Import Competition and Vertical Integration: Evidence from India

Joel Stiebale   Dev Vencappa [1]

August 2019

Abstract

Recent theoretical contributions provide conflicting predictions about the effects of product market competition on firms' organizational choices. This paper uses a rich firm-product-level panel data set of Indian manufacturing firms to analyze the relationship between import competition and vertical integration. Exploiting exogenous variation from changes in India's trade policy, we find that foreign competition, induced by falling output tariffs, increases backward vertical integration by domestic firms. The effects are concentrated in rather homogenous product categories, among firms that mainly operate on the domestic market, and in relatively large firms. Our results are robust towards different sub-samples and hold with or without conditioning on various firm- and product-level characteristics including input tariffs and firm-year fixed effects. We also provide evidence that vertical integration is associated with higher physical productivity, lower marginal costs and increasing markups.

JEL codes: *F14, F61, L25, L22, L23*
Keywords: *Trade Liberalization, Competition, Vertical Integration, Productivity*

---

1 Introduction

There has been a long-standing interest in the adjustment of firms to trade liberalization and foreign competition. Seminal theoretical and empirical contributions stress the reallocation of resources across heterogeneous firms within industries as a mechanism by which international trade raises industry-level productivity (e.g., Melitz, 2003; Pavcnik, 2002). Over time, the focus has shifted towards analyzing adjustments *within* firms via productivity enhancing investment (see, for instance, Bustos, 2011; Lileeva and Trefler, 2010) or changes in firms' product mix (Bernard et al., 2012; Eckel and Neary, 2010) as an important component of potential gains from trade.[1]

Recently, it has been pointed out that international competition can also affect productivity via firms' organizational choices such as their vertical integration intensity (Acemoglu et al., 2010; Alfaro et al., 2016; Aghion et al., 2006; Conconi et al., 2012; Legros and Newman, 2012, 2014).[2] However, the theoretical literature often yields mixed and non-monotonic predictions regarding the effects of product market competition on the organizational structure of domestic firms (Grossman and Helpman, 2002; Legros and Newman, 2014).[3] Therefore, the question of how import competition affects domestic firms' vertical integration ultimately boils down to an empirical matter. An empirical test of this relationship is not only of theoretical interest but also highly relevant for economic policy. There is evidence that vertical integration is often associated with higher productivity, lower prices, and enhanced innovation incentives.[4] Hence, if competition reduces incentives for vertical integration as argued by recent research (e.g., Conconi et al., 2012; Legros and Newman, 2014), policy measures such as deregulation and trade liberalization might have unintended consequences.

This paper provides evidence on the effects of import competition, induced by falling output tariffs, on vertical integration decisions of domestic firms in India's manufacturing industries. The case of India is particularly interesting for several reasons. First, Indian firms have been exposed to a substantial decline in output tariffs. The average most-favored nation tariff across industries—our inverse measure of import competition—fell from more than 100% in 1989 to about 15% towards the end of our sample period and varies substantially across products. Previous empirical evidence shows that variation in this decline of output tariffs across industries has been mostly unaffected by lobbying of domestic firms and displays little correlation with industry performance in years

---

[1]See also the surveys of related empirical literature in Bernard et al. (2012) and Shu and Steinwender (2019).

[2]Recent empirical studies that analyze the role of competition and trade for other aspects of firm organization include Cuñat and Guadalupe (2009); Bloom et al. (2010); Guadalupe and Wulf (2010); Marin and Verdier (2014); Görg and Hanley (2017); Chakraborty and Raveh (2018).

[3]Another strand of literature analyzes vertical integration and outsourcing of multinational firms in international markets (e.g. McLaren, 2000; Grossman and Helpman, 2003; Antràs and Helpman, 2004; Antràs, 2013). In this paper, we are primarily concerned with the effects of import competition on *domestic* firms' vertical integration.

[4]Although vertical integration can have anti-competitive effects via foreclosure, it is often found that the positive effects (from a welfare perspective) dominate (e.g., Hortaçsu and Syverson, 2007; Lafontaine and Slade, 2007).

before tariff changes (e.g., Goldberg et al., 2010a; Topalova and Khandelwal, 2011). Therefore, tariffs provide plausibly exogenous variation to the vertical integration decisions of domestic firms in India. In contrast to previous studies that exploit cross-sectional variation in tariffs across industries or industry-country pairs to analyze the relationship between competition and vertical integration, we provide within-country evidence and are able to control for unobserved heterogeneity across firms, products and industries. Another interesting aspect about Indian firms is that they seem to be characterized by the highest degree of vertical integration across more than 200 countries as reported by Alfaro et al. (2016).[5] Further, previous research has found that the Indian economy has been characterized by substantial misallocation of inputs across firms (Hsieh and Klenow, 2009) and high within-industry dispersion of productivity compared to other countries (see, for instance, Syverson, 2011). Vertical integration is a factor that potentially explains a significant part of this variation in efficiency across firms and time. Finally, in contrast to most other countries, Indian firms are required by law to report information about their production and sales at the product-level which is essential for our empirical approach.

To measure vertical integration, we follow previous empirical studies (e.g., Acemoglu et al., 2009; Alfaro et al., 2016) and construct an index based on the products firms produce and input-output tables. Specifically, we measure the fraction of inputs used in the production of a product that can be produced within the firm.[6] Our results show that this measure is significantly negatively correlated with output tariffs, indicating that foreign competition induces vertical integration by domestic firms. This result is robust towards limiting the analysis to firms' core product or assessing all products a firm produces and towards controlling for firm-product fixed effects, sectoral trends, input tariffs, various other industry- and firm-level control variables and—for a subsample of multi-product firms—even firm-year fixed effects. This allows us to rule out several alternative explanations for our results that are based on unobservable time-varying factors at the firm or industry-level. We also estimate larger effects of tariffs in subsamples where they are likely to play a more important role for competition faced by incumbent firms. For instance, we find that the effects are concentrated among firms that sell the majority of their output on the domestic market and in rather homogenous product categories.

The main purpose of our paper is to estimate a causal effect of trade liberalization on domestic firms' vertical integration rather than to distinguish between competing theories. Nonetheless, we discuss possible mechanisms. In a model inspired by property rights theory, Aghion et al. (2006)

---

[5]Acemoglu et al. (2003) provide a theory to explain that vertical integration is more widespread in countries with high distance from the world technological frontier—which arguably applies to India for most of its manufacturing industries—as these countries focus more on imitation and less on innovation.

[6]We focus on backward (rather than forward) vertical integration, i.e. inputs being integrated into downstream production since, in line with previous evidence (e.g., Acemoglu et al., 2010), we find that this is the empirically more important phenomenon.

predict a U-shaped relationship between competition and vertical integration. According to their theory, starting from a high level of market concentration, a small increase in competition reduces producers' incentives to integrate by improving suppliers' outside option which increases their investment incentives. For high degrees of competition, the model predicts that producers will have a high incentive to vertically integrate because independent suppliers can capture most of the surplus generated by the producer in the absence of integration. In line with this theory, our results indicate that there is indeed a U-shaped relationship between tariffs and vertical integration. However, according to our results, the positive effects of tariffs on vertical integration (implying a negative relationship between competition and vertical integration) only apply to a small part of the sample which faces the highest tariff rates.[7] Consistent with the results being driven by suppliers' and producers' relative bargaining position, we find that input tariffs have opposite effects on vertical integration to those of output tariffs. Input tariffs are positively associated with vertical integration for the majority of our sample and we also provide evidence for an inverted U-shaped relationship which mirrors the effects for output tariffs.

Aghion et al. (2006) find support for a U-shaped relationship using entry rates as measure of competition for a sample of industries in the UK. They further find that foreign entry via greenfield investment is mainly positively associated with vertical integration. However, as discussed by the authors, their results are rather suggestive. Specifically, they do not account for possible endogeneity of entry rates. In our empirical analysis, we address potential endogeneity problems by exploiting plausibly exogenous variation in tariff rates.

Our paper contributes to a growing literature that analyzes the impact of market structure on vertical integration and other organizational choices (see Legros and Newman, 2014, for an overview). Closely related to our empirical analysis is Alfaro et al. (2016) who exploit variation in tariffs across sectors and countries in a cross-section of plants around the world to show that output tariffs—which they use as a proxy for variation in price levels—and vertical integration are positively correlated. In contrast, our paper exploits cross-industry and time series variation within a single country which enables us to control for unobserved heterogeneity across firms and firm-products and to exploit plausibly exogenous changes in trade policy.

Alfaro et al. (2016) explain their findings by a theoretical model based on the organizational industrial organization literature (Legros and Newman, 2014, 2017) in which vertical integration is regarded as productivity enhancing investment. This literature argues that high prices spur the incentives for vertical integration since the benefits from increased productivity increase with

---

[7]In contrast, Grossman and Helpman (2002) predict that a non-linear *inverted* U-shaped relationship between competition—operationalized as the degree of product differentiation—and vertical integration can arise, depending on cost differences between vertically integrated and specialized producers and the relative bargaining power between upstream and downstream firms.

the level of prices while costs of vertical integration are independent of price and output levels.[8] However, these studies analyze the impact of prices in a perfectly competitive setting. While the assumption of price-taking firms might be reasonable for some markets, it is unlikely to hold for Indian manufacturing industries which seem to be characterized by considerable market power and pricing heterogeneity (De Loecker et al., 2016; Stiebale and Vencappa, 2018).

Evidence for a negative effect of competition on vertical integration via prices is also provided by McGowan (2017) who exploits a natural experiment in the US coal mining industry. Acemoglu et al. (2010) find a negative association between a higher number of firms producing on the domestic market and the incidence of backward vertical integration in a cross-section of UK industry-pairs. Buehler and Burghardt (2015) find that a reduction of non-tariff barriers, the introduction of mutual recognition of product standards between Switzerland and the EU, reduced the vertical integration propensity of Swiss plants. In contrast to this paper, we focus on the reduction of tariffs in a univariate trade liberalization which allows to isolate the effects of foreign competition. Liu et al. (forthcoming) find that reductions of output tariffs in Chinese *upstream* industries—induced by China's WTO entry—are associated with a decline in the number of firms in these industries that are acquired by downstream firms. Unlike this paper, we focus on tariffs in *downstream* industries and investigate a broader measure of vertical integration which is not limited to integration via acquisitions.

