PRICES, MARKUPS AND TRADE REFORM

Jan De Loecker
Pinelopi K. Goldberg
Amit K. Khandelwal
Nina Pavcnik

Prices, Markups and Trade Reform
Jan De Loecker, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik
NBER Working Paper No. 17925
March 2012
JEL No. F1,L1

ABSTRACT

This paper examines how prices, markups and marginal costs respond to trade liberalization. We develop a framework to estimate markups from production data with multi-product firms. This approach does not require assumptions on the market structure or demand curves faced by firms, nor assumptions on how firms allocate their inputs across products. We exploit quantity and price information to disentangle markups from quantity-based productivity, and then compute marginal costs by dividing observed prices by the estimated markups. We use India's trade liberalization episode to examine how firms adjust these performance measures. Not surprisingly, we find that trade liberalization lowers factory-gate prices. However, the price declines are small relative to the declines in marginal costs, which fall predominantly because of the input tariff liberalization. The reason is that firms offset their reductions in marginal costs by raising markups. This limited pass-through of cost reductions attenuates the reform's impact on prices. Our results demonstrate substantial heterogeneity and variability in markups across firms and time and suggest that producers benefited relative to consumers, at least immediately after the reforms. To the extent that higher firm profits lead to the new product introductions and growth, long-term gains to consumers may be substantially higher.

Jan De Loecker
Department of Economics
307 Fisher Hall
Princeton University
Princeton, NJ 08544-1021
and NBER
jdeloeck@princeton.edu

Pinelopi K. Goldberg
Yale University
Department of Economics
37 Hillhouse Ave.
P.O. Box 208264
New Haven, CT 06520-8264
and NBER
penny.goldberg@yale.edu

Amit K. Khandelwal
Graduate School of Business
Columbia University
Uris Hall 606, 3022 Broadway
New York, NY 10027
and NBER
ak2796@columbia.edu

Nina Pavcnik
Department of Economics
6106 Rockefeller Center
Dartmouth College
Hanover, NH 03755
and NBER
nina.pavcnik@dartmouth.edu

1 Introduction

Trade reforms have the potential to deliver substantial benefits to economies by forcing a more efficient allocation of resources. A large body of theoretical and empirical literature has analyzed the mechanisms behind this process. When trade barriers fall, aggregate productivity rises as less productive firms exit and the remaining firms expand (e.g., Melitz (2003) and Pavcnik (2002)) and take advantage of cheaper or previously unavailable imported inputs (e.g., Goldberg et al. (2010a) and Halpern et al. (2011)). Trade reforms also have been shown to reduce firms' markups (e.g., Levinsohn (1993) and Harrison (1994)). Based on this evidence, we should expect trade reforms to exert downward pressure on firm prices. However, because firm prices are rarely observed during the period of a trade reforms, we have little direct evidence on how prices respond to liberalization. This paper fills this gap by developing a unified framework to estimate jointly markups and marginal costs from production data, and examine how prices, and their underlying markup and cost components, adjusted during India's comprehensive trade liberalization.

Our paper makes three main contributions. The first contribution is towards measurement. In order to infer markups, one requires estimates of production functions. Typically, these estimates have well-known biases if researchers use revenue data rather than quantity data to estimate production functions. Estimates of "true" productivity (or marginal costs) are confounded by demand shocks and markups, and these biases may be severe (see Foster et al. (2008)). De Loecker (2011) demonstrates that controlling for demand shocks substantially attenuates the productivity increases in response to trade reforms in the European Union textile industry. As a result, approaches to infer markups from production data may be problematic if one uses revenue information. We alleviate this concern by exploiting detailed firm-level data from India that contain the prices and quantities of firms' products over time to infer markups from a quantity-based production function.

The second contribution is to develop a unified framework to estimate markups, marginal costs and productivity of multi-product firms across a broad set of manufacturing industries. Since these performance measures are unobserved, we must impose some structure on the data. However, our approach requires substantially fewer assumptions than is typically required in industrial organization studies. Crucially, our approach does not require assumptions on consumer demand, market structure or the nature of competition. This flexibility enables us to infer the full distribution of markups across firms and products over time in different manufacturing sectors. Moreover, since prices are directly observed, we can directly recover marginal costs from our markup estimates. Our approach is quite general and since data containing this level of detail are becoming increasingly available, this methodology will be useful to researchers studying other countries and industries.[1] The drawback of this approach is that we are unable to perform counterfactual simulations and welfare analysis since we do not explicitly model consumer demand and firm pricing behavior.

Third, existing studies that have analyzed the impact of trade reforms on markups have focused exclusively on the competitive effects from declines in output tariffs (e.g., Levinsohn (1993) and

---

[1]To our knowledge, the following countries contain firm-level panel data that record the prices and quantities of the products that firms manufacture: France, U.S., Colombia, Denmark, Slovenia, and Hungary.

Harrison (1994)). Comprehensive reforms also lower tariffs on imported inputs and previous work, particularly on India, has emphasized this aspect of trade reforms (e.g., see Goldberg et al. (2009)). These two tariff reductions represent distinct shocks to domestic firms. Lower output tariffs increase competition by changing the residual demand that firms face. Conversely, firms benefit from lower costs of production when input tariffs decline. It is therefore important to account for both forms of liberalization in order to understand the total impact of trade reforms on prices and markups. In particular, the decline in markups depends on the extent to which firms pass these cost savings to consumers, which is influenced both by the market structure and the demand curve that firms face. For example, in models with monopolistic competition and CES demand, markups are constant and so by assumption, pass-through of tariffs on prices is complete. Arkolakis et al. (2012) demonstrate that several of the influential trade models assume constant markups and by doing so, abstract away from the markup channel as a potential source of gains from trade. This is the case in Ricardian models that assume perfect competition, such as Eaton and Kortum (2002), and models with monopolistic competition such as Krugman (1980) and its heterogeneous firm extensions like Melitz (2003). There are models that can account for variable markups by imposing some structure on demand and market structure (e.g., Bernard et al. (2003), Melitz and Ottaviano (2008), Feenstra and Weinstein (2010) and Edmonds et al. (2011)). While these studies allow for richer patterns of markup adjustment, the empirical results on markups and pass-through ultimately depend on the underlying parametric assumptions. Ideally, we want to understand how trade reforms affect markups without having to rely on explicit parametric assumptions of the demand systems and/or market structures, which themselves may change with trade liberalization.

The structure of our analysis is as follows. We use production data to infer markups by exploiting the optimality of firms' variable input choices. Our approach is based on De Loecker and Warzynski (2012), but we extend their methodology to account for multi-product firms and to take advantage of observable price data. The key assumption we need to infer markups is simply that firms minimize cost; then, markups are the deviation between the elasticity of output with respect to a variable input and that input's share of total revenue.[2] We obtain this output elasticity from estimates of production functions across many industries. In contrast to many earlier studies, we utilize physical quantity data rather than revenues to estimate the production functions.[3] This alleviates the concern that the production function estimation is contaminated by prices, yet presents different challenges that we discuss in detail in Section 4. Most importantly, using physical quantity data forces us to conduct the analysis at the product level since without a demand system to aggregate across products, prices and physical quantities are only defined at the product level. We also confront the potential concern that (unobserved) input prices vary across firms. This concern is

---

[2]Our framework explicitly accounts for production that relies on fixed factors that may be costly to adjust over time. These include machinery and capital goods, and potentially also labor. This is particularly important in a country like India that has relatively strong labor regulations (Besley and Burgess (2004)) as well as capital market distortions. Our approach requires at least one flexible input that is adjustable in the short run; in our case, this variable input is materials.

[3]Foster et al. (2008) also use quantity data in their analysis of production functions, but they focus on a set of homogeneous products.

usually ignored in earlier work, but we show how to deal with this issue by exploiting variation in observed output prices and input tariffs.

This approach calls for an explicit treatment of multi-product firms. We show how to exploit data on single-product firms along with a sample selection correction to obtain consistent estimates of the production functions. The benefit of using single-product firms in the production function estimation stage is that we do not require assumptions on how firms allocate inputs across products, something we do not observe in our data.[4] While this approach may seem strong, it simply assumes that the *physical* relationship between inputs and outputs is the same for single- and multi-product firms that manufacture the same product. That is, a single-product firm uses the same technology to produce a rickshaw as a multi-product firm. This assumption is already implicitly employed in all previous work that pools data across single- and multi-product firms (e.g., Olley and Pakes (1996) or Levinsohn and Petrin (2003)). Moreover, this assumption of the same physical production structure does not rule out economies of scope, which can operate through lower input prices for multi-product firms (for example, due to discounts associated with bulk purchases of materials), or higher (factor-neutral) productivity of multi-product firms. Once we estimate the production functions from the single-product firms, we show how to back out the quantity-based productivity of multi-product firms and their allocation of inputs across its products.

Our estimation of the output elasticities of the production function provides reasonable results. A noteworthy feature of the production function results is that the estimation on quantity data yields economies of scale in most sectors. This finding is consistent with Klette and Griliches (1996) and De Loecker (2011) who showed that estimation based on revenue data leads to a downward bias in the estimated returns to scale. We obtain the markups for each product manufactured by firms by dividing the output elasticity of materials by the materials share of total revenue (which is in the data).[5] Then, by dividing prices by the markups, we obtain marginal costs.

The performance measures are correlated in intuitive ways and are consistent with recent heterogeneous models of multi-product firms. Markups (marginal costs) are positively (negatively) correlated with a firm's underlying productivity. More productive firms manufacture more products, but firms have lower markups (higher marginal costs) on products that are farther from their core competency.[6] Foreshadowing the impact of the trade liberalizations, we find that changes in marginal costs are not perfectly reflected by changes in prices because of variable markups.

We then analyze how prices, marginal costs, and markups adjust during India's trade liberal-

---

[4]Suppose a firm manufactures three products using raw materials, labor and capital. To our knowledge, no dataset covering manufacturing firms reports information on how much of each input is used for each product. One way around this problem is to assume input proportionality. For example, Foster et al. (2008) allocate inputs based on products' revenue shares. This approach is valid under perfect competition or the assumption of constant markups across all products produced by a firm. While these assumptions are appropriate for the particular homogenous good industries they study, we study a broad class of differentiated products where these assumption may not apply. Moreover, this study aims to estimate markups without imposing such implicit assumptions.

[5]For multi-product firms, we use the estimated input allocations in the markup calculation.

[6]As noted earlier, we obtain these results without any assumptions on the underlying demand system or market structure. However, these correlations are consistent with Mayer et al. (2011). The correlation between product sales rank and marginal costs are also consistent with Eckel and Neary (2010) and Arkolakis and Muendler (2010).

ization. As has been discussed extensively in earlier work, the nature of India's reform provides an identification strategy that can alleviate the usual endogeneity concerns associated with trade liberalization.[7] The combination of an important trade reform that has statistical benefits and the ability to observe not only prices, but their underlying components, is an important contribution of our analysis. Perhaps not surprisingly, we observe a decline in prices during the reform period. However, the distributions of prices before and after the reform are quite similar. When we account for general sectoral-trends, we estimate that the overall trade liberalization (on average, output and input tariffs fell 62 and 24 percentage points, respectively) lowers prices, on average, by 13 percent relative to an industry that experiences no tariff changes. However, we find that relative marginal costs fall on average by nearly 25 percent. The cost declines are primarily driven by the input tariff liberalization, which is consistent with earlier work demonstrating the importance of imported inputs in India's trade reform. Output tariffs have only a small effect on costs, suggesting that reductions of X-inefficiencies were modest.[8] Since our prices decompose exactly into their underlying cost and markup components, we can show that the reason the relatively large decline in marginal costs did not translate to equally large price declines was because firms raised markups. On average, we find that the trade reform increased relative markups by 11 percent. Firms appear to offset the decline in costs coming from tariff reductions by raising markups (we show that their ability to raise markups even further is mitigated by the pro-competitive impact of output tariff declines). The net effect is that the trade reforms have an attenuated impact on prices.

Overall, our results suggest that in the short run, the most likely beneficiaries of the trade liberalization were domestic Indian firms who were able to benefit from lower production costs while raising markups. The short-run gains to consumers are smaller, especially considering that we observe factory-gate prices rather than retail prices. However, we stress that the additional short-run profits to the firms may have spurred innovation in Indian manufacturing, particularly in the introduction of many new products. These new products accounted for about a quarter of overall manufacturing growth (see Goldberg et al. (2010b)). Furthermore, the new product introductions were concentrated in sectors with disproportionally large input tariff declines thereby allowing firms access to new, previously unavailable imported materials (see Goldberg et al. (2010a)). To the extent that the extra profits we document as a result of the input tariff reductions were used to finance the development of new products, the long-term gains to consumers could be substantially larger. This analysis, however, lies beyond the scope of this current paper.

In addition to the papers discussed earlier, our work is related to a wave of recent papers that focus on productivity in developing countries, such as Bloom and Van Reenen (2007) and Hsieh and Klenow (2009). The low productivity in the developing world is often attributed to lack of competition (Bloom and Van Reenen (2007)) or the presence of policy distortions that result in

---

[7]Many studies have exploited the differential changes in industry tariffs following the 1991 trade liberalization. For example, see Topalova (2010), Sivadasan (2009), Topalova and Khandelwal (2011) and Goldberg et al. (2010a,b).

[8]The relative importance of input and output tariffs is consistent with Amiti and Konings (2007) and Topalova and Khandelwal (2011) who find that firm-level productivity changes in Indonesia and India, respectively, were predominantly driven by input tariff declines.

a misallocation of resources across firms (Hsieh and Klenow (2009)). Against this background, it is natural to ask whether there is any evidence that an increase in competition or a removal of distortions increases productivity. India's reforms are an excellent context to study these questions because of the nature of the reform and the availability of detailed data. Trade protection is a policy distortion that distorts resource allocation. Limited competition benefits some firms relative to others, and the high input tariffs are akin to the capital distortions examined by Hsieh and Klenow (2009). Our results suggest that the removal of barriers on inputs indeed lowered production costs. So indeed, we do find that India's trade reforms delivered productivity gains. However, the overall picture is more nuanced as firms do not appear to pass the entirety of the cost savings to consumers via lower prices. Our findings highlight the importance of jointly studying changes prices, markups and costs to understand the full distributional consequences of trade liberalization.

