications include pair fixed effects and control for the log turnover of the client firm k, in columns 1 and 3 we include year fixed effects and in columns 2 and 4 year x supplier fixed effects. In Panel C we also include location x year fixed effects. Standard errors in parentheses are two-way clustered at the level of the product sold by and the location of the client. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table D.5: Additional results on supply chain distortions - alternative samples

	(1)	(2)	(3)	(4)	(5)	(6)
Supplier tax rate t_j:	Medium tax			High tax		
Panel A. Including all clients eligible to the simplified scheme						
VAT client	0.154***	0.118***	0.301***	0.220***	0.188***	0.577***
	(0.014)	(0.013)	(0.035)	(0.029)	(0.027)	(0.068)
VAT client $*$ VAT input share			−0.304***			−0.612***
			(0.054)			(0.071)
p. value of $\beta_1 + \beta_2 = 0$			0.909			0.131
$\rho \approx 1 + \beta / t_j$			6.719			5.039
			(0.665)			(0.476)
Supplier$*$Year FE		x	x		x	x
Observations	2,680,746	2,680,746	2,680,746	639,931	639,931	639,931
Panel B. Restricting to very large suppliers						
VAT client	0.137***	0.108***	0.292***	0.199***	0.173***	0.568***
	(0.015)	(0.014)	(0.040)	(0.029)	(0.028)	(0.063)
VAT client $*$ VAT input share			−0.306***			−0.619***
			(0.063)			(0.065)
p. value of $\beta_1 + \beta_2 = 0$			0.627			0.101
$\rho \approx 1 + \beta / t_j$			6.548			4.976
			(0.760)			(0.441)
Supplier$*$Year FE		x	x		x	x
Observations	1,853,960	1,853,960	1,853,960	469,232	469,232	469,232
$\rho \approx 1 + \beta / t_j$		7		4.9		

The dependent variable is the log input share s_{jkt}. Relative to the sample used in our baseline specification in Table 4 the sample in panel A includes all pairs in which the client has a (minimum) turnover of less than 7 million INR, in panel B the sample includes all pairs in which the supplier has a (minimum) turnover larger than 10 million INR.All specifications include pair fixed effects and control for the log turnover of the client firm k, in columns 1 and 3 we include year fixed effects and in columns 2 and 4 year x supplier fixed effects.Standard errors in parentheses are two-way clustered at the level of the product sold by and the location of the client. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table D.6: Additional results on supply chain distortions - heterogeneous effects by firm size

	(1)	(2)	(3)	(4)	(5)	(6)
Supplier tax rate t_j:	Medium tax			High tax		
Panel A. Medium-small clients (turnover in [2,4] million)						
VAT client	0.182***	0.141***	0.368***	0.259***	0.208***	0.574***
	(0.027)	(0.028)	(0.061)	(0.028)	(0.026)	(0.094)
VAT client $*$ VAT input share			−0.350***			−0.553***
			(0.077)			(0.119)
Supplier$*$Year FE		x	x		x	x
Observations	716,085	716,085	716,085	166,829	166,829	166,829
Panel B. Small clients (turnover less than 2 million)						
VAT client	0.118***	0.095***	0.273***	0.169***	0.154***	0.534***
	(0.017)	(0.015)	(0.047)	(0.032)	(0.030)	(0.062)
VAT client $*$ VAT input share			−0.313***			−0.609***
			(0.077)			(0.069)
Supplier$*$Year FE		x	x		x	x
Observations	1,346,184	1,346,184	1,346,184	338,675	338,675	338,675

The dependent variable is the log input share s_{jkt}. This table splits the sample used in our baseline specification in Table 4 in two groups: Panel A presents results for suppliers with a minimum turnover of more than 2 million, Panel B results for those with a minimum turnover of less than 2 million. All specifications include pair fixed effects and control for the log turnover of the client firm k, in columns 1 and 3 we include year fixed effects and in columns 2 and 4 year x supplier fixed effects. Standard errors in parentheses are two-way clustered at the level of the product sold by and the location of the client. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table D.7: Strategic complementarities: first stage results

	Outcome: In VAT Scheme			
	(1)	(2)	(3)	(4)
Outcome variable:	Share VAT inputs	Share VAT sales	Share VAT inputs	Share VAT sales
Turnover less than	7 million	7 million	4 million	4 million
Pred. share VAT inputs	0.610***	0.010*	0.616***	0.010
	(0.025)	(0.005)	(0.024)	(0.006)
Pred. share VAT sales	−0.002	0.797***	-0.004	0.798***
	(0.015)	(0.017)	(0.016)	(0.017)
Turnover control			x	x
Observations	640,634	640,634	571,498	571,498