Our paper is also related to the literature that studies the effects of competition on innovation and other productivity enhancing investments.[9] This literature provides ambiguous predictions on the relationship between competition and investment. On the one hand, competition reduces profit margins and market shares per firm and thus the returns to investment, the well known Schumpeterian effect. On the other hand, competition also decreases rents in the absence of investment, inducing firms to invest to escape competition.[10] Aghion et al. (2005) develop a theory which combines these countervailing mechanisms and predicts that competition spurs investment for low initial levels of

---

[8]A negative relationship between competition and vertical integration is also predicted by the transaction cost approach where competition reduces asset specificity and therefore the need for vertical integration (Aghion et al., 2006) and by a version of the property rights theory approach developed in Acemoglu et al. (2010) which mainly focuses on the relationship between technology and vertical integration. See the seminal contributions by Williamson (1975), Williamson (1985) for the transaction cost approach and by Grossman and Hart (1986), Hart and Moore (1990) for the property rights theory approach. Lafontaine and Slade (2007) provide a survey on the general determinants and effects of vertical integration.

[9]Consistent with the view that vertical integration is a productivity enhancing investment, we find that vertical integration is associated with higher physical productivity, declining marginal costs and increasing markups, which are derived using recent advances in the estimation of production functions for multi-product firms (De Loecker et al., 2016).

[10]This mechanism goes back to Arrow (1962). Similarly, in in an oligopolistic market, an increase in competition can lead to more investment as it increases the sensitivity of demand to enhanced efficiency (Schmutzler, 2013; Vives, 2008). Another mechanism implying that competition can spur investment occurs when production factors are trapped inside firms and competition lowers the opportunity costs of redeploying these factors towards other uses (Bloom et al., 2016).

competition but reduces investment incentives when markets are already highly competitive. Their theory also implies that competition is more likely to spur investment by industry leaders operating close to the technological frontier. Consistent with this "escape competition effect", we find that negative effects of tariffs on vertical integration are concentrated in initially large firms which are more likely to compete neck-to-neck with foreign producers that are exporting to India.

As discussed by Shu and Steinwender (2019), the majority of empirical studies finds a negative association between international competition and productivity enhancing investments such as R&D in the US while the results for European countries are mixed and most studies based on developing countries estimate a positive relationship, possibly due to differing levels of initial competition intensity and institutional aspects. Interestingly, these differences are consistent with the contrasting results between our analysis and those obtained by other studies like Alfaro et al. (2016) and McGowan (2017).

The rest of this paper is organized as follows. Section 2 discusses the related literature. Our data set and construction of variables are described in section 3. The empirical strategy is detailed in section 4, with results discussed in section 5. Section 6 concludes.

2 Data and Variables

Our empirical analysis draws from several data sources. Our primary data set is the Centre for Monitoring of the Indian Economy (CMIE) Prowess database. We augment this primary data source with a number of additional data sets. We carry out the analysis at the level of National Industrial Classification (NIC) version 2008, and where external data sources use international industrial classifications such as the Harmonised System (HS) codes, we mapped these onto NIC following the concordance tables published by Debroy and Santhanam (1993).

2.1 Firm and Product Level Data

Prowess provides information on company balance sheets and income statements for both publicly listed and unlisted firms across industries in the manufacturing, services, utilities and financial sectors.[11] These firms account for more than 70% of industrial output from the organised sector, 75% of corporate taxes and 95% of excise taxes collected by the government.

The construction of the vertical integration index at the firm level requires us to identify the products produced by the firm. By law, Indian firms are required to report product-level data on

---

[11]This database has been used in a number of recent papers, e.g. Goldberg et al. (2009, 2010a,b); De Loecker et al. (2016).

quantities and values of sales and production.[12] The data thus allows separating production used as an input for other products from intermediate goods sold to other firms. Each product is allocated a twenty-digits code from CMIE's own internal classification of 5908 sub-industries and products. Of these, 4833 products fall under the manufacturing sector. CMIE's own classification is largely based on the Indian National Industrial Classification (NIC) and the HS schedule. Examples of products across different industries include shrimps, corned meat, sponge iron, pipe fittings and rail coaches. Goldberg et al. (2010b) provide a detailed description of the product-level data in Prowess.

We extracted data spanning the period 1989 (the first year firms appear in the database) until 2011 and focus on the manufacturing sector. Hence, we create an unbalanced panel tracking products of each firm every year, mapping the product codes onto India's NIC 2008.[13] In our main estimation sample, we exclude multinational firms since we have no detailed information about the products they produce abroad. Further, these firms may not be affected by import competition in the same way as domestic firms. However, as we discuss in the robustness section, this restriction is not crucial for our results.

2.2   Vertical Integration Indices

We follow Fan and Lang (2000) and Alfaro et al. (2016) in constructing the vertical integration indices using several issues of published Indian input-output (IO) tables. IO tables report transaction coefficients which measure the rupee value of output from industry i required to produce a rupee's worth of output for industry j. Hence, a transaction coefficient of 0.05 means that 5 Indian rupee cents (paise) worth of output in industry i are required to produce one rupee's worth of products in industry j. We combine information on firms' production activities in the Prowess database with IO tables, and construct IO indices for each product produced in industry j by firm f at year t. We define a firm based on unconsolidated accounts and therefore implicitly treat firms that are part of corporate groups as independent entities. This is consistent with recent evidence that physical input flows between firms of corporate groups are limited (Atalay et al., 2014). However, as we discuss in a robustness section, excluding firms that are part of a corporate group from the estimation sample does not affect our results notably.

For our main specification, we use constant IO transaction coefficients which stem from an IO table referring to the fiscal year 1993/94. Constant IO weights across time within industries ensure

---

[12]This is a requirement of the 1956 Companies Act.

[13]In setting up this firm-product-year panel, we checked and adjusted the CMIE product codes to address a number of instances where the same product code was attributed to different products, or where different product codes were allocated to the same product. In addition, we noticed a number of cases where product names varied in spelling and also noted frequent differences in levels of aggregation for what constitutes a product. After cleaning the data, accounting for missing values, and aggregating some products, we were left with 2782 clean and unique CMIE product codes in Prowess.

that our empirical analysis captures variation in firms' production activities rather than time series variation in IO transaction coefficients for firms with a constant product portfolio. However, we also experiment with time varying weights. IO tables are published on an interval of roughly about 5 years, and to create a product-year panel of IO weights, we use the 1993/94 IO transaction coefficients for years 1988-1997, the 1998/99 IO coefficients for adjacent years 1998-2002, the 2003/2004 IO coefficients for adjacent years 2003-06 and the 2007/08 IO coefficients for the remaining years 2007-11 of our sample.[14]

We construct vertical integration indices at the level of the firm as well as at the firm-product level by combining IO transaction coefficients with information on firms' production activities from Prowess. At the firm-product level, the IO index, IO_{fijt}, follows the identity:

$$IO_{fijt} \equiv IO_{ij} \times I_{fijt} \tag{1}$$

IO_{ij} is the IO coefficient for product-industry pair (i, j) and measures how much of i is typically needed to produce one unit of j. The second element, I_{fijt}, is an indicator variable that equals one if the firm produces both i and j at time t. If a firm produces both i and j, it is assumed to supply itself with all the i that is necessary to produce j. Hence, a firm that produces i will be measured to be more integrated in the production of j, the higher IO_{ij}.

The firm's vertical integration index for a product produced in industry j is the sum of IO coefficients across all industries in which it is active:

$$v_{fjt} = \sum_i IO_{fijt} \tag{2}$$

At the firm level, its vertical integration index is as per equation (2), but calculated for its main industry of activity only.[15]

2.3   Tariffs

Tariffs data were sourced from World Integrated Trade Services (WITS). Following Alfaro et al. (2016), we use most favored nation (MFN) tariff rates applied by GATT/WTO members. We select the tariffs data reported at 6 digits HS codes, and map these to NIC codes following the concordance

---

[14]IO 1989-90 tables, which we could have used for years 1988-1992, are unfortunately not available to us.
[15]We define a firm's main product as the product with the highest revenue share throughout the sample period.

tables published by Debroy and Santhanam (1993).[16] We construct simple averages of tariff rates aggregating from six digits HS codes to 3 digits NIC 2008 codes.

We also construct a measure of tariffs applied to imported inputs, which simply weighs the applied MFN tariffs using normalised IO coefficients as weights.

$$inptariff_{jt} = \sum_{i \neq j} tariff_{it} * IO_{ij} \tag{3}$$

where $tariff_{jt}$ are MFN tariff rates and IO_{ij} represents the IO transaction coefficients.

2.4 Further Variables

We construct a number of additional variables which we use as additional control variables or to measure heterogenuous effects where relevant. These include two measures of technological intensity: *investment intensity*, measured as firm investment over sales and *R&Dintensity*, measured as firm R&D relative to sales. Both variables are averaged at the industry level. Another industry-level measure is average *size* of firms within an industry measured by the log of sales. At the firm-level we construct *exports* and *imports* measured relative to sales. A firm's size is capture by *log(sales)*. *log(R&D)* denotes the log(research and development expenditure + 1). *homogenous good* is a dummy variable taking value 1 if the product belongs to a category defined as homogenous and 0 if differentiated. The classification of product categories as differentiated or homogenous follows Rauch (1999).[17]

markup is a product-level markup calculated following the methodology of De Loecker et al. (2016) (see Appendix B for details). Under the assumption that firms minimize costs it can be shown that the markup (ratio of price to marginal costs) equals the material-output elasticity divided by the material to sales ratio. The material-output elasticity is estimated from a translog production function with physical quantity as output and materials, capital and labour as production factors.[18] *marginal cost* is calculated by dividing observed prices (unit values) by estimated markups.