The remainder of the paper is organized as follows. In the next section, we provide a brief overview of India's trade reform and the data used in the analysis. In Section 3, we lay out the general empirical framework that allows us to estimate markups, productivity and marginal costs. In section 4, we explain the estimation and identification strategy employed in our analysis, paying particular attention to issues that arise specifically in the context of multi-product firms. Section 5 presents the results and Section 6 concludes.

2   Data and Trade Policy Background

We begin by describing the Indian data since the nature of the data dictates our empirical methodology. We also describe key elements of India's trade liberalization that are important for our identification strategy. Given that the Indian trade liberalization has been described in a number of papers (including several by a subset of the present authors), we keep the discussion of the reforms brief.

2.1   Production and Price data

We use the Prowess data that is collected by the Centre for Monitoring the Indian Economy (CMIE). Prowess includes the usual set of variables typically found in firm-level production data, but has important advantages over the Annual Survey of Industries (ASI), India's manufacturing census. First, unlike the repeated cross section in the ASI, Prowess is a panel that tracks firm performance over time. Second, the data span India's trade liberalization from 1989-2003. Third, Prowess records detailed product-level information for each firm. This enables us to distinguish between single-product and multi-product firms, and track changes in firm scope over the sample period. Fourth, Prowess collects information on quantity and sales for each reported product, so we can construct the prices of each product a firm manufactures. These advantages make Prowess particularly well-suited for understanding the mechanisms of firm-level adjustments in response to trade liberalizations that are typically hidden in other data sources, and deal with measurement issues that arise in most studies that estimate production functions.

Prowess enables us to track firms' product mix over time because Indian firms are required by the 1956 Companies Act to disclose product-level information on capacities, production and sales in their annual reports. As discussed extensively in Goldberg et al. (2010b), several features of the database give us confidence in its quality. Product-level information is available for 85 percent of the manufacturing firms, which collectively account for more than 90 percent of Prowess' manufacturing output and exports. Since product-level information and overall output are reported in separate modules, we can cross check the consistency of the data. Product-level sales comprise 99 percent of the (independently) reported manufacturing sales.[9] We refer the reader to Goldberg et al. (2010a,b) for a more detailed discussion of the data.

The definition of a product is based on the CMIE's internal product classification. There are a total of 1,526 products in the sample for estimation.[10] Table 1 reports basic summary statistics by two-digit NIC (India's industrial classification system) sector. As a comparison, the U.S. data used by Bernard et al. (2010a), contain approximately 1,500 products, defined as five-digit SIC codes across 455 four-digit SIC industries. Thus, our definition of a product is slightly more detailed than in earlier work that has focused on the U.S. Table 2 provides a few examples of products available in our data set. In our terminology, we will distinguish between "sectors" (which correspond to two-digit NIC aggregates), "industries" (which correspond to four-digit NIC aggregates) and "products" (which correspond to twelve-digit codes); we emphasize that the "product" definition is available at a highly disaggregated level, so that unit values can plausibly interpreted as "prices" in our application.

The data also have some disadvantages. Unlike Census data, the CMIE database is not well suited for understanding firm entry and exit. However, Prowess contains mainly medium large Indian firms, so entry and exit is not necessarily an important margin for understanding the process of adjustment to increased openness within this subset of the manufacturing sector.[11]

We complement the production data with tariff rates from 1987 to 2001. The tariff data are reported at the six-digit Harmonized System (HS) level and were combined by Topalova (2010). We pass the tariff data through India's input-output matrix for 1993-94 to construct input tariffs. We concord the tariffs to India's national industrial classification (NIC) schedule developed by Debroy and Santhanam (1993). Formally, input tariffs are defined as $\tau_{it}^{input} = \sum_k a_{ki} \tau_{kt}^{output}$, where τ_{kt}^{output} is the tariff on industry k at time t, and a_{ki} is the share of industry k in the value of industry i.

---

[9] We deflate all nominal values for our analysis. Materials are deflated by the wholesale price index of primary articles. Wages are deflated by the overall wholesale price index for manufacturing. Capital stock is fixed assets deflated by sector-specific wholesale price indexes. Unit values are also deflated by sector-specific wholesale price indexes.

[10] We have fewer products than in Goldberg et al. (2010b) because we require non-missing values for quantities and revenues rather than just a count of products.

[11] Firms in Prowess account for 60 to 70 percent of the economic activity in the organized industrial sector and comprise 75 percent of corporate taxes and 95 percent of excise duty collected by the Government of India (CMIE).

2.2 India's Trade Liberalization

A key advantage of our approach is that we examine the impact of openness by relying on changes in trade costs induced by a large-scale trade liberalization. India's post-independence development strategy was one of national self-sufficiency and heavy government regulation of the economy. India's trade regime was amongst the most restrictive in Asia, with high nominal tariffs and non-tariff barriers. In response to a balance-of-payments crisis, India launched a dramatic liberalization of the economy as part of an IMF structural adjustment program in August 1991. An important part of this reform was to abandon the extremely restrictive trade policies it had pursued since independence.

Several features of the trade reform are crucial to our study. First, the external crisis of 1991, which came as a surprise, opened the way for market oriented reforms (Hasan et al. (2007)).[12] The liberalization of the trade policy was therefore unanticipated by firms in India and not foreseen in their decisions prior to the reform. Moreover, reforms were passed quickly as sort of a "shock therapy" with little debate or analysis to avoid the inevitable political opposition (see Goyal (1996)). Industries with the highest tariffs received the largest tariff cuts implying that both the average and standard deviation of tariffs across industries fell.

While there was significant variation in the tariff changes across industries, Topalova and Khandelwal (2011) show that tariff changes through 1997 were uncorrelated with pre-reform firm and industry characteristics such as productivity, size, output growth during the 1980s and capital intensity. The tariff liberalization does not appear to have been targeted towards specific industries and appears until 1997 relatively free of usual political economy pressures. Therefore, our findings are less likely to suffer from the usual concerns about the endogeneity of trade policy or correlation of trade policy with pre-existing industry or firm-level trends. We again refer the reader to previous publications that have used this trade reform for a detailed discussion (Topalova and Khandelwal (2011); Topalova (2010); Sivadasan (2009); Goldberg et al. (2010a,b)).

3 Production Technology, Markups and Costs

This section describes the framework to estimate markups and marginal costs from firm-level production data. Our approach to recovering markups follows De Loecker and Warzynski (2012).[13] The main difference from their empirical setting is that we use product-level quantity and price information, rather than firm-level deflated sales, and this enables us to identify markups for each

---

[12]Some commentators (e.g., Panagariya (2008)) noted that once the balance of payments crisis ensued, market-based reforms were inevitable. While the general direction of the reforms may have been anticipated, the precise changes in tariffs were not. Our empirical strategy accounts for this shift in broad anticipation of the reforms, but exploits variation in the sizes of the tariff cuts across industries.

[13]This approach to infer markups from production data began with Hall (1986). An alternative approach to estimating markups would be to assume a price setting behavior of firms and a particular utility structure of consumers and estimate demand curves (e.g., Berry et al. (1995) or Goldberg (1995)). However, we explicitly want to avoid taking a stand on these primitives. This approach is also less feasible when inferring markups across the entire manufacturing sector. We therefore believe that inferring markups from production data is the appropriate path in this setting, and one that can be easily implemented in a variety of contexts.

firm-product-year triplet observation. However, we are forced to confront a number of new issues in order to implement our approach. Once we obtain markups, we compute marginal costs by simply dividing the observed prices by the estimated markups.

Consider the following production function for firm f that manufactures a product:

$$Q_{ft} = F_t(X_{ft}) \exp(\omega_{ft}) \tag{1}$$

where Q is physical output and X is a vector of inputs. To simplify the exposition, we assume for now that this producer is a single-product firm. In Section 4, we explicitly deal with the empirical challenges that arise in the context of multi-product firms. The production function as specified above is general; we can consider alternative functional forms for F and allow the production technology to vary over time. There are only two assumptions that are essential for the subsequent analysis. First, productivity ω enters in log-additive form and is Hicks-neutral. Second, we assume that productivity is firm-specific. This second assumption follows a tradition in the trade literature that models productivity along these lines (e.g., Bernard et al. (2010a)). For single-product firms, this assumption is of course redundant.

Although this production function is entirely standard, the advantage of our setting is that we directly observe output (Q) at the product level. This distinguishes our approach of estimating a production function from the traditional literature where researchers rely on deflated sales data. We discuss the estimation and the identification issues in Section 4.

We assume that producers minimize costs. Let $\mathbf{V}_{ft}$ denote the vector of variable inputs used by the firm. We use the vector $\mathbf{K}_{ft}$ to denote dynamic inputs of production. Any input that faces adjustment costs will fall into this category; capital is an obvious one, and our framework will include labor as well. We consider the firm's conditional cost function where we condition on the set of dynamic inputs $\mathbf{K}_{ft}$. The associated Lagrangian function is:

$$L(\mathbf{V}_{ft}, \mathbf{K}_{ft}, \lambda_{ft}) = \sum_{v=1}^{V} P_{ft}^v V_{ft}^v + \mathbf{r}_{ft}\mathbf{K}_{ft} + \lambda_{ft}\left[Q_{ft} - Q_{ft}(\mathbf{V}_{ft}, \mathbf{K}_{ft}, \omega_{ft})\right] \tag{2}$$

where P_{ft}^v and $\mathbf{r}_{ft}$ denote the firm's input prices for the variable inputs $v = 1, \ldots, V$ and the prices of dynamic inputs, respectively. The first order condition for any variable input free of adjustment costs is

$$\frac{\partial L_{ft}}{\partial V_{ft}^v} = P_{ft}^v - \lambda_{ft}\frac{\partial Q_{ft}(.)}{\partial V_{ft}^v} = 0 \tag{3}$$

where the marginal cost of production at a given level of output is λ_{ft} since $\frac{\partial L_{ft}}{\partial Q_{ft}} = \lambda_{ft}$. Rearranging terms and multiplying both sides by $\frac{V_{ft}}{Q_{ft}}$, provides the following expression:

$$\frac{\partial Q_{ft}(.)}{\partial V_{ft}^v}\frac{V_{ft}^v}{Q_{ft}} = \frac{1}{\lambda_{ft}}\frac{P_{ft}^v V_{ft}^v}{Q_{ft}} \tag{4}$$

The left hand side of the above equation represents the elasticity of output with respect to variable

input V_{ft}^v (the "output elasticity"). The approach simply requires one freely adjustable input into production. This becomes important in settings, such as ours, where there are frictions in adjusting capital and labor. Define the markup μ_{ft} as $\mu_{ft} \equiv \frac{P_{ft}}{\lambda_{ft}}$. As De Loecker and Warzynski (2012) show, the cost-minimization condition can be rearranged to write the markup as:

$$\mu_{ft} = \theta_{ft}^v (\alpha_{ft}^v)^{-1} \tag{5}$$

where θ_{ft}^v denotes the output elasticity on variable input V^v and $\alpha_{ft}^v = \frac{P_{ft}^v V_{ft}^v}{P_{ft} Q_{ft}}$ is its revenue, which is observed in the data. This expression forms the basis for our approach. To compute the markup, we need the output elasticity and the share of the input's expenditure in total sales. We obtain the output elasticity by estimating the production function in (1) and obtain the input's revenue share for single-product firms directly from the data.

Markups for multi-product firms are obtained using the exact same expression; the markup is computed for each product j produced by firm f at time t using:

$$\mu_{fjt} = \theta_{fjt}^v (\alpha_{fjt}^v)^{-1} \tag{6}$$

The only difference between expressions (5) and (6) is the presence of the product-specific subscript j. This seemingly small difference complicates the analysis considerably. In contrast to the single product setting, α_{fjt}^v is not directly observed in the data because firms do not report input expenditures by product.[14] Moreover, we require an estimate of the output elasticity for each product manufactured by each multi-product firm. In the next section, we discuss the identification strategy to obtain the output elasticities for both single- and multi-product firms, and how to recover the firm-product specific input expenditures (the α_{fjt}^v) for the multi-product firms.

This approach provides markups that vary at the firm-product-year level. Since prices are observed at this same level of disaggregation, we can simply calculate the marginal costs from the observed price data:

$$mc_{jft} = \frac{P_{fjt}}{\mu_{fjt}}. \tag{7}$$

We believe that this approach to recover markups is quite flexible and suitable for our particular setting where we want to estimate markups for many industries. The approach requires the presence of at least one input that can be freely adjusted, but allows for frictions in the adjustment of capital and labor. This is important in a country like India that has had heavily regulated labor markets and is often associated with labor and capital market distortions. It does not impose assumptions on the returns to scale nor assumptions on the demand and market structure for each industry. Moreover, we do not require knowledge of the user cost of capital, a factor price that is difficult to measure in a country like India with distorted capital markets. This flexible approach allows us to be *a priori* agnostic about how prices, markups and marginal costs change with trade liberalization.

---

[14]To our knowledge, no dataset reports this information.

4 Empirical Framework

This section discusses our empirical framework. We first discuss the identification strategy and then describe the estimation routine.

Consider the log version of the general production function given in equation (1):

$$q_{fjt} = f^j(\boldsymbol{x}_{fjt}; \boldsymbol{\beta}) + \omega_{ft} + \epsilon_{fjt} \tag{8}$$

where lower case letters denote logs. We introduce the subscript j since many firms in our sample manufacture more than one product. The specification states that quantity of product j in firm f at time t, q_{fjt}, is produced using a set of firm-product-year specific inputs, x_{fjt}. The error term ϵ_{fjt} captures measurement error in recorded output as well unanticipated shocks to output: $q_{fjt} = \ln(Q_{fjt}) \exp(\epsilon_{fjt})$. As noted earlier, the productivity term ω_{ft} is assumed to vary at the firm level. Our goal is to estimate of the parameters of the production function – the output elasticities $\boldsymbol{\beta}$ – for each product.

In the subsequent subsections, we discuss the restrictions we impose on equation (8) given the nature of our data and how we identify the output elasticities. The key challenge is that our data (and virtually all other firm-level data sets, for that matter) do not record how inputs are allocated across outputs within a firm. For single-product firms, this concern is of course not an issue. The challenge is how to deal with the unobserved input allocation among multi-product firms while also controlling for unobserved productivity shocks.