Standard errors in parentheses are two-way clustered at the product and location level. This table presents the first stage results corresponding to the second stage in Table 5, columns 3 and 6. The sample includes all firms with a minimum turnover over the period of less than 7 million INR in columns 1 and 2, and all firms with a minimum turnover of less than 4 million INR in columns 3 and 4. The dependent variable is an indicator equal to 1 if the firm is in the VAT scheme, 0 otherwise. The outcome variables 'share VAT inputs' and 'share VAT sales' are the variables $t_i \sum_k \lambda_{ikt} v_{kt}$ and $\sum_j t_j s_{jit} v_{jt}$, the right-hand-side variables are 'predicted share VAT sales' ($t_i \sum_k \tilde{\lambda_{ik}} v_{k0} e_{kt}$) and 'predicted share VAT inputs' ($\sum_j t_j \tilde{s_{ji}} v_{j0} e_{jt}$). See the text for a description of these variables. All specifications include firm and year$x$product fixed effects, and in columns 3 and 4 the firm's turnover as an additional control. Significance levels: *p<0.1; **p<0.05; ***p<0.01. .

Table D.8: Additional results on strategic complementarities

	Outcome: In VAT Scheme					
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS	(6) IV
A. Lowering sample thresholds						
Weighted share VAT inputs	0.622***		0.900***	0.570***		0.806***
	(0.078)		(0.202)	(0.071)		(0.190)
Weighted share VAT sales	0.209***		0.157**	0.216***		0.170**
	(0.028)		(0.065)	(0.027)		(0.069)
Predicted share VAT inputs		0.554***			0.500***	
		(0.126)			(0.119)	
Predicted share VAT sales		0.123**			0.132**	
		(0.052)			(0.055)	
Turnover control				x	x	x
Observations	623,990	623,990	623,990	551,738	551,738	551,738
B. Increasing sample thresholds						
Weighted share VAT inputs	0.611***		0.881***	0.588***		0.844***
	(0.077)		(0.199)	(0.073)		(0.201)
Weighted share VAT sales	0.204***		0.150**	0.214***		0.127*
	(0.027)		(0.062)	(0.027)		(0.067)
Predicted share VAT inputs		0.538***			0.520***	
		(0.123)			(0.125)	
Predicted share VAT sales		0.118**			0.098*	
		(0.050)			(0.054)	
Turnover control				x	x	x
Observations	654,851	654,851	654,851	587,610	587,610	587,610
C. With year×location fixed effects						
Weighted share VAT inputs	0.593***		0.826***	0.573***		0.801***
	(0.075)		(0.187)	(0.071)		(0.190)
Weighted share VAT sales	0.207***		0.175**	0.216***		0.140*
	(0.028)		(0.075)	(0.028)		(0.077)
Predicted share VAT inputs		0.502***			0.490***	
		(0.113)			(0.116)	
Predicted share VAT sales		0.142**			0.113*	
		(0.061)			(0.062)	
Turnover control				x	x	x
Observations	654,851	654,851	654,851	587,610	587,610	587,610

Standard errors in parentheses are two-way clustered at the product and location level. The sample in panels A, B and C include all firms with a minimum turnover over the period of less than 6,8 and 7 million INR in columns 1 to 3, and all firms with a minimum turnover of less than 3, 4.5 and 4 million INR in columns 4 to 6. The dependent variable is an indicator equal to 1 if the firm is in the VAT scheme, 0 otherwise. The right-hand side variables 'weighted share VAT inputs' and 'weighted share VAT sales' are the variables $t_i \sum_k \lambda_{ikt} v_{kt}$ and $\sum_j t_j s_{jit} v_{jt}$, in columns 3 and 6 these variables are instrumented for using the instruments 'predicted share VAT sales' ($t_i \sum_k \tilde{\lambda_{ik}} v_{k0} e_{kt}$) and 'predicted share VAT inputs' ($\sum_j t_j \tilde{s_{ji}} v_{j0} e_{jt}$). All specifications include firm and year$x$product fixed effects, and in columns 3 to 6 the firm's turnover as an additional control. Significance levels: *p<0.1; **p<0.05; ***p<0.01. .