---

[16]The tariffs data were brought to the common HS 1992 codes and from there on mapped onto NIC codes.
[17]For details on the Rauch classification, see `http://econweb.ucsd.edu/~jrauch/rauch_classification.html`, accessed on August 6, 2019.
[18]See also Stiebale and Vencappa (2018) for details on the construction of product level markups and marginal costs using the same data set as this paper.

2.5 Summary Statistics

Table 1 reports summary statistics for our main variables of interest. In our sample, the median vertical integration index is 0.116 and the mean is 0.134. The distribution of the vertical integration index is depicted in Figure 1. Around 45% of the products considered in our data set are classified as being homogenous goods. The average markup of 2.82 seems quite high, but the median markup is 1.33 for the whole sample. [19] On average, exports make up about 10% of production. Table 2 reports the yearly evolution of output and input tariffs data over the period of our analysis. Average output tariffs fell at a rapid pace in the earlier years around the 1991 reforms and slowed down in later years. A similar observation can be made for input tariffs.

2.6 Exogeneity of Trade Policy

A particular concern around the use of a trade policy measure such as tariffs is the possible endogeneity of this variable. For instance, governments might perceive that specific domestic industries are not sufficiently mature to face import competition, and would seek to protect them. Similarly, inefficient domestic industries might lobby for higher protection from foreign competition. Labour or trade union groups may have concerns around the employment and welfare effects of trade liberalisation and will likely lobby authorities in an attempt to steer the magnitude and pace of trade liberalisation. A number of studies have argued that the 1991 trade reforms in India that arose from the external crisis were drastic and unexpected (e.g., Hasan et al., 2007; Topalova and Khandelwal, 2011). Yet, as Topalova and Khandelwal (2011) points out, variations in trade policy across industries could confound inference if it is correlated with previous industry performance. The authors argue that this might have been the case for more recent trade liberalization episodes in India.

In contrast to Topalova and Khandelwal (2011), our paper uses MFN tariffs and we argue that these can be assumed to be exogenous to vertical integration. As Alfaro et al. (2016) point out, MFN tariffs are agreed following long rounds of multilateral trade negotiations at the end of which every member commits to not exceed agreed tariff bounds. Failure to respect this commitment entitles the affected parties to take matters to the dispute settlement body of the WTO. Once agreed, the tariff rates must be applied in a non-discriminatory way to imports from all WTO members.[20] Pressure for protection from lobby groups is unlikely to be directed to MFN tariffs; governments can instead focus on alternative measures such as antidumping measures and countervailing duties. MFN tariffs are also persistent and vary little over time.

---

[19]These figures are similar to those obtained by De Loecker et al. (2016) who estimate a markup distribution for Indian manufacturing over an earlier time period, reporting an average of 2.70 and a median of 1.34.

[20]Exceptions to this rule are when WTO members form part of a preferential trade agreement, in which case members are allowed to discriminate between members inside and outside the preferential trade agreement.

Although the above are strong arguments for MFN tariff rates being exogenous, we follow Topalova and Khandelwal (2011) and run a variety of checks prior to estimating a causal link of tariffs on our outcome variables. Firstly, we use industry level data from various issues of the Annual Survey of Industries to run regressions of changes in trade policy measures (output and input tariffs) on lagged industrial characteristics for each distinct time period corresponding to India's five-year plans as well as the whole sample period from 1989-2012.[21] Industry performance indicators include employment, output, average wage, concentration, share of skilled workers and the growth of industry output and employment. Specifically, we run the following regressions:

$$tariff_{jt} - tariff_{j0} = \delta_0 + \delta_1 x_{j0} + u_j \tag{4}$$

where $tariff_{jt}$ is either input or output tariff in industry j at time t, and x_{j0} is each of a set of industrial characteristics measured at the beginning of each 5-year time period.

Table 3 reports results of these simple regressions. While for some variables and some time periods, tariff changes appear to be correlated with industry characteristics, most of these correlations are small and—with the exception of average wage levels—statistically insignificant or weakly significant. All in all, there is no evidence that policy makers systematically adjusted tariffs to previous industry performance.

We also follow Topalova and Khandelwal (2011) in a second set of regressions to check whether policy makers adjusted tariffs in response to industry productivity shocks and regress each of the trade policy measures on a one-year lagged industry productivity measure:

$$tariff_{jt} = \zeta \, TFP_{j,t-1} + \mu_t + a_j + u_{jt} \tag{5}$$

where TFP_{jt} denotes average log total factor productivity, estimated using the methodology proposed by Levinsohn and Petrin (2003).[22] The industry-level measure is constructed as a sales-weighted average of (absolute) firm-level TFP.

Table 4 reports the results of these regressions for different time periods with output tariffs in Panel A and input tariffs depicted in Panel B. Coefficients are small and statistically insignificant for the sample period as a whole and insignificant or weakly significant for the different sub-samples. Even the highest coefficient, the correlation between TFP_{t-1} and output tariffs within the time

---

[21]Over the period 1989-2012 India delivered its 7th to 11th five-year plans for the periods 1985-1990, 1992-1997, 1997-2002, 2002-2007 and 2007-2012.

[22]We use sales as a measure of output and material costs, wage bill and fixed assets to measure inputs.

period 1997-2002 suggest that a 10% increase in total factor productivity is associated with only 0.64 percentage points higher output tariffs and has the opposite sign that we would expect when policy makers try to protect low-productivity industries with high tariffs. Taken together, the results from tables 3 and 4 suggest that potential endogeneity concerns around our tariffs variables are mitigated.

3   Empirical Method

The aim of the empirical analysis is to estimate the effects of tariffs on vertical integration. We exploit variation in most-favored nation tariffs across products and years within and between firms. For this purpose, we start with the following regression at the firm-level:

$$v_{f(j)t} = \beta \; tariff_{jt} + X'_{f(j)t}\Pi + \alpha_f + \eta_{kt} + \epsilon_{f(j)t} \tag{6}$$

where $v_{f(j)t}$ denotes the vertical integration index of firm f with main activity in industry j in year t. $tariff_{jt}$ denotes the tariff rate applied to industry j at time t and $X_{f(j)t}$ is a vector of firm-and product specific control variables. The firm fixed effect α_f captures permanent differences among firms including location (which might affect the supply of intermediate inputs). η_{kt} are time dummies which control for changes in market conditions and technology common to all firms which we allow to vary across 2-digit industries (k) in most specifications. Finally, $\epsilon_{f(j)t}$ is an error term.

In a second step, we move the analysis to the firm-product-level and consider all products produced by firms, not only their core product. Therefore, equation (6) becomes:

$$v_{fjt} = \theta \; tariff_{jt} + X'_{jt}\Theta + \alpha_{fj} + \eta_{kt} + \epsilon_{fjt} \tag{7}$$

where v_{fjt} denotes the vertical integration index of product in industry j produced by firm f in year t and α_{fj} is a firm-product specific fixed effect which captures permanent differences in technology and product characteristics. Since tariffs vary across industries within multi-product firm-years, we control additionally for firm-year fixed effects, φ_{ft}, in a further amplification of the model:

$$v_{fjt} = \gamma \; tariff_{jt} + X'_{jt}\Gamma + \alpha_{fj} + \varphi_{ft} + \eta_{kt} + \epsilon_{fjt} \tag{8}$$

An advantage of controlling for firm-year fixed effects is that we can control for time-varying adjustments within firms such as changes in management, financial factors or productivity (as long as these changes are not firm-product-year specific). Firm-year fixed effects also control for the

effects of product-specific tariffs that affect a firm as a whole, for instance via liquidity and credit constraints. A disadvantage of this approach is that we can only run this regression on a selected sample of multi-product (multi-industry) firms which are arguably not a random sample from the population.

Our main identifying assumption is that $E[\epsilon_{fjt}|tariff_{jt}, X_{jt}, \alpha_{fj}, \varphi_{ft}, \eta_{kt}] = 0$. Hence, we assume that unobservables affecting vertical integration decisions which are not captured by firm-year, firm-product or sector-year fixed effects are uncorrelated with variations in tariffs across time within industries. To asses the validity of this approach, we include a set of control variables, X_{ijt}, which are potentially correlated with both vertical integration and tariff rates in some specifications. A potentially important control variable are input tariffs which can affect producers' vertical integration decisions via prices and competition in upstream markets (Acemoglu et al., 2010; Alfaro et al., 2016). These are computed by passing tariff rates through India's input-output matrix. Further, technological characteristics might determine the degree of relation-specificity and investment incentives (Acemoglu et al., 2010). To capture changes in technological characteristics across industries and time, we control for the average of the R&D to sales ratio and the level of investment relative to sales across firms within industries. We also control for variation in average firm size (measured by log sales) to capture general changes in technology. At the firm-level, we control for the ratio of export to sales and imports to sales to account for the fact that firms with access to foreign markets might react differently to changes in import competition. We refrain from controlling for further firm-level variables such as productivity, size, capital or R&D since these variables might be affected by firms' vertical integration choices. However, these variables are implicitly controlled for using firm-year fixed effects in our product-level specifications.

Since firms are arguably more affected by output tariffs if they mainly operate on the domestic market, we follow Alfaro et al. (2016) and also test for heterogeneous effects using the following specification:

$$v_{fjt} = \nu_1 \; tariff_{jt} + \nu_2 \; tariff_{jt} \times domestic_f + X'_{jt}\Psi + \alpha_{fj} + \varphi_{ft} + \eta_{kt} + \epsilon_{fjt} \tag{9}$$

where $domestic_f$ takes on a value of one for firms that export less than half of their output throughout our sample period.[23] We also test for further heterogeneous effects in alternative specifications in which we replace $domestic_f$ by an indicator of firm size or a variable that distinguishes between differentiated and homogenous goods based on the Rauch (1999) classification.