4.1 Identification Strategy: The Use of Single-Product Firms

By writing the production function in equation (8) in terms of physical output rather than revenue, we exploit our ability to observe separate information on quantities and prices for each product manufactured by firms. This advantage alleviates concern about a "price bias" that arises when one must deflate sales by an industry-level price index to obtain firm output (for a detailed discussion, see De Loecker (2011)). For single-product firms, the inputs are allocated to the only product these firms manufacture and so we are not concerned about how to allocate inputs across products. The typical concern that arises in productivity estimation is correlation between the unobserved productivity shock and inputs. Removing this bias has been the predominant focus of the production function estimation literature and following the insights of Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg et al. (2006) we use proxy estimators to deal with this correlation.[15]

For multi-product firms, a *new* identification problem arises since the data do not record how the inputs ($\boldsymbol{x}_{fjt}$) are allocated across the products within a firm. To understand this, denote the log of share of input X in the production of product j as $\rho_{fjt}^X = x_{fjt} - x_{ft}$, for any input $X = \{L, M, K\}$, where L is labor, M is materials and K is capital. We only observe firm-level input information X_{ft} and not how each are allocated across products. Substituting this expression into equation (8)

---

[15]Unobserved input prices might still bias estimates of the production function. We discuss this issue in detail in Section 4.3.2.

yields:

$$q_{fjt} = f^j(\boldsymbol{x}_{ft}; \boldsymbol{\beta}) + \omega_{ft} + A_{fjt}(\rho^X_{fjt}, \boldsymbol{x}_{ft}, \beta) + \epsilon_{fjt} \tag{9}$$

where x_{ft} denotes the log of inputs X_{ft}. For multi-product firms, the production function contains an additional component in the error term, $A(.)$, that will generally be a function of the unobserved input shares (ρ^X_{fjt}), the firm level inputs (x_{ft}) and the production function coefficients, $\boldsymbol{\beta}$. The expression (9) clearly demonstrates that $A(.)$ is correlated - by construction - with the inputs on the right-hand-side of the production function which results in biased estimates of $\boldsymbol{\beta}$.[16] We could deal with this bias by taking a stand on the underlying demand function for each product and model how firms compete in that market, but as discussed earlier, we wish to avoid these assumptions in this paper.[17]

We propose an identification strategy that does not require assumptions on the input allocation across products. The strategy is to obtain estimates of the production function using a sample of single-product firms, since as discussed above, the input allocation problem does not exist for these firms. This sample are those firms that produce a single product at a given point in time; it includes firms that may eventually add additional products later in the sample period. This is feature of the sample is important since many firms start off as single product firms and add products during our sample. Our analysis uses these firms in conjunction with firms that always manufacture a sole product.

This identification strategy uses an unbalanced panel of firms for estimation and the imbalance occurs by removing firms from the estimation if they add a product over the sample period. The unbalanced panel is important because it allows for the non-random event that a firm becomes a multi-product producer based on productivity shocks. However, this non-random event results in a sample selection issue that is analogous to the non-random exit of firms discussed in Olley and Pakes (1996). In that context, they are concerned about the left tail of the productivity distribution and the use of an unbalanced panel. Here, a balanced panel of single-product firms would censor the right tail of the productivity distribution. We improve upon this selection problem by using an unbalanced panel of single-product firms and describe the sample selection correction in 4.3.3.[18]

---

[16]To illustrate the bias, consider a translog production function with a single factor of production, labor. The production function in equation (9) would be: $q_{fjt} = \beta_l l_{fjt} + \beta_{ll} l^2_{fjt} + \omega_{ft} + \epsilon_{fjt}$. We observe labor at the firm level, l_{ft}. The labor allocated for the production of j can be written as: $l_{fjt} = \rho_{fjt} + l_{ft}$. Substitution yields:

$$q_{fjt} = \beta_l l_{ft} + \beta_{ll} l^2_{ft} + \underbrace{\beta_l \rho_{fjt} + \beta_{ll}\left(\rho_{fjt}\right)^2 + 2\beta_{ll}(\rho_{fjt} + l_{ft})}_{=A_{fjt}(\rho_{fjt}, \boldsymbol{x}_{ft}, \boldsymbol{\beta})} + \omega_{ft} + \epsilon_{fjt} \tag{10}$$

Clearly, the unobserved component $A_{fjt}(\rho_{fjt}, \boldsymbol{x}_{ft}, \boldsymbol{\beta})$ is unobserved and would be subsumed in the productivity term, ω_{ft}. This would result in biased estimates of β_l since l_{ft} would be correlated with this composite error.

[17]The few productivity studies that have explicitly dealt with multi-product firms had to make assumptions on how inputs are allocated. For example, De Loecker (2011) allocates inputs equally across products and Foster et al. (2008) allocates inputs according to revenue share. However, both studies are not interested in recovering markups and provide conditions under which the production function coefficients are identified.

[18]Our dataset is not well-suited for analyzing firm exit since it is not a census. Indeed, exit is unlikely to be a major concern since we focus on the medium and larger firms in India. Firms in our sample also rarely drop products. We refer the reader to Goldberg et al. (2010b) for a detailed analysis of product adding and dropping in these data.

4.2 Economies of Scope

In this subsection, we explain that despite relying on single-product firms in our estimation, our strategy does not rule out economies of scope.

This identification strategy assumes that the production technology is product-specific rather than specific to the firm (this is why we have a j subscript on the production function in (8)). Once we obtain the coefficients of the production function from the sample of single-product firms, we use those on the multi-product firms that produce the same product. This identification strategy rules out physical synergies across products. For example, imagine a single-product firm produces a t-shirt using a particular technology, and another single-product firms produces carpets using a different combination of inputs. We assume that a multi-product firm that manufactures both products will use each technology on its respective product, rather than some third technology. This assumption is not unusual. In fact, it implicitly assumed in most productivity studies when researchers pool all firms in the same industry and estimate an industry-level production function.

Furthermore, we stress that this assumption does not rule out economies of scope, which may be quite important for multi-product firms. To be precise, Baumol et al. (1983) define economies of scope in production if the cost function is sub-additive: $c_{ft}(q_1, q_2) \leq c_{ft}(q_1) + c_{ft}(q_2)$ where $c_{ft}(.)$ is a firm's cost curve. So while our framework rules out *production* synergies, it allows for economies of scope through *cost* synergies. For example, economies of scope can emerge if multi-product firms buy inputs in larger quantities than single-product firms and therefore pay lower input prices. Furthermore, we account for potential differences in productivity between the two types of firms that might be reflected through differences in management practices or organizational structures. We feel that this restriction is mild, especially given the current practice within the production function estimation literature and weighed against the costs of assuming a market structure/demand system that would dictate how to allocated inputs across products.

4.3 Estimation

This section discusses the production function estimation on the sub-sample of single-product firms.

4.3.1 The Basic Setup

Since we focus on a subsample of single-product firms for the estimation routine, we drop the subscript j to avoid any confusion.[19] Since we have a physical measure of output, we must control for differences in units across products by using unit fixed effects throughout the estimation procedure.

---

[19] Just to be clear, we observe many firms manufacturing the same product j, but we subsume the subscript j when describing firms that only manufacture one product.

The production function we estimate is of the form:[20]

$$q_{ft} = f(\boldsymbol{x}_{ft}; \boldsymbol{\beta}) + \omega_{ft} + \epsilon_{ft} \tag{12}$$

where we subsume the constant term in productivity, collect all inputs in $\boldsymbol{x}_{ft}$, and $\boldsymbol{\beta}$ is the vector of coefficients describing the transformation of inputs into output. In the case of a translog production function, the vector of log inputs $\boldsymbol{x}_{ft}$ are labor, material and capital, their squares, and their interaction terms, and the coefficient vector is $\boldsymbol{\beta} = (\beta_l, \beta_m, \beta_k, \beta_{ll}, \beta_{mm}, \beta_{kk}, \beta_{lm}, \beta_{lk}, \beta_{mk}, \beta_{lmk})$.

We deal with the potential correlation between unobserved productivity shock and input choices by relying on the approach taken by Ackerberg et al. (2006), who modify Olley and Pakes (1996) and Levinsohn and Petrin (2003).[21]

The law of motion for productivity is:

$$\omega_{ft} = g_{t-1}(\omega_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input}) + \xi_{ft} \tag{13}$$

where $b = \{0, 1\}$. This notation implies that we allow trade liberalization to impact a firm's productivity instantaneously or with a lag.[22] Since in our context, tariffs are unanticipated (see Section 2), we do not need to model the law of motion for the tariffs.

We assume that the dynamic inputs–labor and capital–are chosen prior to observing ξ_{ft}. Materials m_{ft} are chosen when the firm learns its productivity. These timing assumptions treat labor, in addition to capital, as an input that faces adjustment costs, which we feel is appropriate in a country like India where labor markets are not flexible (see Besley and Burgess (2004)).

We proxy for productivity by inverting the materials demand function:

$$m_{ft} = m_t(l_{ft}, k_{ft}, \omega_{ft}, \mathbf{z}_{ft}). \tag{14}$$

The vector $\mathbf{z}_{ft}$ contains all additional variables that affect a firm's intermediate input demand (in our case, this is materials). It includes the input (τ_{it}^{input}) and output tariffs (τ_{it}^{output}) that the firm faces on the product. The subscript i on the tariff variables denotes an industry since to reflect that tariffs vary at a higher level of aggregation than products.[23] The vector also includes product fixed effects and the price of the product.[24] These variables capture demand shocks that affect a

---

[20]Our main specification is the translog and is given by

$$q_{ft} = \beta_l l_{ft} + \beta_{ll} l_{ft}^2 + \beta_k k_{ft} + \beta_{kk} k_{ft}^2 + \beta_m m_{ft} + \beta_{mm} m_{ft}^2 + \beta_{lk} l_{ft} k_{ft} + \beta_{lm} l_{ft} m_{ft} + \beta_{mk} m_{ft} k_{ft} + \beta_{lmk} l_{ft} m_{ft} k_{ft} + \omega_{ft} \tag{11}$$

We do not write this in the text to save space.

[21]By using a static control to proxy for productivity, we do not have to revisit the underlying dynamic model and prove invertibility when modifying Olley and Pakes (1996) for our setting to include additional state variables (e.g., tariffs). See De Loecker (2011) and Ackerberg et al. (2006)for an extensive discussion. In developing countries, many firms report zero investment, which further complicates the use of investment as a proxy.

[22]De Loecker (2010) discusses the importance of including all variables which may affect productivity in the law of motion $g_t(\cdot)$.

[23]For example, all tea products (industry 1549) face the same output and input tariffs.

[24]As we discuss below, we estimate the production functions at the two-digit sector level; the product fixed effects are identified since there are many products within each sector.

firm's output, which in turn affects a firm's demand for materials. For example, the product's price affects the quantity produced which in turn affects a firm's input expenditure. As discussed in Olley and Pakes (1996), the proxy approach does not require knowledge of the market structure for the input markets; it simply states that input demand depends on the firm's state variables, capital and labor, productivity and the aforementioned variables.

As is usual in the proxy approach, our estimation proceeds in two steps.[25] We begin by inverting equation (14) to obtain the proxy for firm productivity, $\omega_{ft} = h_t(m_{ft}, l_{ft}, k_{ft}, \mathbf{z}_{ft})$. Following Ackerberg et al. (2006), in the first stage, we run:

$$q_{ft} = \phi_t(l_{ft}, k_{ft}, m_{ft}, \mathbf{z}_{ft}) + \epsilon_{ft} \tag{15}$$

to obtain estimates of expected output $(\widehat{\phi}_{ft})$ and an estimate for the residual ϵ_{ft}.[26]

The first stage separates the effect of ω_{ft} on output from the effects of unanticipated shocks ϵ_{ft} on output. After the first stage, we compute productivity as $\widehat{\omega}_{ft} = \widehat{\phi}_{ft} - f(\boldsymbol{x}_{ft}; \boldsymbol{\beta})$ for any vector of $\boldsymbol{\beta}$. The second stage provides estimates of all production function coefficients by relying on the law of motion for productivity. By non-parametrically regressing $\widehat{\omega}_{ft}(\boldsymbol{\beta})$ on its lag $\widehat{\omega}_{ft-1}(\boldsymbol{\beta})$, and the set of tariff variables $(\boldsymbol{\tau}_{it})$, we recover the innovation to productivity $\xi_{ft}(\boldsymbol{\beta})$ for a given $\boldsymbol{\beta}$.

To estimate the parameter vector $\boldsymbol{\beta}$, we use moments that are now standard in this literature based on degree of variability of a given input. In our context, we assume that firms freely adjust materials and treat capital and labor as dynamic inputs that face adjustment costs. In other settings, one may choose to treat labor as a flexible input. Since material expenditure is the flexible input, we construct its moments using lagged materials.[27] For labor and capital, we construct moments using current and lagged values. The moments are:

$$E\left(\boldsymbol{\xi}_{ft}(\boldsymbol{\beta})\mathbf{Y}_{ft}\right) = 0 \tag{16}$$

where $\mathbf{Y}_{ft}$ contains lag materials, current and one-year lag labor, current and one-year lag capital, and their higher order terms:

$$\boldsymbol{Y}_{ft} = \{l_{ft-b}, l_{ft-b}^2, m_{ft-1}, m_{ft-1}^2, k_{ft-b}, k_{ft-b}^2, l_{ft-b}m_{ft-1}, l_{ft-b}k_{ft-b}, m_{ft-1}k_{ft-b}, l_{ft-b}m_{ft-1}k_{ft-b}\}$$

with $b = \{0, 1\}$. This method identifies the production function coefficients by exploiting the fact that current shocks to productivity will immediately affect a firm's materials choice while labor and capital do not immediately respond to these shocks; moreover, the degree of adjustment can vary across firms and time.

---

[25]In principle, we could estimate both stages jointly using a system GMM estimator, but in practice this is difficult to implement because the first stage contains many parameters including product fixed effects (see De Loecker (2011) for more details). We therefore adopt a two-stage approach.

[26]The set of estimated unit fixed effects are included in our measure of ϵ_{ft}.

[27]In our setting, input tariffs are serially correlated and since they affect input prices, input prices are serially correlated over time.