Table D.9: Supply chain segmentation under counterfactual scenarios, no location×product fixed effects

	Outcome: Probability in VAT scheme				
	Observed	No trade distortions	No complementarities		Full counterfactual
	(1)	(2)	(3)	(4)	(5)
Share VAT inputs	0.078***	0.056***	0.062***	0.058***	0.036**
	(0.017)	(0.015)	(0.017)	(0.017)	(0.015)
Share VAT sales	0.229***	0.232***	0.228***	0.224***	0.226***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Observations	640,634	640,634	640,634	640,634	640,634

Standard errors in parentheses are two-way clustered at the product and location level. The sample includes all firms with a minimum turnover over the period of less than 7 million INR. The right-hand side variables are the share of inputs (sales) purchased from (sold to) VAT-paying firms. In column 1 we use the observed shares, see the text for a description of how the counterfactual shares in columns 3 to 5 are constructed. All specifications include year fixed effects. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

E Discussion

In this section we explain how we obtain estimates of the change in output and tax revenues that would occur if a reform made it possible for firms in the simplified scheme to deduct VAT paid on their inputs. We use the year 2016 as our baseline. We no longer assume that firms are either strictly downstream or strictly upstream but instead allow all firms to buy from other firms, sell to other firms and to final consumers. We do impose the restriction that all firms in the simplified scheme at baseline are purely downstream (ie do not sell to other firms). This simplifying assumption is motivated by the fact that 99% of the suppliers in our transaction data are in the VAT scheme. All notations are the same as in Section 3 in the paper.

E.1 Effect on output

Theoretical expressions

The change in output of firm i due to the reform can be written as:

$$\frac{\partial q_i}{q_i} = \frac{q_{iF}}{q_i}\frac{\partial q_{iF}}{q_{iF}} + \sum_{k \in K_i} \frac{q_{ik}}{q_i}\frac{\partial q_{ik}}{q_{ik}} \tag{16}$$

where K_i is the set of firm i's suppliers. The first term denotes the change in firm i's sales (in volume) to final consumers, the second the change in its sales to other firms.

The change in firms' sales to final consumers is given by:

$$\frac{\partial q_{iF}}{q_{iF}} = -\sigma\frac{\partial P_i}{P_i} + (\sigma - 1)\frac{\partial P_F}{P_F} \tag{17}$$

and the change in sales from firm i to firm k by:

$$\frac{\partial q_{ik}}{q_{ik}} = \frac{\partial q_k}{q_k} + \rho\left(\frac{\partial P_k}{P_k} + v_i(1 - v_k)t_i\right) \tag{18}$$

where the last term $\rho v_i(1 - v_k)t_i)$ corresponds to the increase in trade between VAT suppliers and clients in the simplified scheme because of the policy reform.

The change in the final price index is given by:

$$\frac{\partial P_F}{P_F} = -\sum_{i \in F} \frac{q_{iF}}{\sum_{j \in F} q_{jF}}(1 - v_i)\frac{\partial P_i}{P_i} \tag{19}$$

where F is the set of firms that sell to final consumers.

Finally, the change in firm i's input cost index is given by:

$$\frac{\partial P_i}{P_i} = -(1-v_i)\left(\tilde{s}_{vi}\bar{t}_i + \tilde{s}_{vi}\bar{t}_i^2 - \frac{\rho}{2}\tilde{s}_{vi}(1-\tilde{s}_{vi})\bar{t}_i^2\right) \approx -(1-v_i)\tilde{s}_{vi}\bar{t}_i \tag{20}$$

Calibration parameters

To calibrate these expressions we use the observed v_i, t_i and $\bar{t}_i$ for each firm. We compute the share of firms' sales sold to other firms, by combining our firm and transaction data and use it to proxy for both $\sum_{k\in K_i}\frac{q_{ik}}{q_i}$ and $\frac{q_{iF}}{q_i}$, which we assume is one minus that share (this is equivalent to assuming that firms' clients that are not tax-registered all have the same price elasticity of demand σ). We use our estimate of $\rho \approx 5$ and the average value of σ estimated by DellaVigna and Gentzkow (2019) amongst US retailers, 2.5.

Finally, we take $\tilde{s}_{Vi} = s_{Vi}$ for firms in the VAT scheme. This equality holds as long as there are no firms in the simplified scheme upstream of firms in the VAT scheme, which is approximately the case in our context. If i is in the simplified regime, we invert equation (15) to obtain an expression for $\tilde{s}_{Vi}$ as a function of parameters that we observe:

$$\tilde{s}_{Vi} = \frac{s_{Vi}}{(1-\bar{t}_i)^{\rho-1} + s_{Vi}[1-(1-\bar{t}_i)^{\rho-1}]} \tag{21}$$

Calibrated effects on prices

These elements enable us to calibrate for each firm it's change in input price index due to the reform using expression (20). This is null for all firms except for firms in the simplified tax scheme, for which P_i falls by 2.6%. From this we obtain the overall change in the price index faced by the final consumer using expression (19): the index falls by a very small percentage (0.002%) because firms in the simplified scheme represent a very small share of the final consumption basket.