As an additional test, we replace sector-year with product-year fixed effects (measured at the

---

[23]This variable is firm but not firm-product-specific since, unfortunately, our data does not provide a breakdown of firm-level exports by products.

same level as tariffs) which allows identifying the effect of the interaction term $tariff_{jt} \times domestic_f$ while controlling for other time-varying market conditions that might be correlated with both tariffs and vertical integration.

$$v_{fjt} = \chi \, tariff_{jt} \times domestic_f + X'_{jt}\Xi + \alpha_{fj} + \varphi_{ft} + \eta_{jt} + \epsilon_{fjt} \tag{10}$$

In all specifications, we use two-way clustered standard errors. First, we cluster at the product-level since our main variable of interest, tariffs, varies at the product-level while the dependent variable is firm-product specific. Second, we cluster at the firm-level since vertical integration decisions might be correlated within firms across products and years.

4 Results

4.1 Main results

Table 5 reports the results our firm-level regressions based on estimation of equation (6). Column (1) shows results of regressions that control for firm fixed effects and year dummies. The coefficient indicates that a 100 percentage point increase in output tariffs is associated with a decline of the vertical integration index by 0.022 which corresponds to about 19% of the median value of the index. While the estimated effect becomes somewhat smaller when we control for sector (2-digit industry)-year fixed effects in columns (2), it increases again once we add further control variables. In column (3), we control for input tariffs, while column (4) adds additional control variables at the industry-level to the regression which include investment and R&D intensities (as a proxy for technology) and average firm size. In column (5), we control for two additional variables at the firm-level, import and export shares. The results for output tariffs remain highly statistically significant and indicate economically important effects as well. According to the coefficient estimates in column (6) and descriptive statistics in Table 1, an increase in output tariffs by one standard deviation decreases the expected vertical integration index by 0.0064 which is about 5.5% of the median vertical integration index. As we discuss below, the effects are even more pronounced for some subsamles. Among the control variables, mainly investment intensity turns out to be significant. This is in line with Acemoglu et al. (2010) who argue that technological intensity in the producers' industry increases the incentives for vertical integration. Both input tariffs and R&D intensity are positively associated with changes in vertical integration, but the results are not statistically significant.

Results of firm-product level regressions based on estimation of equation (7) are depicted in Table 6. Column (1) shows results of regressions that control for firm-product fixed effects and

year dummies. In column (2), we add sector fixed effects and column (3) controls for input tariffs. The estimated effects are again statistically significant and in absolute terms even somewhat higher than in firm-level regressions. A more restrictive test of the effects of tariffs on vertical integration is whether particular firms are more likely to vertically integrate in products/industries with lower tariff rates. For instance, variation in tariff rates across time and industries might be correlated with unobserved changes in corporate culture or company wide trends in investment, management or productivity. To assess this possibility, we add firm-year fixed effects to the model. Identification of the effects of tariffs in this specification is limited to firms that produce in at least two different industries. The estimates become even stronger. For instance, column (4) indicates that a one-standard deviation increase in tariffs is associated with declines in vertical integration of 0.012 which is more than 10% of the median vertical integration index. The results remain significant if we control for sector-year fixed effects and input tariffs in columns (5) and (6). In column (6), the coefficient for input tariffs becomes statistically significant. As we will see below, the lack of significance of the input tariff coeffiient in some specifications is mainly due to heterogeneous responses to input tariffs across firms and industries. While the coefficient for input tariffs is in absolute terms much larger than the coefficient for output tariffs, the magnitude of the relative impact is rather similar. A one standard deviation increase in input tariffs increases vertical integration by 0.011.

In Table 7, we present the results of heterogeneous effects of tariffs across homogenous and differentiated goods. The effects of tariffs on domestic firms should be more pronounced in homogenous product categories where the elasticity of substitution across varieties is arguably higher. This is exactly what we find. Estimates in column (1) indicate that a 100 percentage point increase in output tariffs decrease the vertical integration index of firm-products in homogenous good industries by 0.086, more than 50% of the mean or median vertical integration index in these product categories.[24] The coefficient is more than six times higher than the coefficient for differentiated goods industries documented in column (2). When we add firm-year fixed effects to these regressions, the differences across product categories even become more pronounced and the coefficient for differentiated goods loses its significance. For homogenous goods industries, there is clear evidence that both input and output tariffs affect vertical integration. This is consistent with international competition affecting vertical integration incentives via the relative bargaining position of suppliers and buyers.

We investigate further heterogeneous effects in Table 8. For instance, the response to import competition might be quite different for industry leaders compared to smaller and less productive firms. Column (1) uses an interaction term between tariffs and a dummy variable which takes on

---

[24]The average (median) VI index equals 0.160 (0.163) in homogenous and 0.114 (0.093) in differentiated product categories. The average (median) tariff rate equals 0.366 (0.317) in non-differentiated and 0.305 (0.253) in differentiated industries.

value one when sales exceed those of the median firm within an industry, measured in the year a firm enters the sample. The results show that the sensitivity of vertical integration decisions to changes in tariffs by relatively large firms is almost twice as high compared to smaller firms. Larger firms also respond to changes in input tariffs significantly. In column (2), we perform a separate regression for firms that export on average more than half of their output within our sample period. If the coefficients for output tariffs reflect a causal effect of import competition on Indian firms, the estimated effects should be much less pronounced for firms with high export shares. The results confirm this hypothesis as the coefficient for output tariffs becomes statistically and economically insignificant. However, exporters seem to respond much more to changes in import tariffs than domestic firms. A possible explanation is that firms with high export shares have better access to foreign input suppliers and use these to substitute domestic in-house production of intermediates.[25] We analyze differences in the responses of firms with high export shares and domestic firms in more detail in columns (4) and (5) based on estimation of equations (9) and (10). The estimates confirm that domestic firms' vertical integration propensity significantly decreases with the level of output tariffs. This effect even becomes stronger when we control for product-year fixed effects in column (5).

So far, we have assumed a monotonic relationship between tariffs and vertical integration. However, the effect of tariffs might be non-linear. For instance, Aghion et al. (2006) predict a U-shaped relationship between competition and vertical integration. Column (1) of Table 9 shows results of adding a squared term of tariffs to the regression. The coefficients indeed indicate a non-linear relationship where the negative effects of tariffs on vertical integration is decreasing in the level of tariffs. Since tariffs are an inverse measure of foreign competition, the estimates are consistent with a U-shaped relationship between import competition and vertical integration. The predicted turning point, where further increases in tariffs have positive effects on vertical integration, is at a tariff rate of 115% percent which approximately equals the 95-percentile of the distribution of tariffs. Hence, the effect of foreign competition seems to be positive for the vast majority of firms in our sample. We also experimented with alternative functional forms for the relationship between tariffs and vertical integration. For instance, following Alfaro et al. (2016), we regress $\ln(v + 1)$ on $\ln(tariff + 1)$ and—within the subsample of firm-product-years with positive values of vertical integration and tariffs—we relate $\ln(v)$ to $\ln(tariff)$. The results depicted in column (2) and (3) confirm the negative relationship between output tariffs and vertical integration. The specification in column (4) adds a squared value of $\ln(tariff + 1)$ to the specification. Again, the results indicate

---

[25]The share of importers among firms which export more than half of their output is almost 90% while it is below 40% for other firms. Unfortunately, our data does not provide detailed information about the goods that firms import which prevents us from analyzing this channel in more detail.

a turning point above the 95-percentile of tariffs.[26]

4.2   Vertical integration and efficiency

In most theoretical models we discussed, vertical integration is interpreted as a productivity enhancing investment. It is thus natural to ask whether it is indeed associated with improved performance in our sample. For this purpose, we correlate vertical integration with a firm-level measure of total factor productivity (TFP) and firm-product-level measures of markups and marginal costs. Following De Loecker and Warzynski (2012), markups are estimated as the ratio of the material-output elasticity to the cost share of materials. The material-output elasticity is estimated from a translog production function based on physical units of output using the method suggested by De Loecker et al. (2016). Estimation of the production function also yields a measure of physical TFP at the firm-level. Marginal costs are recovered from dividing estimated markups by observed unit values. We describe this method in detail in Appendix B.[27] Results displayed in Table 10 show that within-firm-variation in vertical integration is indeed negatively correlated with marginal costs and positively correlated with markups and TFP. The coefficients for markups and costs even increase (in absolute terms) when we control for product-year or firm-year fixed effects. Nonetheless, we would like to emphasize that these estimates only identify correlations since vertical integration might be endogenous to costs, efficiency and markups.[28]

4.3   Additional robustness checks

We checked the robustness of our results towards different empirical methods and estimation samples which are documented in Appendix A. Table A1 shows results using alternative estimation methods which account for zero values in the dependent variable. Particularly, we estimate a Tobit model with a censoring point at zero, fractional response Logit and Probit models and a quasi-maximum likelihood estimator of an exponential mean model based on a Poisson distribution. Due to the incidental parameters problem and computational difficulties, we cannot control for firm-product fixed effects in these estimations. Instead, we control for industry dummies at the 3-digit level to ensure that our estimates identify variation in tariffs and vertical integration within industries across time. Columns (1) to (3) show marginal effects from the Tobit, fractional Logit and fractional Probit

---

[26]Since ln(*tariff*) is negative for most observations, a squared term of ln(*tariff*) is not appropriate for describing a U-shaped relationship. Hence, we do not report the results of such a specification.

[27]See also De Loecker et al. (2016) and Stiebale and Vencappa (2018) for further details.

[28]Note that rising markups do not necessarily imply that vertical integration increases market power. For instance, markups can increase due to incomplete pass-through of costs to prices and vertical integration might also yield higher product quality. Identifying the sources of variation in markups and the causal effect of vertical integration on firm performance is beyond the scope of this paper.

estimations respectively. In column (4), coefficients from an exponential mean model, which can be interpreted as semi-elasticities, are depicted. All these alternative estimation methods confirm the negative effects of output tariffs on vertical integration.