We estimate the model using a GMM procedure on a sample of firms that manufacture a single product for at least three consecutive years. We choose three years since the moment conditions require at least two years of data because of the lagged values; we add an extra year to allow for potential measurement error in the precise timing of a new product introduction. In principle, one could run the estimation separately for each product, In practice, our sample size is such that we estimate (12) at the two-digit sector level.

4.3.2   Input Price Variation Across Firms

Although we observe quantity and price data for firm outputs, we do not observe this information for firm inputs and so we follow the standard practice in the productivity literature and deflate input expenditures using price indexes. If input prices vary across firms, this practice results in unobserved input price variation across firms generating a concern that is analogous to the one that arises if one deflates output revenue by a common deflator. This concern is present even when data on worker wages are available, as researchers typically do not have information on firm-specific materials prices and never observe the firm-specific user cost of capital. But while this issue always arises when researchers do not observe input-level prices and quantities, the problem is exacerbated in our setting because we use physical quantity data rather than deflated revenues.

To understand this concern, suppose we want to estimate the productivity of two shirt producers. Imagine that the only difference between the firms is that one firm manufactures shirts made from expensive silk while the other manufactures shirts made from cheaper cotton. Otherwise, the two firms have identical productivity: they utilize the same quantity of inputs (balls of yarn, number of workers and number of sewing machines) to produce the same quantity of shirts. In our data, we would not observe the price of the materials (or workers or capital) and so we would deflate the firms' input expenditures by a common deflator. Since we do not account for variation in input prices, we will estimate a lower productivity of the silk shirt manufacturer because it sells the same quantity as its cotton producer counterpart, but has higher (deflated) input expenditures. Our estimation procedure would yield downward biased estimates on the production function coefficients and find productivity difference between the two firms, even though they are exactly identical.

In this setting, the output price of the firm can help control for input price variation across firms. For example, using data from Colombia that uniquely records price information for both inputs and outputs, Kugler and Verhoogen (2011) find that more expensive producers use more expensive inputs. In a world where input prices are passed through to the output price, we can exploit the output price data to control for variation in input prices across firms using the procedure described below, thereby resolving this issue of mapping expenditures into physical quantities.[28]

Formally, we address this problem by first modifying our notation to account for the fact that we observe deflated input expenditures, rather than input quantities. Let $\tilde{\mathbf{x}}_{ft}$ denote the firm's vector of deflated input expenditures and re-write the production function in (12) as:

---

[28]See also the related discussion in Katayama et al. (2009).

$$q_{ft} = f(\tilde{\boldsymbol{x}}_{ft}; \boldsymbol{\beta}) + \omega_{ft} + B_{ft}(\tilde{\mathbf{x}}_{ft}; \mathbf{W}_{ft}; \boldsymbol{\beta}) + \epsilon_{ft}. \tag{17}$$

The term B_{ft} appears because the input deflators are sector-, rather than firm-, specific. In our translog specification, this term captures deviations of the firm input prices from the sector deflators, $\mathbf{W}_{ft}$, and their interactions with $\tilde{\mathbf{x}}_{ft}$ and the parameter vector $\boldsymbol{\beta}$. It is clear from (17) that B_{ft} is correlated with the deflated inputs and will result in biased estimates of the production function.[29]

The presence of B_{ft} is not a concern for the first stage of the estimation since it is not correlated with random unanticipated shocks to production ϵ_{ft}. The purpose of the first stage is simply to estimate expected output $\widehat{\phi}_{ft}$, net of unanticipated shocks ϵ_{ft}.

However, in the second stage of estimation, we need to recover an estimate of the true innovation to productivity ξ_{ft} in order to form moments in equation (16) that identify the production function coefficients. The presence of B_{ft} in (17) violates the orthogonality conditions. To see this problem, define $\tilde{\omega}_{ft}$ as as the composite term:

$$\tilde{\omega}_{ft} = \omega_{ft} + B_{ft}. \tag{18}$$

The expected output from first stage of estimation of production function (17) will yield a second-stage estimate of productivity $\tilde{\omega}_{ft}$ that contains the unobserved variation in input prices. Using the law of motion for productivity in equation (13), let the *measured* innovation to productivity be $\tilde{\xi}_{ft} = \tilde{\omega}_{ft} - g_{t-1}(\tilde{\omega}_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input})$ where $b = \{0,1\}$. Using (18) and the law of motion for productivity (13), we can express the measured innovation to productivity $\tilde{\xi}_{ft}$ as a function of true innovation in productivity ξ_{ft}:

$$\tilde{\xi}_{ft} = \xi_{ft} + B_{ft} - e(B_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input}) \tag{19}$$

where $e(B_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input}) = g_{t-1}(\tilde{\omega}_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input}) - g_{t-1}(\omega_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input})$. Equation (19) illustrates that $\tilde{\xi}_{ft}$ contains current and lagged input prices, which are obviously correlated with the current and lagged input expenditures: the moment conditions are violated. The term $B_{ft} - e(B_{ft-1}, \tau_{it-b}^{output}, \tau_{it-b}^{input})$ creates the correlation between the measured innovation $\tilde{\xi}_{ft}$ and measured input use.

We account for the bias from unobserved variation in input prices in $\tilde{\xi}_{ft}$ by conditioning on additional variables when we form our moment conditions in the second stage. The conditioning set includes flexible polynomials of the product's output price, the input tariff on the prod-

---

[29]To illustrate this bias, consider again the translog production function with a single factor of production, labor, that we specified in Footnote 16. Here, we focus on the bias that arises on single product firms due to unobserved input price variation. We observe the total wage bill of the firm which we deflate with a sector-specific index to obtain the firm's deflated (log) labor expenditure, $\tilde{l}_{ft}$. Let $\tilde{l}_{ft} = l_{ft} + w_{ft}^L$, where w_{ft}^L is the (log) *firm*-specific price of labor (e.g., the firm's wage). Since we don't observe w_{ft}^L, we would estimate:

$$q_{ft} = \beta_l \tilde{l}_{ft} + \beta_{ll} \tilde{l}_{ft}^2 + \underbrace{\beta_l w_{ft}^L + \beta_{ll} \left(w_{ft}^L\right)^2 + 2\beta_{ll}(w_{ft}^L + \tilde{l}_{ft})}_{=B_{ft}(\tilde{\mathbf{x}}_{ft}; \mathbf{W}_{ft}, \boldsymbol{\beta})} + \omega_{ft}.$$

Clearly, the unobserved component $B_{ft}(\tilde{\mathbf{x}}_{ft}; \mathbf{W}_{ft}; \boldsymbol{\beta})$ is unobserved and would be subsumed in the productivity term, ω_{ft}. This would result in biased estimates of β_l since l_{ft} would be correlated with this composite error.

uct, and their interactions with deflated input expenditures.[30] This polynomial, which we denote $d(p_{ft}, \tau_{it-b}^{input}, \tilde{\mathbf{x}}_{ft}; \boldsymbol{\delta})$ with $b = \{0, 1\}$, picks up potential input price variation (and its interaction with input expenditures arising from the translog specification) that is unaccounted for when using sector-specific input price deflators. The use of observed output prices, which we denote by p_{fjt}, to control for unobserved input prices can be rationalized within several models that generate a positive association between input and output prices.[31]

Let us illustrate the approach by focusing on one specific moment condition, the one for lagged materials. The modified moment condition is $E[\tilde{\xi}_{ft}\tilde{m}_{ft-1} | d(p_{ft}, \tau_{it-b}^{input}, \tilde{\mathbf{x}}_{ft})] = 0$, where $b = \{0, 1\}$ and the polynomial $d(.)$ controls for firm-specific input price variation in a flexible way. The coefficients on the variables in $d(.)$ are themselves not of interest. All other moment conditions in equation (16) are modified in a similar way.

Conditioning on $d(.)$ in the moment conditions allows us to obtain consistent estimates of the production function coefficients - our primary objective at this stage of the empirical analysis - without relying on additional instruments that are not available in this context. This issue is similar to the one discussed in Klette and Griliches (1996) in the context of unobserved output prices.

We note that the variables in $d(.)$ cannot be used as alternative instruments when forming moment conditions. The reason is that candidate instruments– p_{ft}, τ_{it}^{ouput} and τ_{it}^{input}–are all correlated with the measured productivity innovation. Current tariffs and output prices are correlated with the unobserved input price variation contained in $\tilde{\xi}_{it}$. This violates the exclusion restriction required for the instruments. While lagged values satisfy the exclusion restriction, there is no reason for them to be correlated with the current period innovation. This implies that the potential lagged instruments may be weak.

While our approach allows us to obtain consistent estimates of the production function coefficients, it does not allow us us to identify separately a firm's physical productivity from its input prices. For example, the coefficient on the current output price p_{ft} in $d(.)$ will capture both its correlation with input prices *and* its correlation with productivity (recall equation (18)). This is precisely why we cannot separately recover firm productivity ω_{ft} and input prices without imposing further assumptions.[32] We can only recover the combined term $\tilde{\omega}_{ft}$. Our productivity measure, unlike productivity measures in most productivity papers, does not capture output price variation (since we use quantities rather than revenues), but we are unable to separate physical productivity from input prices. This is not a limitation in our paper since we can directly analyze marginal costs rather than rely on this potentially contaminated productivity estimate.

---

[30]Differences in input prices across firms reflect not only quality differences (which are captured by output prices), but also other factors. Here, it is natural to include input tariffs since this policy variable affects the prices of imported intermediate inputs.

[31]For example, Kremer (1993) uses an O-ring production function to show that lower and higher quality inputs are not perfectly substitutable and higher quality output requires the use of higher quality inputs. See also the discussion in Verhoogen (2008) and Kugler and Verhoogen (2011)).

[32]For example, one could assume that the polynomial $d(.)$ is not just correlated with the input price term, but correctly measures it. This would allow us to obtain separate measures of productivity and input prices at the firm level.

To summarize, the estimation of the production function coefficients relies on adjusted moment conditions that are formed by conditioning the moment conditions discussed earlier on $d(.)$:

$$E\left(\tilde{\boldsymbol{\xi}}_{ft}(\boldsymbol{\beta})\mathbf{Y}_{ft}|d(.)\right) = 0 \tag{20}$$

If there is no input price variation across firms, the estimates obtained by these adjusted moment conditions would be identical to those obtained when using moment conditions in (16).

4.3.3 The Sample Selection Correction

In this subsection, we discuss the sample selection correction that accounts for the unbalanced panel of single-product firms used in the production function estimation.

Rather than relying on firms that remain single-product producers for the entire sample, our estimation uses an unbalanced panel of single-product firms. As in Olley and Pakes (1996), an unbalanced panel improves on the selection problem since it includes firms that may eventually become multi-product producers in response to productivity shocks.[33] We address the bias arising from sample selection by introducing a correction for sample selection and modifying the law of motion for productivity in an analogous way as Olley and Pakes (1996).

The underlying model behind our sample selection correction is one where the number of products manufactured by firms increases with productivity. There are several multi-product firm models that generate this correlation, and the one that matches our setup most closely is Mayer et al. (2011). In that model, the number of products a firm produces is an increasing step function of the firms' productivity (see Figure 1 in that paper). Firms have a productivity draw which determines their core product. Conditional on entry, the firm produces this core product and incurs an increasingly higher marginal cost of production for each additional product it manufactures. This structure generates a "competence ladder" that is characterized by a set of cutoff points, each associated with the introduction of an additional product.[34]

The cutoff point that is relevant to our sample selection procedure is the one associated with the introduction of a *second* product. Let us denote this cutoff by $\bar{\omega}_{ft}$. Firms with productivity that exceeds $\bar{\omega}_{ft}$ are multi-product firms that produce two (or more) products while firms below $\bar{\omega}_{ft}$ remain single-product producers and are included the estimation sample.

If the threshold $\bar{\omega}_{ft}$ were independent of the variables entering the right hand side of the production function, there would be no selection bias and we would obtain consistent estimates of production function coefficients (as long as we use the unbalanced panel). A bias arises when the threshold is a function of capital and/or labor. For example, it is possible that even conditional on

---

[33]If we modeled the behavior of multi-product firms that shed products to become single-product producers, the number of products would be a natural state variable to include in $\boldsymbol{z}$ in procedure below. So while we abstract away from product deletions since they are rare in our data, our approach can accommodate this case as well.

[34]Alternative models of multi-product firms, such as Bernard et al. (2010b), introduce firm-product-specific demand shocks that generate product switching (e.g., product addition and dropping) in each period. We avoid this additional complexity in our setup since product dropping is not a prominent feature of our data (Goldberg et al. (2010b)). Moreover, in the empirical section we find strong support that firms' marginal costs are lower on their core competent products (products that have higher sales shares).

productivity, a firm with more capital finds it easier to finance the introduction of an additional product; or, a firm that employs more workers may have an easier time expanding into new product lines. In these cases, firms with more capital and/or labor are less likely to be single-product firms, even conditional on productivity and this generates a generate a negative bias in the capital and labor coefficients.

To address the selection bias, we allow the threshold $\bar{\omega}_{ft}$ to be a function of the state variables and the firm's information set at time $t-1$, when we assume the decision is made to add a product. The state variables in our setting include capital, labor, productivity, and all variables in $\mathbf{z}_{ft}$ that are serially correlated (i.e., input and output tariffs, output prices and product dummies). The selection rule requires that the firm make its decision to add a product based on a forecast of these variables in the future. Define an indicator function χ_{ft} to be equal to 1 if the firm remains single-product (SP) and 0 otherwise, and let $\bar{\omega}_{ft}$ be the productivity threshold a firm has to clear in order to produce more than one product.