Calibrated effects on output

With these two elements we then calibrate the change in q_{iF} for all firms that sell to final consumers using expression (17). This gives us for each firm $\frac{\partial \hat{q}_{iF}}{q_{iF}}$. We then compute the change in trade between firms and their suppliers k using expression (18) using the following expression:

$$\frac{\partial q_{ik}}{q_{ik}} = \frac{\partial \hat{q}_{iF}}{q_{iF}} + \rho\left(\frac{\partial P_k}{P_k} + v_i(1-v_k)t_i\right) \tag{22}$$

where $\frac{\partial \hat{q}_{iF}}{q_{iF}} = 0$ if firm i does not sell to final consumers. This change in trade leads to a change in the suppliers' output, and therefore a change in their trade with their own

suppliers j. Effects therefore spread upwards throughout the supply chain: when a client k's output changes by $\frac{\partial q_k}{q_k}$ its supplier j's output changes by $\frac{q_{jk}}{q_j}\frac{\partial q_k}{q_k}$, and this supplier's supplier l's output changes by $\frac{q_{lj}}{q_l}\frac{q_{kj}}{q_j}\frac{\partial q_k}{q_k}$, and so on. As we move up the supply chain these effects become smaller, we stop considering additional supply chain effects when they represent less than 0.1% of the total effect.

Table E.10 presents the results. We see that firms in the simplified scheme are only affected through the fact that their decrease in input costs enables them to sell at a lower price to final consumers and therefore sell more (second column, direct effect). This is because the vast majority of these firms do not sell to other firms. Firms in the VAT scheme are on the contrary hardly (and negatively) affected via their sales to final consumers - because they compete with firms in the simplified scheme whose price has fallen - , but positively through supply chain effects (third column). This is because a large share of them (47%) belong to the same supply chain as at least one firm in the simplified scheme, and therefore benefit from these firms' higher demand for inputs (directly or through their own clients).

Table E.10: Effect of the reform on firms' output, average by firm type

Sample	Total effect	Direct effect	Supply chain effect
Simplified scheme	7.55	7.55	0.00
	(8.76)	(8.76)	(0.00)
VAT scheme (small)	0.09	-0.03	0.12
	(1.14)	(0.01)	(1.14)
VAT scheme (large)	0.15	-0.02	0.18
	(0.75)	(0.01)	(0.75)

Mean (standard deviation). All numbers are in percentage points. Row 1 includes all firms in the simplified tax scheme, row 2 all firms in the VAT scheme with a turnover under 5 million INR, row 3 all firms in the VAT scheme with a turnover over 5 million INR. The first column presents the average calibrated total change in output, the second column the change in output due to a change in demand from final consumers, and the third the change in output due to all supply chain effects.

E.2 Effect on tax revenues

We abstract here from the mechanical effect of the reform: a transfer between the government and firms, which corresponds to less than 0.7% of the government's total VAT revenues. The main behavioral effect of the policy reform on tax revenues comes from changes in firms' choice of tax scheme. By letting firms deduct VAT paid on their inputs from their tax liabilities regardless of their tax scheme, the reform implies that firms with large input shares purchased from VAT suppliers no longer have an incentive to themselves pay the VAT. This leads to some firms exiting the VAT scheme, and less VAT revenues (see below for a discussion of smaller effects that will tend to have the opposite effect on revenues). To quantify this effect we use our estimate of δ_2, which captures by how much a 1 percentage point increase in the weighted share of inputs purchased from VAT suppliers increase a firm's propensity to pay VAT at baseline. This effect no longer holds after the policy reform, so firm i's new propensity to pay VAT is given by:

$$\hat{v}_i = 1 - \delta_2 \sum_{j \in J_i} t_j s_{ji} v_j \tag{23}$$

Using our estimate of δ_2 in Table 5 (0.814) and the observed $\sum t_j s_{ji} v_j$ we obtain an estimate $\hat{v}_i$ for each firm eligible to choose between the two tax schemes. The average value is 0.849, a 3.2 percentage point decrease compared to the average probability that these firms pay VAT at baseline. Note that there are no reasons to expect that the fact that some firms exit

the VAT will lead to other firms exiting the VAT here: the policy reform implies that firms no longer take into account the tax scheme of their suppliers when choosing their scheme, or the tax scheme of their corporate clients (because tax-registered clients can deduct VAT paid on their inputs from their tax liabilities regardless of their tax scheme- see below).

To quantify the change in VAT revenues from this decrease in the share of firms in the VAT scheme, we rank firms paying VAT at baseline and eligible to choose their tax scheme by their estimated new propensity to pay VAT, $\hat{v}$. The amount of VAT revenues lost is the amount of VAT paid on sales by the bottom 3.2% of these firms. This corresponds to 0.8% of the total VAT paid in 2016.