We performed further robustness checks which relate to the selection of the estimation sample. First, we limited the sample to stand-alone firms since it is not clear whether our vertical integration index describes the production activities of firms within corporate groups accurately. Results, depicted in Table A2, show that excluding firms that are part of corporate groups does not affect our conclusions. In our main estimation sample, we exclude multinational firms since we have no detailed information about their foreign production activity. However, as documented in Table A3, including these firms and their production activities in India in the estimation sample does not affect our main results either. Finally, we checked the robustness towards aggregating all production activities to the 3-digit industry level which corresponds to the level of aggregation of our tariff variables. Results, depicted in Table A4 show very similar coefficients as in the baseline regressions. All in all, the results are robust towards different control variables, functional forms, econometric methods and estimation samples and indicate that there is a strongly significantly negative effect of import competition on the vertical integration propensity of domestic firms.

5 Conclusion

Recent theoretical contributions provide conflicting predictions about the relationship between competition and firms' organizational choices. In this paper, we use a rich firm-product level panel data set of Indian manufacturing firms to analyse the relationship between import competition and vertical integration. Our identification strategy exploits the fact that foreign competition faced by domestic firms increased substantially due to India's trade liberalisation which reduced average MFN output tariffs from more than 100% in 1989 to about 15% in 2011. We provide evidence that variation in the decline of MFN rates across industries is only weakly correlated with lagged industry performance measures such as productivity and output growth indicating that declines in tariffs provide plausibly exogenuous variation to the vertical integration decisions of domestic firms in India.

Following previous empirical studies, we construct an index of vertical integration based on firms' products and IO tables, which captures the fraction of inputs used in the production of a product that can be produced inhouse. Relating this measure to output tariff rates, our inverse measure of import competition, we find a strong and statistically significant negative relationship, suggesting that foreign competition induces vertical integration by domestic firms. This result holds

whether we consider all the products of a firm or focus only on their core product. Our finding are similarly robust to controlling for firm-product fixed effects, sectoral trends, input tariffs, various other industry- and firm-level control variables and even firm-year fixed effects.

While the negative effect of tariffs on vertical integration dominate in our sample, we also provide evidence for a U-shaped relationship driven by industries and time periods with the highest tariffs rate. This non-linearity is consistent with a model inspired by the property rights theory approach developed by Aghion et al. (2006). Consistent with the results being driven by suppliers' and producers' relative bargaining position, we find that input tariffs have the opposite effects on backward vertical integration.

We also find evidence that the responses to import competition vary with firm and industry characteristics. For instance, our results indicate that the effects of tariffs on vertical integration are concentrated in firms that mainly operate on the domestic market, in relatively large firms and in rather homogenous product categories. There is also evidence that higher input tariffs induce backward vertical integration, particularly for homogenous products, firms with high export sales, and firms of large size.

From an economic policy point of view, our results indicate that concerns that increased competition may have a negative effect on firm performance via reducing incentives for vertical integration do not seem to be justified, at least in the context of India. This result contrasts with recent empirical evidence and it will be an interesting topic for future research to investigate whether these differences are due to the specificity of an emerging market like India or due to exploiting different sources of variation in import competition.

References

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti, "Vertical integration and distance to frontier," *Journal of the European Economic Association*, 2003, *1* (2-3), 630–638.

_ , **Rachel Griffith, Philippe Aghion, and Fabrizio Zilibotti**, "Vertical integration and technology: theory and evidence," *Journal of the European Economic Association*, 2010, *8* (5), 989–1033.

_ , **Simon Johnson, and Todd Mitton**, "Determinants of vertical integration: financial development and contracting costs," *Journal of Finance*, 2009, *64* (3), 1251–1290.

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Aghion, Philippe, Rachel Griffith, and Peter Howitt, "Vertical integration and competition," *American Economic Review*, 2006, *96* (2), 97–102.

Aghion, Phillipe, Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and innovation: An inverted U relationship," *Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

Alfaro, Laura, Paola Conconi, Harald Fadinger, and Andrew F Newman, "Do Prices Determine Vertical Integration?," *Review of Economic Studies*, 2016, *83*, 855–888.

Antràs, Pol, "Grossman–Hart (1986) Goes Global: Incomplete Contracts, Property Rights, and the International Organization of Production," *Journal of Law, Economics, and Organization*, 2013, p. ews023.

_ **and Elhanan Helpman**, "Global sourcing," *Journal of Political Economy*, 2004, *112* (3), 552–580.

Arrow, KJ, "Economic Welfare and the Allocation of Resources to Invention. The Rate and Direction of Economic Activity: Economic and Social Factors. U.-NBC f. ER a. C. o. EG ot SSR Council," 1962.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson, "Vertical integration and input flows," *The American Economic Review*, 2014, *104* (4), 1120–1148.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott, "The empirics of firm heterogeneity and international trade," *Annual Review of Economics.*, 2012, *4* (1), 283–313.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," *Review of Economic Studies*, 2016, *83* (1), 87–117.

_ , **Raffaella Sadun, and John Van Reenen**, "Does product market competition lead firms to decentralize?," *American Economic Review*, 2010, *100* (2), 434–38.

Buehler, Stefan and Dirk Burghardt, "Globalization and vertical firm structure: An empirical investigation," Technical Report, Mimeo, University of St. Gallen 2015.

Bustos, Paula, "Trade liberalization, exports, and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms," *American Economic Review*, 2011, *101* (1), 304–340.

Chakraborty, Pavel and Ohad Raveh, "Input-trade liberalization and the demand for managers: Evidence from India," *Journal of International Economics*, 2018, *111*, 159–176.

Conconi, Paola, Patrick Legros, and Andrew F Newman, "Trade liberalization and organizational change," *Journal of International Economics*, 2012, *86* (2), 197–208.

Cuñat, Vicente and Maria Guadalupe, "Globalization and the provision of incentives inside the firm: The effect of foreign competition," *Journal of Labor Economics*, 2009, *27* (2), 179–212.

Debroy, B and A. T. Santhanam, "Matching Trade Codes with Industrial Codes," *Foreign Trade Bulletin*, 1993, *24* (1).

Eckel, Carsten and J Peter Neary, "Multi-product firms and flexible manufacturing in the global economy," *Review of Economic Studies*, 2010, *77* (1), 188–217.

Fan, J. P. H. and L. H. Lang, "The Measurement of Relatedness: An Application to Corporate Diversication," *Journal of Business*, 2000, *73*, 629–660.

Goldberg, Pinelopi K, Amit K Khandelwal, Nina Pavcnik, and Petia Topalova, "Trade liberalization and new imported inputs," *American Economic Review*, 2009, *99* (2), 494–500.

_ , _ , _ , **and** _ , "Imported intermediate inputs and domestic product growth," *Quarterly Journal of Economics*, 2010, *125* (4), 1727–1767.

_ , _ , _ , **and** _ , "Multiproduct firms and product turnover in the developing world: Evidence from India," *Review of Economics and Statistics*, 2010, *92* (4), 1042–1049.

Görg, Holger and Aoife Hanley, "Firms' global engagement and management practices," *Economics Letters*, 2017, *155*, 80–83.

Grossman, Gene M and Elhanan Helpman, "Integration versus outsourcing in industry equilibrium," *The Quarterly Journal of Economics*, 2002, *117* (1), 85–120.

_ **and** _ , "Outsourcing versus FDI in industry equilibrium," *Journal of the European Economic Association*, 2003, *1* (2-3), 317–327.

Grossman, Sanford J and Oliver D Hart, "The costs and benefits of ownership: A theory of vertical and lateral integration," *Journal of Political Economy*, 1986, *94* (4), 691–719.

Guadalupe, Maria and Julie Wulf, "The flattening firm and product market competition: The effect of trade liberalization on corporate hierarchies," *American Economic Journal: Applied Economics*, 2010, *2* (4), 105–27.

Hart, Oliver and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 1990, *98* (6), 1119–1158.

Hasan, Rana, Devashish Mitra, and Krishnarajapet V Ramaswamy, "Trade reforms, labor regulations, and labor-demand elasticities: Empirical evidence from India," *The Review of Economics and Statistics*, 2007, *89* (3), 466–481.

Hortaçsu, Ali and Chad Syverson, "Cementing relationships: Vertical integration, foreclosure, productivity, and prices," *Journal of Political Economy*, 2007, *115* (2), 250–301.

Hsieh, C. T. and P. J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Lafontaine, Francine and Margaret Slade, "Vertical integration and firm boundaries: The evidence," *Journal of Economic Literature*, 2007, *45* (3), 629–685.

Legros, Patrick and Andrew F Newman, "A price theory of vertical and lateral integration," *Quarterly Journal of Economics*, 2012, pp. 725–770.

_ **and** _ , "Contracts, ownership, and industrial organization: past and future," *Journal of Law, Economics, and Organization*, 2014, *30* (suppl 1), i82–i117.

_ **and** _ , "Demand-driven integration and divorcement policy," *International Journal of Industrial Organization*, 2017, *53*, 306–325.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Lileeva, Alla and Daniel Trefler, "Improved access to foreign markets raises plant-level productivity... for some plants," *Quarterly Journal of Economics*, 2010, *125* (3), 1051–1099.

Liu, Qing, Larry D Qiu, and Chaoqun Zhan, "Trade liberalization and domestic vertical integration: Evidence from China," *Journal of International Economics*, forthcoming.

Loecker, Jan De and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–2471.

_ , **Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik**, "Prices, markups and trade reform," *Econometrica*, 2016, *84* (4), 445–510.

Marin, Dalia and Thierry Verdier, "Corporate hierarchies and international trade: Theory and evidence," *Journal of International Economics*, 2014, *94* (2), 295–310.

McGowan, Danny, "Digging deep to compete: Vertical integration, product market competition and prices," *The Journal of Industrial Economics*, 2017, *65* (4), 683–718.

McLaren, John, "" Globalization" and vertical structure," *American Economic Review*, 2000, *90* (5), 1239–1254.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Pavcnik, Nina, "Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants," *Review of Economic Studies*, 2002, *69* (1), 245–276.

Rauch, James E, "Networks versus markets in international trade," *Journal of International Economics*, 1999, *48* (1), 7–35.

Schmutzler, Armin, "Competition and investment – A unified approach," *International Journal of Industrial Organization*, 2013, *31* (5), 477–487.