The selection rule can be rewritten as:

$$
\begin{aligned}
\Pr(\chi_{ft} = 1) &= \Pr\left[\omega_{ft} \leq \bar{\omega}_{ft}(l_{ft}, k_{ft}, \mathbf{z}_{ft}) | \bar{\omega}_{ft}(l_{ft}, k_{ft}, \mathbf{z}_{ft}), \omega_{ft-1}\right] \quad (21) \\
&= \varkappa_{t-1}(\bar{\omega}_{ft}(l_{ft}, k_{ft}, \mathbf{z}_{ft}), \omega_{ft-1}) \\
&= \varkappa_{t-1}(l_{ft-1}, k_{ft-1}, i_{ft-1}, \mathbf{z}_{ft-1}, \omega_{ft-1}) \\
&= \varkappa_{t-1}(l_{ft-1}, k_{ft-1}, i_{ft-1}, \mathbf{z}_{ft-1}, m_{ft-1}) \equiv S_{ft-1}
\end{aligned}
$$

We use the fact that the threshold at t is predicted using the firm's state variables at $t-1$, the accumulation equation for capital[35], and $\omega_{ft-1} = h_t(l_{ft-1}, k_{ft-1}, \mathbf{z}_{ft-1}, m_{ft-1})$ to arrive at the last equation.[36] The law of motion for productivity now becomes:

$$
\omega_{ft} = g'_{t-1}(\omega_{ft-1}, \tau^{input}_{it-b}, \tau^{output}_{it-b}, \bar{\omega}_{ft}) + \xi_{ft} \quad (22)
$$

where $b = \{0, 1\}$.

As in Olley and Pakes, we have two different indexes of firm heterogeneity, the productivity and the productivity cutoff point. Note that $S_{ft-1} = \varkappa_{t-1}(\omega_{ft-1}, \bar{\omega}_{ft})$ and therefore $\bar{\omega}_{ft} = \varkappa^{-1}_{t-1}(\omega_{ft-1}, S_{ft-1})$. Plugging this last expression into the modified law of productivity gives:

$$
\omega_{ft} = g'_{t-1}(\omega_{ft-1}, \tau^{input}_{it-b}, \tau^{output}_{it-b}, S_{ft-1}) + \xi_{ft} \quad (23)
$$

This is the law of motion we use to form the moments in the second stage of the estimation. To obtain S_{ft-1}, we estimate by industry, a probit that regresses a firm's single-product status on materials, capital, labor, the tariff variables, the output price of the product the firm currently produces, product and time dummies. Once the probit is estimated, we construct the propensity

---

[35]The accumulation equation for capital is: $k_{ft} = (1-\delta)k_{ft-1} + i_{ft-1}$, where δ is the depreciation rate of capital.

[36]This specification takes into account that firms hire and/or fire workers based on their labor force at time $t-1$ and their forecast of future demand and costs captured by $\mathbf{z}$ and ω. So all variables entering the non-parametric function $\kappa_{t-1}(.)$ help predict the firm's employment at time t.

score $\hat{S}_{ft-1}$–the predicted probability that the firm remains single-product at time t given the firm's information set at time$t-1$–and insert it in the modified law of motion for productivity.

4.4 Markups and Marginal Costs

4.4.1 Single-Product Firms

We can now apply our framework from Section 3 to compute markups and marginal costs using the estimates of the production function. This computation requires information on a variable input free of adjustment costs, which is materials in our setting. Using equation (5), we compute markups $\widehat{\mu}_{ft} = \widehat{\theta}_{ft}^{M}(\alpha_{ft}^{M})^{-1}$, where the estimated output elasticities $\widehat{\theta}_{ft}^{M}$ on materials are computed using the estimated coefficients of the production function[37] and α_{ft}^{M}, the revenue share of materials, is data. We then use the markup definition in conjunction with the price data to recover the firm's marginal cost, mc_{ft}, at each point in time as $\widehat{mc}_{ft} = \frac{P_{ft}}{\widehat{\mu}_{ft}}$.[38]

4.4.2 Multi-Product Firms

We now discuss how we obtain markups and marginal costs for the multi-product firms.

As shown in equation (6) and (7), computing markups and marginal costs requires the product-specific output elasticity on materials and product-specific revenue shares for materials. We obtain the output elasticity from the single-product firms, but we do not know the product-specific revenue shares of inputs for multi-product firms. Here, we show how to compute the input allocations across products of a multi-product firm in order to construct α_{fjt}^{M}. This approaches applies the structure of the model and the estimated coefficients and does not assume input proportionality.

Let $\rho_{fjt} = \ln\left(\frac{\tilde{X}_{fjt}}{\tilde{X}_{ft}}\right)$ be product j's input cost share, where $\tilde{X}_{ft}$ denotes total deflated expenditures on each input by firm f at time t. We assume that this share does not vary across inputs. We solve for ρ_{fjt} as follows. We first eliminate unanticipated shocks and measurement error from the output data by following the same procedure as in the first stage of our estimation routine in Section 4.3.1 for the single-product firms. We project output quantity, q_{fjt}, on interactions of all inputs, output and input tariffs, the output price, product dummies and time dummies and obtain the predicted values. We next compute a firm-product-specific term $\widehat{\omega}_{fjt}$ that is the residual of expected output and the production function, net of productivity: $\widehat{\omega}_{fjt} \equiv E(q_{fjt}) - f(\tilde{\boldsymbol{x}}_{ft}; \hat{\boldsymbol{\beta}})$.[39] From (9), this becomes:

$$\widehat{\omega}_{fjt} = \omega_{ft} + A_{fjt}(\rho_{fjt}, \boldsymbol{x}_{ft}, \hat{\boldsymbol{\beta}}) \tag{24}$$

$$= \omega_{ft} + \hat{a}_{ft}\rho_{fjt} + \hat{b}_{ft}\rho_{fjt}^2 + \hat{c}_{ft}\rho_{fjt}^3 \tag{25}$$

---

[37]The expression for the materials output elasticity is: $\widehat{\theta}_{ft}^{M} = \hat{\beta}_m + 2\hat{\beta}_{mm}m_{ft} + \hat{\beta}_{lm}l_{ft} + \hat{\beta}_{mk}k_{ft} + \hat{\beta}_{lmk}l_{ft}k_{ft}$.

[38]Output elasticities will vary across sectors because we estimate $\hat{\boldsymbol{\beta}}$ separately for each sector. Moreover, given our translog specification, even single-product firms *within* a sector will have different output elasticities.

[39]The estimated parameter vector$\hat{\boldsymbol{\beta}}$ is already purged of the bias arising from input price variation discussed in Section 4.3.2.

where the second equation follows from applying our translog functional form. The terms $\hat{a}_{ft}$, $\hat{b}_{ft}$, and $\hat{c}_{ft}$ are functions of the estimated parameter vector$\hat{\boldsymbol{\beta}}$.[40]

With an estimate of $E(q_{fjt})$, we can construct $\widehat{\omega}_{fjt}$ for each multi-product firm observation (firm-year-product triplet). For each year, we obtain the firm's productivity and input allocations, the $J+1$ unknowns $(\omega_{ft}, \rho_{f1t}, \ldots, \rho_{fJt})$, by solving a system of $J+1$ equations:

$$\widehat{\omega}_{1ft} = \omega_{ft} + \hat{a}_{ft}\rho_{f1t} + \hat{b}_{ft}\rho_{f1t}^2 + \hat{c}_{ft}\rho_{f1t}^3 \tag{26}$$

$$\ldots \quad \cdots \tag{27}$$

$$\widehat{\omega}_{fJt} = \omega_{ft} + \hat{a}_{ft}\rho_{fJt} + \hat{b}_{ft}\rho_{fJt}^2 + \hat{c}_{ft}\rho_{fJt}^3 \tag{28}$$

$$\sum_{j=1}^{J} \exp\left(\rho_{fjt}\right) = 1, \qquad \exp(\rho_{fjt}) \le 1 \forall j \tag{29}$$

This system imposes the economic restriction that each input share can never exceed one and they must together sum up to one across products in a firm. We solve this system for each firm in each year using Matlab.[41]

We now have all the ingredients to calculate markups and the implied marginal costs for the multi-product firms according to equation (6):

$$\hat{\mu}_{fjt} = \hat{\theta}_{fjt}^M \frac{P_{fjt}Q_{fjt}}{\exp\left(\hat{\rho}_{fjt}\right)P_{ft}^M V_{ft}^M} \tag{30}$$

The product-specific output elasticity for materials, $\hat{\theta}_{fjt}^M$, is a function of the production function coefficients, the product-specific output (which is data) and the materials allocated to product j. Hence, it can be easily computed once the allocation of inputs across products has been recovered.[42] Marginal costs for the products made by multi-product firms are then recovered by dividing prices by the markup according to equation (7).

---

[40] For the translog production function:

$$\hat{a}_{ft} = \hat{\beta}_l + \hat{\beta}_m + \hat{\beta}_k + 2\left(\hat{\beta}_{ll}l_{ft} + \hat{\beta}_{mm}m_{ft} + \hat{\beta}_{kk}k_{ft}\right) + \hat{\beta}_{lm}\left(l_{ft} + m_{ft}\right)$$
$$\quad + \hat{\beta}_{lk}\left(l_{ft} + k_{ft}\right) + \hat{\beta}_{mk}\left(m_{ft} + k_{ft}\right) + \hat{\beta}_{lmk}(l_{ft}m_{ft} + l_{ft}k_{ft} + m_{ft}k_{ft})$$
$$\hat{b}_{ft} = \hat{\beta}_{ll} + \hat{\beta}_{mm} + \hat{\beta}_{kk} + \hat{\beta}_{lm} + \hat{\beta}_{lk} + \hat{\beta}_{mk} + \hat{\beta}_{lmk}\left(l_{ft} + m_{ft} + k_{ft}\right)$$
$$\hat{c}_{ft} = \hat{\beta}_{lmk}$$

[41] We find that the input allocations across products are highly correlated with, but not identical to, allocating inputs according to product revenue shares. We experiment with various starting values for the unknowns and find that conditional on converging to an inside solution (e.g., all the product's input shares are between 0 and 1, non-inclusive), the solution is unique. Out of the total 12,958 multi-product firm-year pairs, we hit a corner solution in 745 cases and so we exclude observations from these firm-year pairs from the analysis in Section 5.

[42] The expression for the materials output elasticity for product j at time t is: $\widehat{\theta}_{fjt}^M = \hat{\beta}_m + 2\hat{\beta}_{mm}\left[\hat{\rho}_{fjt} + m_{ft}\right] + \hat{\beta}_{lm}\left[\hat{\rho}_{fjt} + l_{ft}\right] + \hat{\beta}_{mk}\left[\hat{\rho}_{fjt} + k_{ft}\right] + \hat{\beta}_{lmk}\left[\hat{\rho}_{fjt} + l_{ft}\right]\left[\hat{\rho}_{fjt} + k_{ft}\right]$. For single product firms, $exp(\rho_{fjt}) = 1$ and the output elasticity becomes the one reported in Footnote 37.

5 Results

5.1 Output Elasticities

In this subsection, we present the output elasticities recovered from the estimation procedure outlined in 4. We also describe how failing to correct for input price variation or account for the selection bias affects the parameters.

The output elasticities for the single-product firms used in the estimation are reported in the first panel of Table 3.[43] A nice feature of the translog is that unlike in a Cobb-Douglas production function, output elasticities can vary across firms (and across products within firms). We therefore report both the average and standard deviation of the elasticities across sectors. The last column in the table reports the returns to scale. The main message of the table is that the average returns to scale exceed one in most sectors and that there is a distribution around the average.

The first panel of Table 4 re-runs the estimation without implementing the correction for the unobserved input price variation discussed in Section 4.3.2. It is clear that the uncorrected procedure yields nonsensical estimates of the production function. For example, the labor output elasticities are often negative and range from -1.44 to 2.4 across sectors, while the returns to scale are either very low or extremely high. As we discussed earlier, these results are not surprising given that we estimate a quantity-based production function using deflated input expenditures. It is clear that failing to account for input price variation across firms yields distorted estimates. For example, a negative labor elasticity stems from the fact that there are firms that produce similar quantities of output using more expensive labor, resulting in a downward biased labor elasticity. In general, it is difficult to sign this bias because there are three inputs which interact in complicated ways in the translog, but it is clear that one needs to correct for input price variation across firms using the procedure described above.

In the second panel of Table 4, we present the output elasticities from estimation of the production function on a balanced panel of single-product firms. This estimation does not account for the sample selection correction described in Section 4.3.3. As discussed earlier, we expect the output elasticity for capital, and to a lesser extent the output elasticity for labor, to change if we use a balanced panel. We find that moving from the balanced panel to the specification we use in Table 3(unbalanced panel and selection correction) increases the returns to scale estimate in 9 out of 14 industries. In several instances, the output elasticity of capital increases. This is consistent with our expectations and the underlying selection equation: all things equal (including productivity), firms with more capital have a higher propensity to become multi-product producers over time. This leads to a negative correlation between capital and the firm specific productivity threshold resulting in a downward bias capital elasticity in the balanced panel estimation. The output elasticities of labor and especially materials do not change as drastically which is consistent with our premise that materials immediately adjust to productivity and demand shocks and are therefore not a state

---

[43]The output elasticity for materials is reported in Footnote 37. The output elasticities for labor and capital are
$\widehat{\theta}_{ft}^{L} = \hat{\beta}_l + 2\hat{\beta}_{ll}l_{ft} + \hat{\beta}_{lm}m_{ft} + \hat{\beta}_{lk}k_{ft} + \hat{\beta}_{lmk}m_{ft}k_{ft}$ and $\widehat{\theta}_{ft}^{K} = \hat{\beta}_k + 2\hat{\beta}_{kk}k_{ft} + \hat{\beta}_{lk}l_{ft} + \hat{\beta}_{mk}m_{ft} + \hat{\beta}_{lmk}l_{ft}k_{ft}$.

variable for firms.

5.2 Prices, Markups and Marginal Costs Patterns

Our data and methodology deliver measures of firm performance–prices, markups, marginal costs, and productivity, which are usually not simultaneously considered in the existing literature. This section describes the relationship between prices, markups, marginal costs, and productivity.

The markup estimates are reported in Table 5. There is considerable variation across sectors. The mean and median markups are 1.67 and 1.04, respectively. As discussed in 3, our methodology yields measures of markups and marginal costs without a priori assumptions on the returns to scale. However, our production function estimates in Section 5.1 suggest increasing returns to scale in most industries. Once we back out marginal costs, we would therefore expect to observe an inverse relationship between a product's marginal cost and quantity produced.[44] In Figure 2 we plot marginal costs against production quantities (we de-mean each variable by product-year fixed effects in order to facilitate comparisons across firms). The figure shows indeed that marginal costs vary inversely with production quantities. The right panel of the figure shows that quantities and markups are positively related indicating that firms producing more output also enjoy higher markups (due to their lower marginal costs).