This estimate is an upper-bound on the total amount of tax revenues lost because of the reform, for three reasons. First, we ignore the fact that firms that leave the VAT scheme for the simplified scheme will still pay some taxes on their turnover, albeit a very small amount. This tax gain represents less than 3% of the amount of VAT no longer paid by these firms. Second, the small increase in the output of firms in the VAT scheme due to the reform will likely increase the amount of VAT they pay. Again, this effect will be very small because firm growth effects are concentrated amongst smaller firms, as explained above. Third, the reform will also *increase* some firms incentives to pay VAT via complementarities. Before the reform, firms in the simplified scheme that mostly sold to other firms in this scheme had an incentive not to pay VAT. This incentive is removed by the reform, because it ensures that VAT is never paid on transactions between tax-registered firms (VAT is still paid by consumers and non-tax registered corporate clients, so firms still take into account the nature of their clients when choosing their tax scheme). This will increase firms' propensity to pay VAT by an amount $\delta_1 t_i \sum_{k \in K_i} \lambda_{ik}(1 - v_k)$: firms in the simplified scheme at baseline that sell large amounts to other firms in this scheme no longer have a disincentive to pay VAT. We cannot quantify this effect in the absence of data on transactions between firms in the simplified scheme. However, we note that this effect is also likely to be very small, because complementarities effect with respect to clients' tax status are small (see Table 5) and because firms in the simplified scheme are mostly found in downstream stages of supply chains (see Table B.1) so it is unlikely that there is much trade between firms in the simplified scheme.

E.3 Effect on supply chain segmentation

Table E.11 presents results obtained from running a regression of firms' propensity to pay VAT (at baseline this is equal to 1 if the firm is in the VAT scheme, 0 otherwise) on their trade with VAT-paying clients and suppliers under different scenarios. All specifications

include location$x$product and year fixed effects. Column (1) presents baseline results (see column 4 of Table 2) and column (4) results obtained after simulating the total effect of the reform as described above. Columns (2) and (3) decompose this effect by considering first a hypothetical world in which the reform has changed firm-to-firm trade but left firms' choice of tax scheme unchanged (column 2) then a world in which trade is unchanged but firms have adjusted their tax scheme choice (column 3). We see that the reform would decrease the level of segmentation observed 'upstream' in supply chains (the correlation between the share of firms' inputs purchased from VAT suppliers and their tax scheme) by roughly 50%. The level of segmentation 'downstream' (the correlation between the share of firms' sales sold to VAT clients and their tax scheme) would also decrease, but by less - 15% - and mostly because of the change in tax scheme. The change in trade hardly affects this correlation, mostly because most firms' sales to non-VAT-paying firms are to final consumers (or non-tax registered firms), not firms in the simplified scheme. These sales are unchanged by the reform. In addition, whilst there is an increase in trade between clients in the simplified scheme and VAT-paying suppliers, it has a much smaller effect on the suppliers' sales than on the clients' inputs. This is because firms in the simplified scheme are substantially smaller than their VAT-paying suppliers.

Note that results in Table E.11, like the rest of our analysis, ignores the fact that the reform would also increase some firms' propensity to choose the VAT scheme (see discussion above). This likely leads us to under-estimate how much the reform would decrease supply chain segmentation 'downstream' in supply chains. This effect would lead to some firms in the simplified scheme at baseline and selling a large share of their sales to others in this scheme entering the VAT scheme. These firms would be in the VAT scheme after the reform and by definition have a low share of VAT sales, so taking this change into account would decrease the correlation further.

Table E.11: Supply chain segmentation under counterfactual scenarios

	Outcome: Probability in VAT scheme			
	Observed	Change in trade	Change in tax scheme	Change in both
	(1)	(2)	(3)	(4)
Share VAT sales	0.133***	0.137***	0.111***	0.113***
	(0.019)	(0.020)	(0.019)	(0.019)
Share VAT inputs	0.088***	0.061***	0.071***	0.045***
	(0.014)	(0.011)	(0.012)	(0.009)
Observations	640,634	640,634	640,634	640,634

Standard errors in parentheses are two-way clustered at the product and location level. The sample includes all firms with a minimum turnover over the period of less than 7 million INR. The right-hand side variables are the share of inputs (sales) purchased from (sold to) VAT-paying firms. In column 1 we use the observed shares and firms' tax schemes, see the text for a description of how the counterfactual shares and tax schemes in columns 2 to 5 are constructed. All specifications include location x product and year fixed effects. Significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.