Shu, Pian and Claudia Steinwender, "The impact of trade liberalization on firm productivity and innovation," in Josh Lerner and Scott Stern, eds., *NBER Book Series Innovation Policy and the Economy*, Vol. 19 2019.

Stiebale, Joel and Dev Vencappa, "Acquisitions, Markups, Efficiency, and Product Quality: Evidence from India," *Journal of International Economics*, 2018, *112*, 70–87.

Syverson, Chad, "What determines productivity?," *Journal of Economic Literature*, 2011, *49* (2), 326–365.

Topalova, Petia and Amit Khandelwal, "Trade liberalization and firm productivity: The case of India," *Review of economics and statistics*, 2011, *93* (3), 995–1009.

Vives, Xavier, "Innovation and competitive pressure," *Journal of Industrial Economics*, 2008, *56* (3), 419–469.

Williamson, Oliver E, "Markets and hierarchies," *New York*, 1975, pp. 26–30.

_ , *The economic intstitutions of capitalism*, Simon and Schuster, 1985.

Wooldridge, Jeffrey M, "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics Letters*, 2009, *104* (3), 112–114.
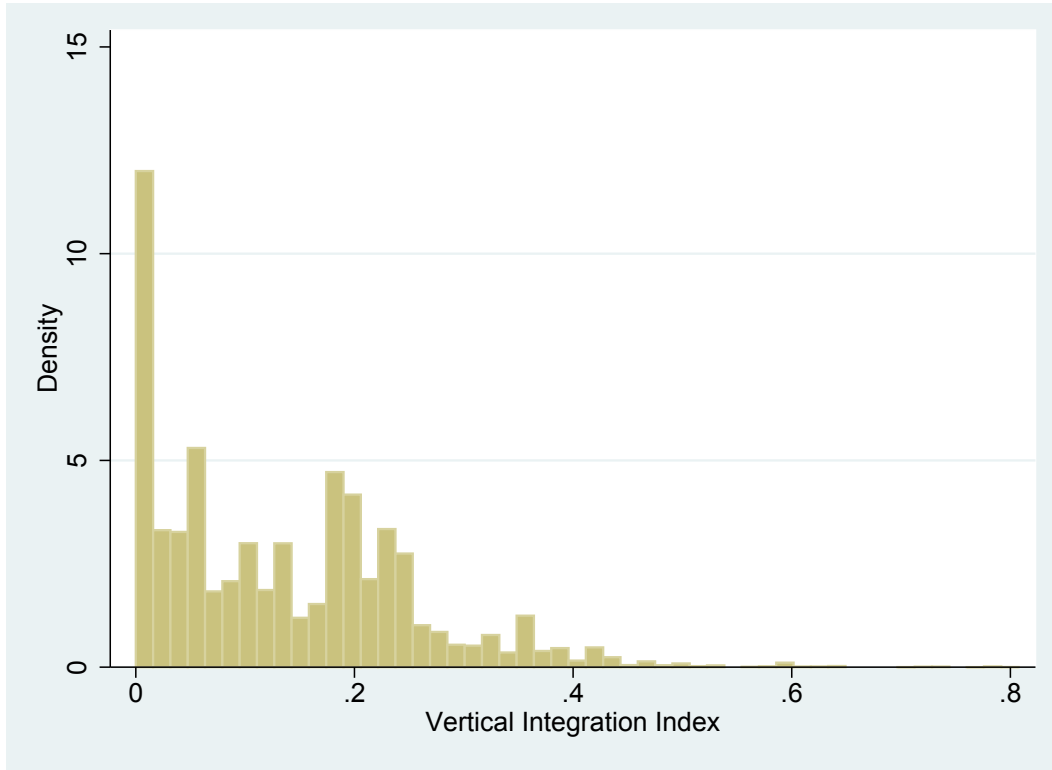
Figure 1: Vertical Integration Index

Table 1: Summary Statistics

Variable	Definition	Mean	Median	Std.dev.
v_{fjt}^p	vertical integration index	0.134	0.116	0.115
$tariff_{jt}$	most-favored nations (MFN) tariff	0.332	0.297	0.279
$input\_tariff_{jt}$	MFN tariffs weighted by IO transaction coefficients	0.084	0.064	0.078
$investment\_intensity_{jt}$	investment / sales (average at the industry-level)	0.061	0.057	0.032
$R\&D\_intensity_{jt}$	R&D / sales (average at the industry-level)	0.002	0.001	0.002
$industry\_size_{jt}$	log(sales), (average at the industry-level)	6.283	6.182	0.720
$homogenous\_good_p$	=1 if homogenous good (Rauch 1999 classification)	0.450	0.000	0.497
$exports_{ft}$	export / sales	0.094	0.004	0.195
$imports_{ft}$	imports / input expenditures	0.239	0.093	0.319
log_sales_{ft}	log(sales),	6.477	6.429	2.049
$log\_R\&D_{ft}$	log (R&D expenditures + 1)	0.584	0.000	1.273
$markup_{fpt}$	price / marginal cost	2.828	1.339	4.831
$marginal\_cost_{fpt}$	marginal cost, deviation from product-year average	0.000	-0.002	149.1

f, p, j denote variable measured at firm, product and industry level respectively. t represents year.

Table 2: Tariffs over time

	output tariff			input tariff		
year	mean	median	sd	mean	median	sd
1989	1.119	1.168	0.485	0.308	0.295	0.186
1990	0.964	0.968	0.387	0.263	0.254	0.154
1991	0.797	0.730	0.289	0.217	0.220	0.121
1992	0.632	0.609	0.220	0.170	0.186	0.089
1993	0.573	0.556	0.213	0.153	0.169	0.081
1994	0.514	0.502	0.202	0.135	0.153	0.072
1995	0.461	0.423	0.212	0.120	0.136	0.063
1996	0.410	0.370	0.227	0.104	0.119	0.055
1997	0.325	0.298	0.188	0.085	0.095	0.044
1998	0.341	0.321	0.180	0.078	0.084	0.044
1999	0.355	0.343	0.166	0.083	0.090	0.047
2000	0.356	0.366	0.133	0.082	0.094	0.047
2001	0.349	0.339	0.149	0.080	0.088	0.044
2002	0.318	0.298	0.140	0.072	0.078	0.042
2003	0.293	0.250	0.140	0.083	0.093	0.047
2004	0.318	0.300	0.139	0.091	0.106	0.051
2005	0.212	0.156	0.170	0.055	0.060	0.034
2006	0.190	0.124	0.178	0.045	0.046	0.029
2007	0.196	0.132	0.173	0.055	0.057	0.034
2008	0.153	0.085	0.173	0.038	0.036	0.030
2009	0.159	0.086	0.185	0.040	0.038	0.030
2010	0.148	0.082	0.172	0.037	0.036	0.029
2011	0.151	0.082	0.180	0.041	0.039	0.030
all	0.329	0.294	0.272	0.085	0.064	0.078

Table 3: Changes in output tariffs and industrial characteristics

Industrial Characteristics	output tariffs						input tariffs					
	1989-1992	1992-1997	1997-2002	2002-2007	2007-2012	1989-2012	1989-1992	1992-1997	1997-2002	2002-2007	2007-2012	1989-2012
Log average wage	-0.035	-0.104	0.041	-0.072***	-0.011	-0.178*	-0.067***	-0.049***	-0.008	-0.006	-0.012***	-0.166***
	(0.113)	(0.097)	(0.045)	(0.017)	(0.009)	(0.100)	(0.023)	(0.010)	(0.008)	(0.010)	(0.003)	(0.037)
Share of non-production workers	0.574	-0.979	0.581*	-0.135	-0.042*	-0.289	-0.237	-0.230***	-0.049	0.004	-0.025*	-0.754**
	(0.768)	(0.638)	(0.297)	(0.112)	(0.179)	(0.085)	(0.081)	(0.074)	(0.014)	(0.311)	(0.025)	(0.744)
Capital labour ratio	-0.008	-0.025	0.002	-0.001	-0.001	-0.046	0.013	-0.000	0.000	0.000*	-0.000	0.011
	(0.040)	(0.020)	(0.003)	(0.001)	(0.000)	(0.049)	(0.010)	(0.004)	(0.001)	(0.000)	(0.000)	(0.020)
Log output	-0.008	-0.018	0.022**	0.006	-0.001	0.007	-0.014	-0.007	-0.003	0.004	-0.000	-0.009
	(0.010)	(0.005)	(0.003)	(0.003)	(0.001)	(0.022)						
Factory size	-0.000	0.000	0.000	-0.000	0.000	-0.000	0.000***	0.000	0.000	0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log employment	0.004	0.003	0.011	0.007	-0.002	0.028	-0.010	-0.003	-0.004	0.005	0.001	0.003
	(0.029)	(0.010)	(0.008)	(0.006)	(0.002)	(0.030)	(0.013)	(0.006)	(0.003)	(0.003)	(0.002)	(0.025)
Growth in output	0.105	-0.180	-0.047**	-0.020	0.001	-0.014	0.115	0.020	0.009	-0.029	-0.001	-0.010
	(0.250)	(0.166)	(0.021)	(0.023)	(0.001)	(0.016)	(0.086)	(0.022)	(0.012)	(0.021)	(0.002)	(0.026)
Growth in employment	-0.537	-0.134	-0.056***	-0.018	0.001	-0.023	-0.285	-0.002	0.001	-0.056**	-0.003	-0.027
	(0.564)	(0.113)	(0.020)	(0.021)	(0.001)	(0.015)	(0.244)	(0.048)	(0.009)	(0.025)	(0.002)	(0.021)
Observations	68	68	69	67	66	65	50	50	43	45	45	43

Bivariate regressions of change in output tariffs on each industry characteristic, weighted by the number of factories in each industry.

Apart from growth in output and employment, industrial characteristics are at beginning of period values, i.e. 1989 for period 1989-1992, 1992 for period 1992-1997,etc.

Growth in output and employment is lagged relative to sample period, i.e. 1992-1997 for period 1997-2002.