We next examine the relationships between the performance measures. Table 6 reports the raw pairwise correlation matrix across prices, markups, marginal costs and firm-level productivity. As we expect, prices are positively correlated with marginal costs and markups, but negatively correlated with firm productivity. Marginal costs are negatively correlated with markups and productivity, and there is a positive correlation between markups and productivity.[45]

Figure 1 presents an alternative way to see the inter-relatedness of the firm performance measures by plotting firms' productivity against the marginal costs and markups of their products. In order to compare these variables across firms, we regress the firm-level productivity measures on the firm's main industry-year pair fixed effects and recover the residuals. This allows us to compare productivity across firms within industries. We compare the markups and marginal costs across products by de-meaning each by product-year fixed effects. The figure plots the de-meaned markups and marginal costs against the de-meaned productivities, and removes outliers above and below the 97th and 3rd percentiles. Again, the variables are correlated in ways that one would expect: productivity is positively correlated with markups and strongly negatively correlated with marginal costs.

The previous tables and figures demonstrate correlation patterns that arise across firms. We also

---

[44]It is important to note here that we do not derive marginal cost based on the duality of production and cost functions. Instead, we bring in additional information on prices and derive marginal cost based on its definition as the ratio of price to markup. This implies that even though we find returns to scale in the production function estimation, it does not follow automatically that our estimated marginal costs will be declining in quantity.

[45]Our productivity measure does not encompass variation in output prices. However, as we discussed earlier, like in most existing measures of productivity, our measure of productivity encompasses both variation in producer-specific input prices and firm-level physical productivity. The above correlations between productivity and other performance measures might in part reflect producer-specific input prices.

examine how markups and marginal costs vary across products *within* a firm. Our analysis here is guided by the recent literature on multi-product firms. Our correlations are remarkably consistent with the predictions of this literature, especially with those of the multi-product firm extension of Melitz and Ottaviano (2008) developed by Mayer et al. (2011). A key assumption in that model is that multi-product firms each have a "core competency". The "core" product has the lowest (within a firm) marginal cost. For the other products, marginal costs rise with a product's distance from the "core competency". A firm's product scope is determined by the point at which the marginal revenue of a product is equal to the marginal cost of manufacturing that product. This structure on the production technology is also shared by the multi-product firm models of Eckel and Neary (2010) and Arkolakis and Muendler (2010). In all three models, more productive firms manufacture more products, a pattern clearly confirmed by Figure 3.

Mayer et al. (2011) also assume a linear demand system which implies that firms have non-constant markups across products. Further, firms have their highest markups on their "core" products with markups declining as they move away from their main product. Figures 4 and 5 provide evidence supporting both implications. In these two figures, markups and marginal costs are de-meaned by product-year and firm-year fixed effects in order to make these variables comparable across products within firms. Figure 5 plots the de-meaned markups and marginal costs against the sales share of the product within each firm, and 4 plots the same variables against the product's rank. We observe that the marginal costs rise as a firm moves away from its core competency while the markups fall. In other words, the firm's most profitable product (excluding any product-specific fixed costs) is its core product. Despite not imposing any assumptions on the market structure and demand system in our estimation, these correlations are remarkably consistent with Melitz and Ottaviano (2008).

These descriptive results highlight the advantage of jointly analyzing firm performance measures in empirical work. The variables are correlated in predictable ways that are consistent with recent models of heterogeneous multi-product firms. Foreshadowing our results in the next subsection, we also find evidence of imperfect pass-through of costs on prices because of variable markups. When we relate prices to marginal costs, we find a pass-through coefficient of 0.20.[46] This implies that any shock to marginal costs, for example through trade liberalization, will not translate to a proportional change in factory-gate prices because of changes in markups. We examine this markup adjustment in detail in the subsequent section.

---

[46]We obtain this pass-through coefficient by regressing (log) prices on (log) marginal costs while controlling for firm-product fixed effects, so the pass-through coefficient is identified based on variation over time. We obtain a higher pass-through coefficient–0.35–by regressing prices on marginal costs controlling for product-year fixed effects. In either case, pass-through is far from complete. Note that the marginal cost coefficient should be interpreted with caution, since marginal costs are estimates derived from prices; therefore, any measurement error in prices is transmitted to marginal costs. This positively correlated measurement error, however, should only bias us against finding incomplete pass-through.

5.3   Prices, Markups and Trade Liberalization

We now examine how prices, markups and marginal costs adjusted as India liberalized its economy. As discussed in Section 2, we restrict the analysis to 1989-1997 since tariff movements after this period appear correlated with industry characteristics.

We begin by simply plotting the distribution of prices in 1989 and 1997 in Figure 6. Here, we include only firm-product pairs that are present in both years, and we compare the prices over time by regressing them on firm-product pair fixed effects and plotting the residuals. As before, we remove outliers in the bottom and top 3rd percentiles. This comparison of the same firm-product pairs over time exploits the same variation as our regression analysis below. The figure shows that the distribution of (real) prices is virtually *unchanged* between 1989 and 1997. This is a surprising result given nature of India's economic reforms during this period that were designed to reduce entry barriers and increase competition in the manufacturing sector. At a first pass, the figure suggests that prices did not move much despite these reforms.

Of course, the figure includes only firm-product pairs that are present at the beginning and end of the sample, and does not control for macroeconomic factors that could influence prices beyond the trade reforms. We next relate log prices to output tariffs:[47]

$$p_{fjt} = \delta_{fj} + \delta_t + \delta_1 \tau_{it}^{output} + \eta_{fjt}. \tag{31}$$

By including firm-product fixed effects (δ_{fj}), we exploit variation in prices and output tariffs within a firm-product over time. We control for macroeconomic fluctuations through year fixed effects δ_t. Since the trade policy measure varies at the industry level, we cluster our standard errors at this level. The price regression is reported in column 1 of Table 7. The coefficient on the output tariff is positive implying that a 10 percentage point decline is associated with a very small–1.11 percent–decline in prices.[48] Between 1989 and 1997, output tariffs fell on average by 62 percentage points; this resulted in a precisely estimated average price decline of 6.9 percent (=62*.111). This is a small effect of the trade reform on prices and it is consistent with the raw distributions plotted in Figure 6. The basic message remains the same if we control for secular trends within each sector using sector-year fixed effects (column 2). The results imply that the average decline in output tariffs led to a 7.9 (=62*.127) percent relative drop in prices.

These results show that although the trade liberalization led to lower factory-gate prices, the decline is more modest than we would have expected given the magnitude of the tariff declines. Our previous work (Goldberg et al. (2010a), Topalova and Khandelwal (2011)) has emphasized the

---

[47]One could try to capture the net impact of tariff reforms using the effective rate of protection measure (ERP) proposed by Corden (1966). However, this measure is derived in a setting with perfect competition and an infinite export-demand and import-supply elasticities which imply perfect pass-through. As we show below, these assumptions are not satisfied in our setting, so that the concept of the "effective rate of protection" is not well defined in our case. For the interested reader, we note that if we regress prices on the ERP, we find that a 10 percentage point decline in the effective rate of protection reduces prices by .44 percent; however, for the reasons given above, we do not attempt to interpret this change.

[48]Our result is consistent with Topalova (2010) who finds that a 10 percentage point decline in output tariffs results in a 0.96 percent decline in wholesale prices in India during this period.

importance of declines in input tariffs in shaping firm performance. It is useful to separate the effects of output tariffs and input tariffs on prices. Output tariff liberalization reflects primarily an increase in competition, while the input tariff liberalization should provide access to lower cost (and more variety of) inputs. We run the analog of the regression in (31), but separately include input and output tariffs:

$$p_{fjt} = \delta_{fj} + \delta_{st} + \delta_1 \tau_{it}^{output} + \delta_2 \tau_{it}^{input} + \eta_{fjt}. \tag{32}$$

We use sector-year fixed effects (δ_{st}) in this specification and report the results in column 1 of Table 8.[49] There are two interesting findings that are important for understanding how trade affected prices in this liberalization episode. First, there is a positive and statistically significant coefficient on output tariffs. This result is consistent with the common intuition that increases in competitive pressures through lower output tariffs will lead to price declines. The effect is traditionally attributed to reductions in markups and/or reductions in X-inefficiencies within the firm. The point estimates imply that a 10 percentage point decline in output tariffs led to a 1.19 percent decline in prices. On the other hand, the coefficient on input tariffs is very noisy. In a counterfactual analysis that holds input tariffs fixed and reduces output tariffs, we would observe a fairly precisely estimated decline in prices. However, given the noisy effect of the input tariffs, the overall changes of prices taking into account both types of tariff declines are not well estimated. Over the sample period, output tariffs and input tariffs fall by 62 and 24 percentage points, respectively. Using the point estimates in column 1, this implies that prices fall on average by 13.1 percent, but this decline is no longer statistically significant.

Our methodology that recovers markups and marginal costs allows us to explore the mechanisms behind these moderate and often noisy changes in factory-gate prices. We begin by plotting the distribution of markups and costs in Figure 7. Like Figure 6, this figure considers only firm-product pairs that appear in both 1989 and 1997 and de-means the observations by their time average. The figure demonstrates that between 1989 and 1997, the marginal cost distribution shifted left indicating an efficiency gain. However, this marginal cost decline is almost exactly offset by a corresponding rightward shift in the markup distribution. Since (log) marginal costs and (log) markups exactly sum to (log) prices, the net effect is virtually no change on prices.

In columns 2-3 of Table 8, we use the specification in (32) to relate marginal costs and markups to input and output tariffs. Since prices decompose exactly to the sum of marginal costs and markups, the coefficients in columns 2 and 3 sum to their respective coefficients in column 1. The estimates are somewhat noisy, but suggest a clear pattern. The positive coefficient on output tariffs in the marginal costs regression (column 2) suggests that X-inefficiencies were moderately reduced when output tariffs fell. More importantly, although the coefficient on input tariffs is not statistically significant, it has a large positive point estimate indicating that improved access to cheaper and more variety of imported inputs resulted in cost declines. The final row of Table 8 reports the average effect on marginal costs using the average declines in input and output tariffs. On average,

---

[49]The results with year fixed effects are qualitatively similar and available upon request.

marginal costs fell 24.6 percent (significant at the 10.4 percent level).

This magnitude of the marginal costs declines is fairly sizable and would translate to larger than observed prices declines if markups remained constant. However, Figure 7 suggests that markups rose during this period, and in column 3 of Table 8, we directly examine how input and output tariffs affected markups. The results show a large negative coefficient on input tariffs (significant at the 13 percent level) which implies that input tariff liberalization resulted in higher markups. Together with column 2, the table suggests that firms offset the beneficial cost reductions from improved access to imported inputs by raising markups. The overall effect, taking into account the average declines in input and output tariffs between 1989 and 1997, is that markups, on average, increased by 11.4 percent. This increase offsets about half of the average decline in marginal costs, and as a result, the overall effect of the trade reform on prices is moderated.

The markup regression in column 3 of Table 8 warrants additional discussion. We might expect the coefficient on output tariffs to be larger as competitive pressures through lower output tariffs change the residual demand for domestic products and reduce markups. However, from column 2, we observe that output tariffs simultaneously reduced marginal costs (through reductions of X-inefficiencies) and earlier results establish that firms offset cost declines by raising markups. This would imply a negative correlation between output tariffs and markups which would attenuate the coefficient in column 3. So while both effects due to output tariffs result in lower prices (column 1), the two channels offset each other and as a result, we observe a small effect of output tariffs on markups. To demonstrate this effect, we re-examine the relationship between markups and output tariffs while controlling for marginal costs and report the findings in column 1 of Table 9. By conditioning on marginal costs, the output tariff coefficient now captures only the direct pro-competitive effect of output tariffs on markups. Indeed, we now observe a positive and significant coefficient on output tariffs which provides direct evidence of the pro-competitive effects that output tariffs have on markups. In column 2, we also include input tariffs. As discussed earlier, input tariffs should affect markups only through imperfect pass-through of their effects on marginal costs through cheaper imported inputs. Once we control for marginal costs, input tariffs should have no effect on markups, and this is what we find. The coefficient on input tariffs falls dramatically compared to column 3 of Table 8. We acknowledge that it is difficult to interpret the coefficient on marginal costs in Table 9, but these results shed light on why we find a small effect of output tariffs on markups in column 3 of Table 8.

While the overall message of Table 8 is consistent with Figures 6 and 7, some of the results are somewhat noisy. One concern is that the marginal cost and markup variables are computed from estimates of the production function. Column 1 of Table 3 indicates that some sectors used many more observations in the production function estimation than others (e.g., for example, there are 264 single-product firms in apparel products, and 1,521 such firms in the food products sector). We re-examine our results by weighing the regression in (32) by the number of observations used to estimate the production functions for each sector reported in Table 3. That is, we assign more weight to the observations in the food product sector than in the apparel sector since we are more

confident in our estimates of the output elasticities, and consequently the estimates of markups and marginal costs, for sectors where we have more data. The results are reported in the right panel of Table 8. The weighted regressions yield more precise estimates of the input tariff coefficients. In particular, the precision on the markup regression improves. More importantly, however, the pattern and message of the results remain as before. On average, the combined impact of the trade liberalization was to reduce marginal costs, with the source of these improvements coming predominantly from imported inputs. However, firms simultaneously raised markups and only about half of the efficiency gains were passed through.

We also investigate the price response by looking at three-year differences since annual fluctuations in prices may be small. The long difference specification based on (32) is given by:

$$\Delta p_{fjt} = \delta_{st} + \delta_1 \Delta \tau_{it}^{output} + \delta_2 \Delta \tau_{it}^{input} + \Delta \eta_{fjt}, \tag{33}$$

where the Δ operator is a three-year difference between t and $t-3$. We present the long difference results in Table 10; the left panel has the unweighted regressions and the right panel weights by the observations from the production function estimations. The price regressions in columns 1 and 4 indicate that output tariff liberalization led to a statistically significant decline in prices. The input tariff coefficient, however, is noisy. As a result, the average price decline due to the trade liberalization is not precisely estimated. As with the annual regressions, we find that input tariffs led to opposite patterns in the responses of costs and markups. The average declines in marginal costs due to lower input tariffs are essentially offset by increases in markups. The net effect is an incomplete pass-through of the trade reforms on factory-gate prices.

Our results suggest that input tariff liberalization lowered the marginal costs of productions for the Indian firms. However, firms did not pass through their entire declines in costs on factory-gate prices. This suggests that producers benefited relatively more from the trade reforms, at least in the short run, than consumers. However, given results from our earlier work in Goldberg et al. (2010a), we also know that firms introduced many new products–accounting for about a quarter of output growth–during this period. In Table 11, we report results that relate a firm's product additions to changes in its average markup across its products. In column 1, we regress an indicator if a firm adds a product in period t on the change in (log) average markups between $t-1$ and t, while controlling for firm and year fixed effects. In column 2, we use the change in the (log) number of products as the dependent variable. In both cases, increases in markups are correlated with new product introductions.[50] This suggests that that firms use the input tariff reductions to finance the development of new products and so the long-term gains to consumers could be substantially larger. A complete analysis of this mechanism, however, lies beyond the scope of this current paper.

---

[50]These findings are consistent with Peters (2012) who develops a model with imperfect competition that generates heterogeneous markups which determine innovation incentives. We note however that we did not obtain these results by committing to a particular structure (demand, market structure and competition) - our findings could also be consistent with alternative mechanisms.

6 Conclusion

This paper examines the adjustment of prices, markups and marginal costs in response to trade liberalization. We take advantage of detailed price and quantity information to estimate markups from quantity-based production functions. Our approach does not require any assumptions on the market structure or demand curves that firms face. This feature of our approach is important in our context since we want to analyze how markups adjust to trade reforms without imposing *ex ante* restrictions on their behavior. An added advantage of our approach is that since we observe firm-level prices in the data, we can directly compute firms' marginal costs once we have estimates of the markups.

Estimating quantity-based production functions for a broad range of differentiated products introduces new methodological issues that we must confront. We propose an identification strategy to estimate production functions on single-product firms. The advantage of this approach is that we do not need to take a stand on how inputs are allocated across products within multi-product firms. We also demonstrate how to correct for a bias that arises when researchers do not observe input price variation across firms, an issue that becomes particularly important when estimating quantity-based production functions.

The large variation in markups suggests that trade models that assume constant markups may be missing an important channel when quantifying the gains from trade. Furthermore, our results highlight the importance of analyzing the effects of both output and input tariff liberalization. We observe large declines in marginal costs, particularly due to input tariff liberalization. However, prices do not fall by as much. This imperfect pass-through occurs because firms offset the cost declines by raising markups; on average, we observe that only about half of the cost declines attributed to the trade reform are passed through in the form of lower prices. These results suggest that trade liberalization can have large, yet nuanced effects, on productivity and markups. Understanding these distributional consequences of input and output tariff reforms through variable markups is an important channel for future research.

Our results have broader implications for thinking about the trade and productivity across firms in developing countries. The methodology produces quantity-based productivity measures that can be compared with revenue-based productivity measures. Hsieh and Klenow (2009) discuss how differences between these two measures can inform us about distortions and the magnitude of misallocation within an economy. Importantly, our quantity-based productivity measure is purged of substantial variation in markups across firms which potentially improves upon our understanding of the impact of misallocation on productivity dispersion. We leave the analysis of the role of misallocation on the distribution of these performance measures for future research.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2006). Structural identification of production functions. *Mimeo, UCLA.*

Amiti, M. and J. Konings (2007). Trade liberalization, intermediate inputs, and productivity: Evidence from indonesia. *American Economic Review 97*(5), 1611–1638.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Arkolakis, C. and M.-A. Muendler (2010). The extensive margin of exporting products: A firm-level analysis. *NBER Working Paper 16641*.

Baumol, W. J., J. C. Panzar, and R. D. Willig (1983). Contestable markets: An uprising in the theory of industry structure: Reply. *American Economic Review 73*(3), 491–96.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003). Plants and productivity in international trade. *American Economic Review 93*(4), 1268–1290.

Bernard, A. B., S. J. Redding, and P. K. Schott (2010a). Multiple-product firms and product switching. *American Economic Review 100*(1), 70–97.

Bernard, A. B., S. J. Redding, and P. K. Schott (2010b). Multiple-product firms and product switching. *American Economic Review 100*(1), 70–97.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–90.

Besley, T. and R. Burgess (2004). Can labor regulation hinder economic performance? evidence from india. *The Quarterly Journal of Economics 119*(1), 91–134.

Bloom, N. and J. Van Reenen (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics 122*(4), 1351–1408.

Corden, W. M. (1966). The structure of a tariff system and the effective protective rate. *Journal of Political Economy 74*(3), 221–237.

De Loecker, J. (2010). A note on detecting learning by exporting. *NBER Working Paper 16548*.

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica 79*(5), 1407–1451.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *American Economic Review*. forthcoming.

Debroy, B. and A. Santhanam (1993). Matching trade codes with industrial codes. *Foreign Trade Bulletin 24*(1), 5–27.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

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

Edmonds, C., V. Midrigan, and Y. Xu (2011). Competition, markups and the gains from international trade. *Mimeo, Duke University*.

Feenstra, R. C. and D. Weinstein (2010). Globalization, markups, and the U.S. price level. *NBER Working Paper 15749*.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review 98*(1), 394–425.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the u.s. automobile industry. *Econometrica 63*(4), 891–951.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2009). Trade liberalization and new imported inputs. *American Economic Review 99*(2), 494–500.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010a). Imported intermediate inputs and domestic product growth: Evidence from india. *The Quarterly Journal of Economics 125*(4), 1727–1767.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010b). Multiproduct firms and product turnover in the developing world: Evidence from india. *The Review of Economics and Statistics 92*(4), 1042–1049.

Goyal, S. (1996). Political economy of India's economic reforms. *Institute for Studies in Industrial Development Working Paper*.

Hall, R. E. (1986). Market structure and macroeconomic fluctuations. *Brookings Papers on Economic Activity 17*(2), 285–338.

Halpern, L., M. Koren, and A. Szeidl (2011). Imports and productivity. *CEPR Discussion Paper 5139*.

Harrison, A. E. (1994). Productivity, imperfect competition and trade reform : Theory and evidence. *Journal of International Economics 36*(1-2), 53–73.

Hasan, R., D. Mitra, and K. Ramaswamy (2007). Trade reforms, labor regulations, and labor-demand elasticities: Empirical evidence from india. *The Review of Economics and Statistics 89*(3), 466–481.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Katayama, H., S. Lu, and J. R. Tybout (2009). Firm-level productivity studies: Illusions and a solution. *International Journal of Industrial Organization 27*(3), 403–413.

Klette, T. J. and Z. Griliches (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics 11*(4), 343–61.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics 108*(3), 551–75.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review 70*(5), 950–59.

Kugler, M. and E. Verhoogen (2011). Prices, plant size, and product quality. *The Review of Economic Studies*.

Levinsohn, J. (1993). Testing the imports-as-market-discipline hypothesis. *Journal of International Economics 35*(1-2), 1–22.

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

Mayer, T., M. J. Melitz, and G. I. Ottaviano (2011). Market size, competition, and the product mix of exporters. *NBER Working Paper 16959*.

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

Melitz, M. J. and G. I. P. Ottaviano (2008). Market size, trade, and productivity. *Review of Economic Studies 75*(1), 295–316.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Panagariya, A. (2008). *India: The Emerging Giant.* Number 9780195315035 in OUP Catalogue. Oxford University Press.

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants. *Review of Economic Studies 69*(1), 245–76.

Peters, M. (2012). Heterogeneous mark-ups and endogenous misallocation. *Mimeo, Massachussets Institute of Technology*.

Sivadasan, J. (2009). Barriers to competition and productivity: Evidence from india. *The B.E. Journal of Economic Analysis & Policy 9*(1), 42.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from india. *American Economic Journal: Applied Economics 2*(4), 1–41.

Topalova, P. and A. K. Khandelwal (2011). Trade liberalization and firm productivity: The case of india. *The Review of Economics and Statistics 93*(3), 995–1009.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector. *The Quarterly Journal of Economics 123*(2), 489–530.

Tables and Figures

Table 1: Summary Statistics

Sector	Share of Sample Output (1)	All Firms (2)	Single-Product Firms (3)	Products (4)
15 Food products and beverages	8%	298	137	135
17 Textiles	8%	328	196	78
18 Wearing apparel	0%	23	20	14
21 Paper and paper products	2%	76	59	32
23 Coke, refined petroleum products	25%	26	11	22
24 Chemicals	19%	434	216	474
25 Rubber and Plastic	4%	146	102	83
26 Non-metallic mineral products	5%	116	88	60
27 Basic Metal	8%	214	143	99
28 Fabricated metal products	1%	75	55	45
29 Machinery and equipment	5%	165	84	183
31 Electrical machinery and apparatus	4%	87	55	96
32 Radio, TV and communication	2%	50	39	87
34 Motor vehicles, trailers	8%	59	40	94
Total	100%	2,097	1,245	1,502

Notes: Table reports summary statistics for the sample. The first column reports the share of output by sector in 1995. Columns 2 and 3 report the number of firms and number of single-product firms manufacturing products in the sector in 1995. Column 4 reports the number of products over the full sample, 1989-2003.

Table 2: Example of Sector, Industry and Product Classifications

Examples of Industries, Sectors and Products	
NIC Code	Description
27	Basic Metal Industries (Sector *s*)
2710	Manufacture of Basic Iron & Steel (Industry *i*)
130101010000	Pig iron
130101020000	Sponge iron
130101030000	Ferro alloys
130106040800	Welded steel tubular poles
130106040900	Steel tubular structural poles
130106050000	Tube & pipe fittings
130106100000	Wires & ropes of iron & steel
130106100300	Stranded wire
2731	Casting of iron and steel (Industry *i*)
130106030000	Castings & forgings
130106030100	Castings
130106030101	Steel castings
130106030102	Cast iron castings
130106030103	Maleable iron castings
130106030104	S.G. iron castings
130106030199	Castings, nec

(Products (*j*) brace spans the product rows for NIC 2710 and NIC 2731 respectively.)

Notes: This table is replicated from Goldberg et al. (2010b). For NIC 2710, there are a total of 111 products, but only a subset are listed in the table. For NIC 2731, all products are listed in the table.

Table 3: Output Elasticities of Translog Production Function, by Sector

Sector	Obs. in Single-Product Firm Estimation (1)	Labor (2)	Materials (3)	Capital (4)	Returns to Scale (5)
15 Food products and beverages	1,521	0.31 [0.06]	0.61 [0.05]	0.13 [0.05]	1.06 [0.07]
17 Textiles	1,900	0.45 [0.07]	0.59 [0.10]	0.10 [0.01]	1.14 [0.15]
18 Wearing apparel	264	0.33 [0.04]	0.60 [0.00]	0.12 [0.02]	1.05 [0.04]
21 Paper and paper products	702	0.40 [0.09]	0.63 [0.08]	0.13 [0.06]	1.16 [0.22]
23 Coke, refined petroleum products	159	0.30 [0.02]	0.60 [0.00]	0.10 [0.01]	0.99 [0.03]
24 Chemicals	2,174	0.36 [0.05]	0.61 [0.09]	0.15 [0.08]	1.11 [0.20]
25 Rubber and Plastic	1,005	0.17 [0.03]	0.59 [0.03]	0.11 [0.02]	0.88 [0.07]
26 Non-metallic mineral products	909	0.26 [0.05]	0.57 [0.05]	0.12 [0.03]	0.95 [0.11]
27 Basic Metal	1,286	0.49 [0.11]	0.57 [0.11]	0.11 [0.03]	1.17 [0.21]
28 Fabricated metal products	595	0.35 [0.04]	0.59 [0.04]	0.10 [0.00]	1.04 [0.07]
29 Machinery and equipment	902	0.38 [0.04]	0.59 [0.02]	0.09 [0.00]	1.06 [0.06]
31 Electrical machinery and apparatus	574	0.32 [0.03]	1.04 [0.33]	0.10 [0.02]	1.46 [0.35]
32 Radio, TV and communication	401	0.29 [0.03]	0.60 [0.04]	0.10 [0.02]	0.99 [0.09]
34 Motor vehicles, trailers	554	0.30 [0.03]	0.58 [0.03]	0.32 [0.14]	1.20 [0.11]

Notes: Table reports the output elasticities from the baseline production function. The first column reports the number of observations for each production function estimation. As noted in the text, these observations are the number of firms that remain single-product for three consecutive years. Columns 2-4 report the average estimated output elasticity with respect to each factor of production for the translog production function for these single-product firms. Standard deviations of the output elasticities reported in brackets are below the mean values. The 5th column reports the average returns to scale, which is the sum of the mean values from the preceding three columns.

Table 4: Output Elasticities, Input Price Variation and Sample Selection

Sector	Estimates without Correcting for Input Price Variation				Estimates on a Balanced Panel of Single-Product Firms			
	Labor (1)	Materials (2)	Capital (3)	Returns to Scale (4)	Labor (5)	Materials (6)	Capital (7)	Returns to Scale (8)
15 Food products and beverages	1.22	0.50	0.36	2.09	0.25	0.65	0.14	1.04
17 Textiles	3.17	-0.24	-2.50	0.43	0.42	0.63	0.09	1.13
18 Wearing apparel	0.74	0.14	0.33	1.21	0.37	0.54	0.06	0.97
21 Paper and paper products	0.00	1.30	1.19	2.49	0.26	0.55	0.14	0.96
23 Coke, refined petroleum products	0.53	0.10	-0.04	0.59	0.33	0.61	0.14	1.08
24 Chemicals	0.00	1.36	0.42	1.78	0.30	0.58	0.08	0.96
25 Rubber and Plastic	-1.12	1.19	1.88	1.96	0.32	0.64	0.07	1.03
26 Non-metallic mineral products	-2.53	1.96	0.28	-0.29	0.36	0.63	0.21	1.19
27 Basic Metal	-1.04	0.73	1.15	0.85	0.46	0.55	0.07	1.08
28 Fabricated metal products	2.01	0.44	-0.02	2.43	0.29	0.59	0.09	0.97
29 Machinery and equipment	0.37	-0.02	0.39	0.74	0.37	0.60	0.11	1.08
31 Electrical machinery and apparatus	-0.71	-0.03	-0.70	-1.44	0.30	0.73	0.03	1.06
32 Radio, TV and communication	-3.01	0.54	2.20	-0.27	0.29	0.65	0.12	1.06
34 Motor vehicles, trailers	0.43	0.47	0.25	1.16	0.29	0.63	0.21	1.13

Notes: The first column of the table reports the number of observations for each production function estimation. To save space, we do not report the stardard deviations of the output elasticities. As noted in the text, these observations are the number of firms that remain single-product for three consecutive years. The first panel reports the output elasticities and returns to scale for the estimation procedure that does not account for input price variation. The second panel reports the output elasticities and returns to scale for the estimation procedure that uses a balanced set of single-product firms from 1989-2003. For this panel, the observations associated with each sector are less than the values reported in column 1 of Table 3.