Robust standard errors in parentheses

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

Table 4: Tariffs and lagged productivity

	1989-1992	1992-1997	1997-2002	2002-2007	2007-2012	1989-2012
		Panel A: Output tariffs				
TFP_{t-1}	0.066	-0.002	0.064*	-0.016	0.011*	0.004
	(0.076)	(0.007)	(0.033)	(0.015)	(0.006)	(0.013)
		Panel B: Input tariffs				
TFP_{t-1}	0.033	0.002	0.005	-0.010	0.004*	0.001
	(0.022)	(0.003)	(0.006)	(0.008)	(0.002)	(0.010)
Observations	132	216	225	222	212	860
Industry fixed effects	yes	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes	yes

Standard errors, in parentheses, are clustered by industry.

All regressions are weighted by the number of firms in each industry

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

Table 5: Vertical integration and tariffs, firm-level

	(1)	(2)	(3)	(4)	(5)
tariff	-0.022**	-0.017**	-0.025***	-0.025***	-0.025***
	(0.010)	(0.007)	(0.009)	(0.009)	(0.009)
input tariff			0.057	0.047	0.047
			(0.044)	(0.043)	(0.043)
Investment intensity				0.164***	0.164***
				(0.047)	(0.047)
R&D intensity				0.497	0.498
				(0.517)	(0.517)
Industry size				-0.001	-0.001
				(0.005)	(0.005)
Export share					0.001
					(0.002)
Import share					-0.000
					(0.001)
Observations	66664	66664	66664	66664	66664
Firm fixed effects	yes	yes	yes	yes	yes
Sector-year fixed effects	no	yes	yes	yes	yes
Year fixed effects	yes	no	no	no	no

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{it}, the vertical integration index in a firm's main industry.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Size, Investment and R&D intensity are measured at the industry-level.

Export and import share are measured at the firm-level.

Table 6: Vertical integration and tariffs, firm-product-level

	(1)	(2)	(3)	(4)	(5)	(6)
tariff	-0.027**	-0.024***	-0.034***	-0.064***	-0.024***	-0.042***
	(0.012)	(0.007)	(0.009)	(0.014)	(0.006)	(0.010)
input tariff			0.072			0.138***
			(0.046)			(0.049)
Observations	121443	121443	121443	81275	81275	81275
Firm-product fixed effects	yes	yes	yes	yes	yes	yes
Sector-year fixed effects	no	yes	yes	no	yes	yes
Firm-year fixed effects	no	no	no	yes	yes	yes
Year fixed effects	yes	no	no	yes	no	no

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Table 7: Vertical integration and tariffs: differentiated vs. homogenous goods

	(1)	(2)	(3)	(4)
Product category	homogenous	differentiated	homogenous	differentiated
tariff	-0.086***	-0.013**	-0.112***	-0.009
	(0.021)	(0.006)	(0.028)	(0.007)
input tariff	0.488**	-0.032	0.956***	-0.061
	(0.199)	(0.032)	(0.245)	(0.042)
Observations	55063	65466	32193	38057
Firm-product fixed effects	yes	yes	yes	yes
Sector-year fixed effects	yes	yes	yes	yes
Firm-year fixed effects	no	no	yes	yes

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Differentiated and homogenous are defined according to Rauch (1999)'s liberal classification.

Table 8: Vertical integration and tariffs: heterogeneous responses

Sample	(1) all firms	(2) non-domestic	(3) all firms	(4) all firms
tariff	-0.026***	-0.002	-0.002	
	(0.009)	(0.006)	(0.007)	
large firm * tariff	-0.021***			
	(0.005)			
input tariff	0.042	0.675***	0.675***	
	(0.047)	(0.217)	(0.212)	
large firm * input tariff	0.078***			
	(0.019)			
domestic * tariff			-0.031***	-0.046***
			(0.007)	(0.016)
domestic * input tariff			-0.614***	0.248***
			(0.204)	(0.085)
Observations	121443	6204	121443	121443
Firm-product fixed effects	yes	yes	yes	yes
Sector-year fixed effects	yes	yes	yes	no
Product-year fixed effects	no	no	no	yes

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

large firm indicates firm with sales above the industry median.

domestic indicates firms with export shares of at most 0.5.

non-domestic indicates firms with export shares of more than 0.5.

large, domestic, and non-domestic are based on firm-specific averages.

Table 9: Vertical integration and tariffs: non-linearities

	(1) v	(2) ln(v)	(3) ln(v+1)	(4) ln(v+1)
tariff	-0.079*** (0.025)			
tariff^2	0.034*** (0.011)			
input tariff	0.401*** (0.088)			
input tariff^2	-0.655*** (0.120)			
ln(tariff)		-0.270*** (0.102)		
ln(input tariff)		0.885*** (0.087)		
ln(tariff+1)			-0.052*** (0.015)	-0.104** (0.045)
ln(input tariff+1)			0.076 (0.048)	0.457*** (0.088)
ln(tariff+1)^2				0.081** (0.036)
ln(input tariff+1)^2				-1.023*** (0.163)
Observations	121443	103149	121443	121443

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable in column (1) is v_{ijt}.

The dependent variable is $\ln(v_{ijt})$ $(\ln(v_{ijt} + 1))$ in column 2 ((3) and (4)).

v_{ijt} is the vertical integration index of a firm-product in year t.

tariff is the MFN tariff rate.

input tariff is the weighted tariff rate of supplying industries.

Table 10: Vertical integration, marginal costs, markups and TFP

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
	marg. cost	markup	marg. cost	markup	marg. cost	markup	TFPQ
vertical integration index	-0.488***	0.292**	-1.118***	1.121***	-0.790**	0.912***	0.537**
	(0.149)	(0.136)	(0.301)	(0.275)	(0.336)	(0.337)	(0.259)
Observations	114,621	114,621	114,621	114,621	82,005	82,005	29,742
Firm-product fixed effects	yes	yes	yes	yes	yes	yes	no
Product-year fixed effects	no	no	yes	yes	no	no	no
Firm-year fixed effects	no	no	no	no	yes	yes	no
Year fixed effects	yes	yes	no	no	no	no	yes
Firm fixed effects	no	no	no	no	no	no	yes

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable in columns (1), (3) and (5) is the log of marginal costs at the firm-product-level.

The dependent variable in columns (2), (4) and (6) is the log of the markup at the firm-product-level.

The dependent variable in column (7) is a measure of physical TFP at the firm-level.

Appendix A: Additional Tables

Table A1: Alternative estimation methods

| | (1) | (2) | (3) | (4) |
| | Tobit | Fractional Probit | Fractional Logit | Exponential pseudo ML |
Method				
tariff	-0.076***	-0.091***	-0.090***	-0.721***
	(0.002)	(0.003)	(0.003)	(0.025)
input tariff	0.367***	0.456***	0.443***	3.594***
	(0.009)	(0.012)	(0.012)	(0.095)
Observations	123340	123340	123340	123340
Industry and year dummies	yes	yes	yes	yes

Numbers are average marginal effects in columns (1)-(3) and coefficients in column (4).

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Table A2: Excluding corporate groups

	(1)	(2)	(3)	(4)	(5)	(6)
tariff	-0.025**	-0.026***	-0.038***	-0.039***	-0.020***	-0.034***
	(0.012)	(0.008)	(0.011)	(0.013)	(0.008)	(0.010)
input tariff			0.104*			0.132**
			(0.053)			(0.063)
Observations	81398	81398	81398	51991	51991	51991
Firm-product fixed effects	yes	yes	yes	yes	yes	yes
Sector-year fixed effects	no	yes	yes	no	yes	yes
Firm-year fixed effects	no	no	no	yes	yes	yes
Year fixed effects	yes	no	no	yes	no	no

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Table A3: Including multinationals

	(1)	(2)	(3)	(4)	(5)	(6)
tariff	-0.027**	-0.024***	-0.033***	-0.063***	-0.027***	-0.043***
	(0.010)	(0.006)	(0.009)	(0.013)	(0.006)	(0.009)
input tariff			0.066			0.103**
			(0.045)			(0.045)
Observations	148994	148983	148983	103642	103618	103310
Firm-product fixed effects	yes	yes	yes	yes	yes	yes
Sector-year fixed effects	no	yes	yes	no	yes	yes
Firm-year fixed effects	no	no	no	yes	yes	yes
Year fixed effects	yes	no	no	yes	no	no

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Table A4: Products aggregated to the 3-digit industry level

	(1)	(2)	(3)	(4)	(5)	(6)
tariff	-0.023**	-0.021***	-0.027***	-0.056***	-0.024***	-0.040***
	(0.010)	(0.007)	(0.009)	(0.012)	(0.006)	(0.009)
input tariff			0.029			0.135***
			(0.043)			(0.049)
Observations	88523	88508	88508	35360	35276	35163
Firm-product fixed effects	yes	yes	yes	yes	yes	yes
Sector-year fixed effects	no	yes	yes	no	yes	yes
Firm-year fixed effects	no	no	no	yes	yes	yes
Year fixed effects	yes	no	no	yes	no	no

Robust standard errors, clustered two-way by firm and industry, in parentheses.

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

The dependent variable is v_{ijt}, the vertical integration index of a firm-product.

tariff is the MFN tariff rate. input tariff is the weighted tariff rate of supplying industries.

Appendix B: Estimation of productivity, markups, and marginal costs

To estimate productivity, markups, and marginal costs, we follow the methodology recently introduced by De Loecker et al. (2016), henceforth LGKP. This method accounts for endogeneity of productions inputs similar to standard techniques in the productivity literature (Ackerberg et al., 2015; Levinsohn and Petrin, 2003; Olley and Pakes, 1996). In addition, it relies on the availability of quantities and prices at the product level to separate true efficiency from revenue based productivity. As most (if not all) firm-product-level data sets, Prowess does not include complete information on prices of all inputs and has no information about how inputs are allocated across products for multi-product firms. The main innovations of the LGKP approach are the introduction of a control function for unobserved input prices and a method to recover the allocation of inputs across products. We describe the methodology below.

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

$$Q_{ijt} = F_j(M_{ijt}, K_{ijt}, L_{ijt})\Omega_{it} \tag{11}$$

where Q_{ijt} denotes physical output, M_{ijt} denotes a freely adjustable input (materials in our case), K_{ijt} and L_{ijt} are capital stock and labor input respectively and Ω_{it} denotes total factor productivity (TFP). All production inputs are defined in physical units. A firm minimizes costs product-by-product subject to the production function and input costs.

As shown by De Loecker and Warzynski (2012) and LGKP, this cost minimization yields an expression for the firm-product specific markup as:

$$\mu_{ijt} = \left(\frac{P_{ijt}Q_{ijt}}{W_{ijt}^M M_{ijt}}\right)\frac{\partial Q_{ijt}(.)}{\partial M_{ijt}}\frac{M_{ijt}}{Q_{ijt}} = \frac{\theta_{ijt}^M}{\alpha_{ijt}^M} \tag{12}$$

where P_{ijt} denotes the output price, W_{ijt}^M is the input price of materials, α_{ijt}^M is the ratio of expenditures on input M_{ijt} to a product's revenue and θ_{ijt}^M is the elasticity of output with respect to this input. Intuitively, the output elasticity equals the input's revenue share only in the case of perfect competition. Under imperfect competition, the output elasticity will exceed the revenue share. θ_{ijt}^M can be estimated from a production function and α_{ijt}^M can be calculated, once the allocation of inputs across a firms' product has been estimated. Marginal costs (mc_{ijt}) can then be calculated as the

ratio of observed prices to estimated markups:

$$mc_{ijt} = \frac{P_{ijt}}{\mu_{ijt}} \tag{13}$$

The basis for productivity estimation is the logarithmic version of equation (11) with an additive error term, ϵ_{ijt} which captures measurement error:

$$q_{ijt} = f_j(\boldsymbol{v}_{ijt}; \boldsymbol{\beta}) + \omega_{it} + \epsilon_{ijt} \tag{14}$$

where $\boldsymbol{v}_{ijt}$ denotes a vector of logarithmic physical inputs (capital k_{ijt}, labor l_{ijt} and materials m_{ijt}) allocated to product j and ω_{it} is the log of TFP. For our application, we use a translog production function, hence:

$$\begin{aligned} f_j(\boldsymbol{v}_{ijt}; \boldsymbol{\beta}) =&\beta_l l_{ijt} + \beta_m m_{ijt} + \beta_k k_{ijt} + \beta_{lm} l_{ijt} m_{ijt} + \beta_{lk} l_{ijt} k_{ijt} + \beta_{mk} m_{ijt} k_{ijt} \\ &+ \beta_{ll} l_{ijt}^2 + \beta_{mm} m_{ijt}^2 + \beta_{kk} k_{ijt}^2 + \beta_{lmk} l_{ijt} m_{ijt} k_{ijt} \end{aligned} \tag{15}$$

The translog production function yields a physical output-material elasticity:

$$\theta_{ijt}^M = \beta_m + \beta_{lm} l_{ijt} + \beta_{mk} k_{ijt} + 2\beta_{mm} m_{ijt} + \beta_{lmk} l_{ijt} k_{ijt} \tag{16}$$

which varies across firms within industries and nests a Cobb-Douglas production function as a special case.

Physical inputs can be expressed as $v_{ijt} = \rho_{ijt} + \tilde{v}_{it} - w_{ijt}$ where $\tilde{v}_{it}$ denotes observed input expenditures at the firm-level, ρ_{ijt} is the log of the input share allocated to product j and w_{ijt} denotes the log of an input price index (defined as deviations from industry-specific deflators). When the log of input allocations, ρ_{ijt}, is captured by a function $A(\rho_{ijt}, \tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta})$ and the log of the unobserved input price index, w_{ijt}, are captured by a function $B(w_{ijt}, \rho_{ijt}, \tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta})$, output can be rewritten as a function of firm-specific input expenditures instead of unobserved product-specific input quantities (see LGKP for the exact functional form of $A(.)$ and $B(.)$ for the translog case):

$$q_{ijt} = f_j(\tilde{\boldsymbol{v}}_{ijt}; \boldsymbol{\beta}) + A(\rho_{ijt}, \tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta}) + B(w_{ijt}, \rho_{ijt}, \tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta}) + \omega_{it} + \epsilon_{ijt} \tag{17}$$

Estimation of the parameters of the production function is based on a sample of single product firms for which $A(.)$ can be ignored. Unobserved input prices w_{it} in $B(.)$ are approximated by output prices (p_{it}), market shares (s_{it}), product dummies ($\boldsymbol{D_j}$), and export status (ex_{it}) to account for differences in product quality and local input markets. We also include our vertical integration

measure ($\boldsymbol{vi}_{it}$), as we want to allow for the possibility that vertical integration is associated with different input prices.

Material demand is assumed to be a function of productivity, other inputs, output prices, market share, product, export and vertical integration, hence: $\tilde{m}_{it} = m(\omega_{it}, \tilde{k}_{it}, \tilde{l}_{it}, p_{it}, \boldsymbol{D}_j, s_{it}, ex_{it}, \boldsymbol{vi}_{it})$. Inverting the material demand function yields and expression for productivity: $\omega_{it} = h(\tilde{\boldsymbol{v}}_{it}, \boldsymbol{c}_{it})$ where $\boldsymbol{c}_{it}$ includes all variables from the input demand function except input expenditures.

The use of single product firms induces a further complication of endogenous sample selection since single-product firms might be less productive compared to multi-product firms. Analogous to the exit correction proposed by Olley and Pakes (1996), the probability of remaining a single product firm (SP_{it}) is a function of previous year's productivity and an unobserved productivity cutoff.[29]

For the evolution of productivity, the following law of motion is assumed:

$$\omega_{it} = g(\omega_{i,t-1}, ex_{it}, \boldsymbol{vi}_{i,t-1}, SP_{it}) + \varsigma_{it} \tag{18}$$

In addition to export status and the probability of remaining a single product firm, we allow the evolution of productivity to depend on the degree of vertical integration.

Since for single product firms, we do not face the problem of unobserved input allocation across products and can drop the product-specific subscript of the production function, equation (17) becomes:

$$q_{ijt} = f(\tilde{\boldsymbol{v}}_{ijt}; \boldsymbol{\beta}) + B(w_{ijt}, \rho_{ijt}, \tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta}) + \omega_{it} + \epsilon_{ijt} \tag{19}$$

One can combine $f(.)$, $B(.)$ and $g(.)$ into a function $\phi(\tilde{\boldsymbol{v}}_{ijt}, \boldsymbol{c}_{it})$ such that output can be expressed as a function of observable variables and measurement errors: $q_{it} = \phi(\tilde{\boldsymbol{v}}_{it}, \boldsymbol{c}_{it}) + \epsilon_{it}$.

$\phi(.)$ is approximated by a linear combination of all its elements and a polynomial in all continuous variables. While this expression does not identify any parameters of the production and input price functions, it identifies output net of measurement error ϵ_{it} which is denoted by $\hat{\phi}_{it}$. Productivity can then be expressed as:

$$\omega_{it} = \hat{\phi}_{it} - f(\tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta}) - B(\boldsymbol{c}_{it}, \boldsymbol{c}_{it} \times \tilde{\boldsymbol{v}}_{it}, \boldsymbol{\beta}, \boldsymbol{\delta}) \tag{20}$$

where $\boldsymbol{\delta}$ are the parameters of the input price function to be estimated. For identification of param-

---

[29] SP_{it} is estimated by a Probit regression of a dummy variable for remaining a single-product firm on $\tilde{\boldsymbol{v}}_{i,t-1}$ $\boldsymbol{c}_{i,t-1}$, investment, year and industry dummies.

eters, equation (18) can be used to construct moment conditions:

$$E[\varsigma_{it}(\boldsymbol{\beta}, \boldsymbol{\delta})\boldsymbol{Z}_{it}] = 0 \tag{21}$$

$\boldsymbol{Z}_{it}$ is a vector which includes current values of labour and capital, lagged values of materials and their higher order and interaction terms as they appear in the production function. It further includes lagged values of market shares and prices as well as interactions of lagged prices with lags of production factors and market share. We treat labor as a dynamic input that is characterized by adjustment costs due to the rather rigid Indian labor market. Estimation is undertaken using the GMM procedure suggested by Wooldridge (2009) which is based on moment conditions on the combined error term $\varsigma_{it} + \epsilon_{it}$.

This estimation procedure yields estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$, hence, it identifies all parameters from the production and input price functions. We estimate $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ separately for each industry to allow for industry-specific production technologies and input prices. Under the assumption that $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are the same for multi- and single-product firms within industries, input allocations across products within multi-product firms can be recovered which allows estimation of markups and marginal costs for each firm-product-year. Note that as discussed by LGKP, this assumption does not rule out differences in productivity levels between single- and multi-product. Since productivity is modeled to be factor-neutral, differences in TFP do not imply differences in $\boldsymbol{\beta}$ or output elasticities. The approach also allows for TFP to depend on the number of products which can imply (dis)economies of scope. Under the assumption of a common production technology within industries, one can express predicted output as: $\hat{q}_{ijt} = f(\tilde{\boldsymbol{v}}_{ijt}, \boldsymbol{\beta}, \hat{w}_{ijt}, \rho_{ijt}) + \omega_{it}$ and divide the production function into two parts, f_1 and f_2, such that only f_2 depends on input allocations across products. This yields a system of equation for each firm-year which allows identifying productivity ω_{it} for each firm-year and the input share allocation ρ_{ijt} for each firm-product-year:

$$\hat{q}_{ijt} - f_1(\tilde{\boldsymbol{v}}_{ijt}, \boldsymbol{\beta}, \hat{w}_{ijt}) = f_2(\tilde{\boldsymbol{v}}_{ijt}, \hat{w}_{ijt}, \rho_{ijt}) + \omega_{it} \tag{22}$$

$$\sum_j \exp(\rho_{ijt}) = 1$$