Table 5: Markups, by Sector

Sector	Markups	
	Mean	Median
15 Food products and beverages	1.64	1.09
17 Textiles	1.48	1.27
18 Wearing apparel	1.10	1.10
21 Paper and paper products	0.96	0.95
23 Coke, refined petroleum products	2.07	0.84
24 Chemicals	1.09	0.62
25 Rubber and Plastic	1.64	1.07
26 Non-metallic mineral products	3.41	1.79
27 Basic Metal	1.49	1.23
28 Fabricated metal products	2.16	1.21
29 Machinery and equipment	2.40	1.05
31 Electrical machinery	2.01	1.35
32 Radio, TV and communication	3.11	1.16
34 Motor vehicles	2.00	1.10
Average	1.68	1.04

Notes: Table displays the mean and median markup by sector for the sample 1989-2003. The table trims observations with markups that are above and below the 3rd and 97th percentiles within each sector.

Table 6: Pairwise Correlation Matrix between Prices, Markups, Marginal Costs and Productivity

	Prices	Markups	Marginal Costs	Productivity
Prices	1.00			
Markups	0.14	1.00		
Marginal Costs	0.93	-0.23	1.00	
Productivity	-0.35	0.02	-0.35	1.00

Notes: Table reports the pairwise correlation matrix across four performance variables: prices, markups, marginal costs and productivity. All variables are expressed in logs. Prices, markups and marginal costs vary at the firm-product level while productivity varies at the firm level. The table trims observations with markups that are above and below the 3rd and 97th percentiles within each sector.

Table 7: Prices and Output Tariffs, Annual Regressions

	Log Prices$_{fjt}$	Log Prices$_{fjt}$
	(1)	(2)
Output Tariff$_{it}$	0.111**	0.127***
	0.055	0.046
R-squared	0.00	0.02
Observations	19,768	19,768
Firm-Product FEs	yes	yes
Year FEs	yes	no
Sector-Year FEs	no	yes
Overall Impact of Trade	-6.9	-7.9***
Liberalization	3.4	2.8

Notes: The dependent variable is a firm-product's (log) price. The left column includes year fixed effects and the right column includes sector-year fixed effects; all specifications have firm-product fixed effects. The regressions exclude outliers in the top and bottom 3rd percent of the markup distribution. All regressions include firm-product and are run from 1989-1997, and standard errors are clustered at the industry level. The final row uses the average 62% decline in output tariffs from 1989-1997 to compute the mean and standard error of the impact of trade liberalization on prices. In other words, for each column the mean impact is equal to the -0.62*100*{coefficient on output tariffs}. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 8: Prices and Tariffs, Annual Regressions

	Log Prices$_{fjt}$	Log Marginal Cost$_{fjt}$	Log Markup$_{fjt}$	Log Prices$_{fjt}$	Log Marginal Cost$_{fjt}$	Log Markup$_{fjt}$
	(1)	(2)	(3)	(4)	(5)	(6)
Output Tariff$_{it}$	0.119**	0.098	0.022	0.136**	0.099	0.037
	0.050	0.089	0.072	0.057	0.101	0.074
Input Tariff$_{it}$	0.234	0.763	-0.529	0.344	1.082	-0.738*
	0.437	0.627	0.338	0.489	0.755	0.382
R-squared	0.02	0.04	0.03	0.02	0.04	0.03
Observations	19,768	19,768	19,768	19,768	19,768	19,768
Firm-Product FEs	yes	yes	yes	yes	yes	yes
Sector-Year FEs	yes	yes	yes	yes	yes	yes
Weights (see footnote)	no	no	no	yes	yes	yes
Overall Impact of Trade Liberalization	-13.1	-24.6	11.4	-16.8*	-32.4*	15.6
	9.2	15.5	10.0	9.9	17.9	11.0

Notes: The dependent variable is noted in the columns. Within each panel, the sum of the coefficients from the markup and marginal costs regression equals their respective coefficient in the unit value regression. The regressions exclude outliers in the top and bottom 3rd percent of the markup distribution. All regressions include firm-product fixed effects and sector-year fixed effects. The right panel uses the number of observations from the production function estimation (see column 1 of Table 3) as weights for each sector. The regressions are run from 1989-1997 and standard errors are clustered at the industry level. The final row uses the average 62% and 24% declines in output and input tariffs from 1989-1997, respectively, to compute the mean and standard error of the impact of trade liberalization on each performance measure. In other words, for each column the mean impact is equal to the -0.64*100*{coefficient on output tariff} + -0.24*100*{coefficient on input tariff}. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 9: Markups and Pass-through

	Log Markup$_{fjt}$	
	(1)	(2)
Log Marginal Cost$_{fjt}$	-0.726***	-0.726***
	0.063	0.063
Output Tariff$_{it}$	0.093**	0.093**
	0.038	0.042
Input Tariff$_{it}$		0.025
		0.309
R-squared	0.62	0.62
Observations	19,768	19,768
Firm-Product FEs	yes	yes
Sector-Year FEs	yes	yes

Notes: The dependent variable is (log) markup. The regressions exclude outliers in the top and bottom 3rd percent of the markup distribution. All regressions include firm-product fixed effects and sector-year fixed effects. The regressions are run from 1989-1997 and standard errors are clustered at the industry level. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 10: Prices and Tariffs, Long Difference Regressions

|	Log Prices$_{fjt}$	Log Marginal Cost$_{fjt}$	Log Markup$_{fjt}$	Log Prices$_{fjt}$	Log Marginal Cost$_{fjt}$	Log Markup$_{fjt}$
	(1)	(2)	(3)	(4)	(5)	(6)
ΔOutput Tariff$_{it}$	0.078*	0.067	0.011	0.096**	0.074	0.022
	0.044	0.088	0.072	0.048	0.095	0.068
ΔInput Tariff$_{it}$	0.107	0.587	-0.480	0.136	0.886	-0.749**
	0.402	0.569	0.332	0.430	0.690	0.364
R-squared	0.04	0.02	0.02	0.04	0.02	0.01
Observations	8,263	8,263	8,263	8,263	8,263	8,263
Sector-Year FEs	yes	yes	yes	yes	yes	yes
Weights (see footnote)	no	no	no	yes	yes	yes
Average Impact of Trade	-7.4	-18.4	11.0	-9.3	-26.1	16.8*
Liberalization	8.7	14.6	9.3	9.2	17.0	10.1

Notes: The dependent variable is the three-year change of each variable noted in the columns. Within each panel, the sum of the coefficients from the markup and marginal costs regression equals their respective coefficient in the unit value regression. The regressions exclude outliers in the top and bottom 3rd percent of the markup distribution. All regressions include sector-year fixed effects. The right panel uses the number of observations from the production function estimation (see column 1 of Table 3) as weights for each sector. The regressions are run from 1989-1997 and standard errors are clustered at the industry level. The final row uses the average 62% and 24% declines in output and input tariffs from 1989-1997, respectively, to compute the mean and standard error of the impact of trade liberalization on each performance measure. In other words, for each column the mean impact is equal to the -0.64*100*{coefficient on output tariff} + -0.24*100*{coefficient on input tariff}. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table 11: Markups and Product Scope

|	Add Dummy (1)	Change in Log Products (2)
Change in Log Average Markup$_{ft}$	0.031***	0.013***
	(0.005)	(0.003)
Firm FEs	yes	yes
Year FEs	yes	yes
R-squared	0.36	0.23
Obserations	9,007	9,007

Notes: The independent variable in each regression is the change between *t-1* and *t* of a firm's average (unweighted) log markup. The average is contructed by first dropping outlier markups below and above the 3rd and 97th percentile, respectively, and then computing the (unweighted) average markup across products for each firm and then taking first differences. The dependent variable in column one is an indicator if the firm adds a product between *t-1* and *t*. The dependent variable in column 2 is the change in the firm's log number of products. All regressions include firm and year fixed effects. Significance: * 10 percent, ** 5 percent, *** 1 percent.
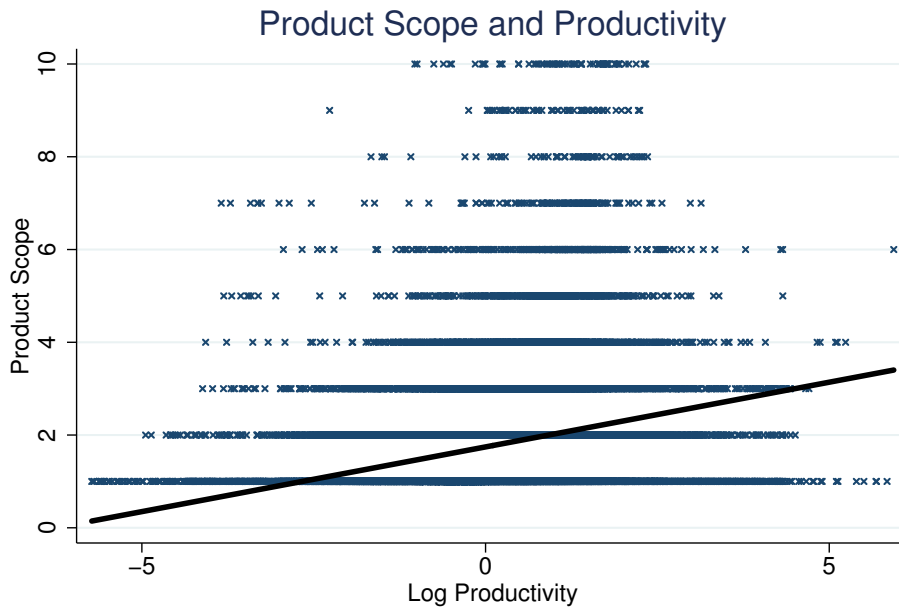
Figure 1: Marginal Costs and Productivity
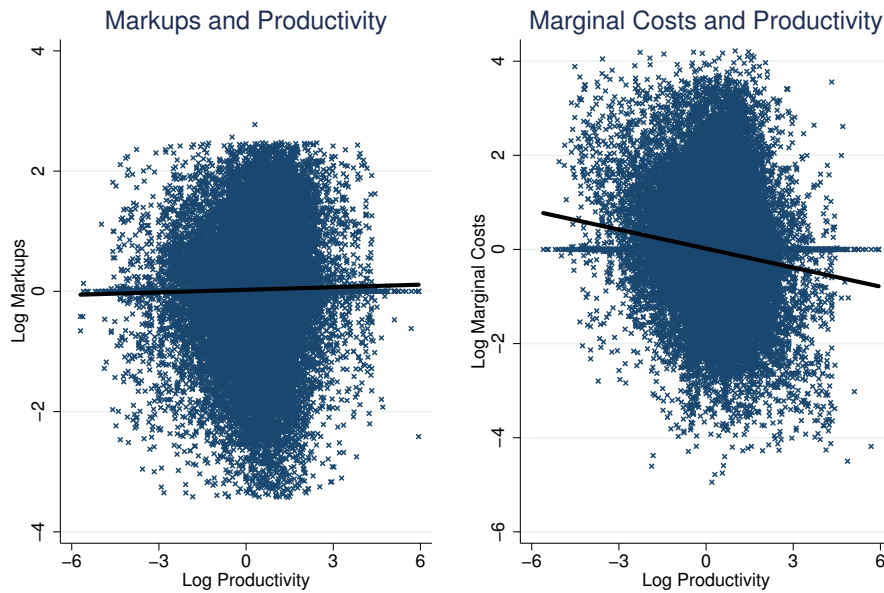
Markup and Marginal costs are de−meaned by product−year FEs. Firm productivity is demeaned by the firm's main industry−year FE. For each variable, outliers are trimmed below and above 3rd and 97th percentiles.

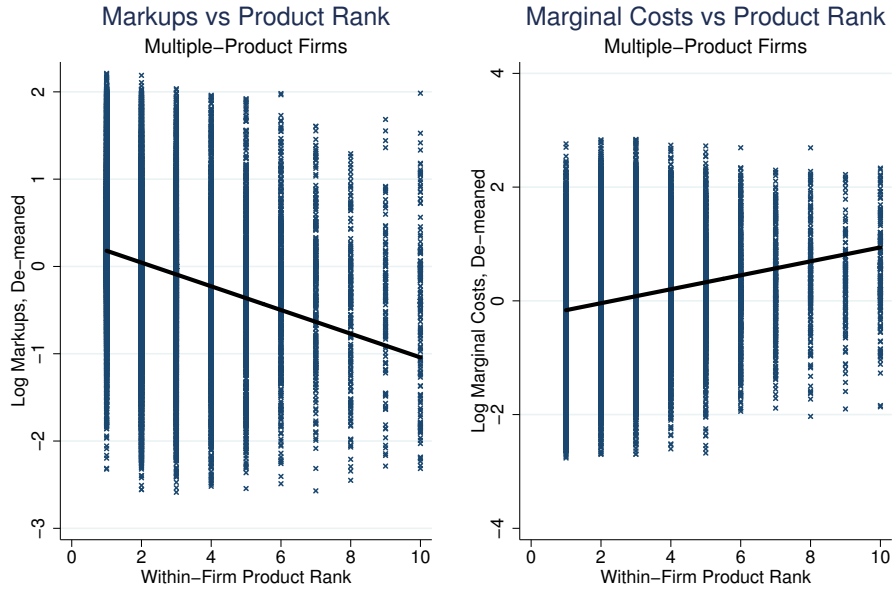Figure 2: Marginal Costs and Quantities
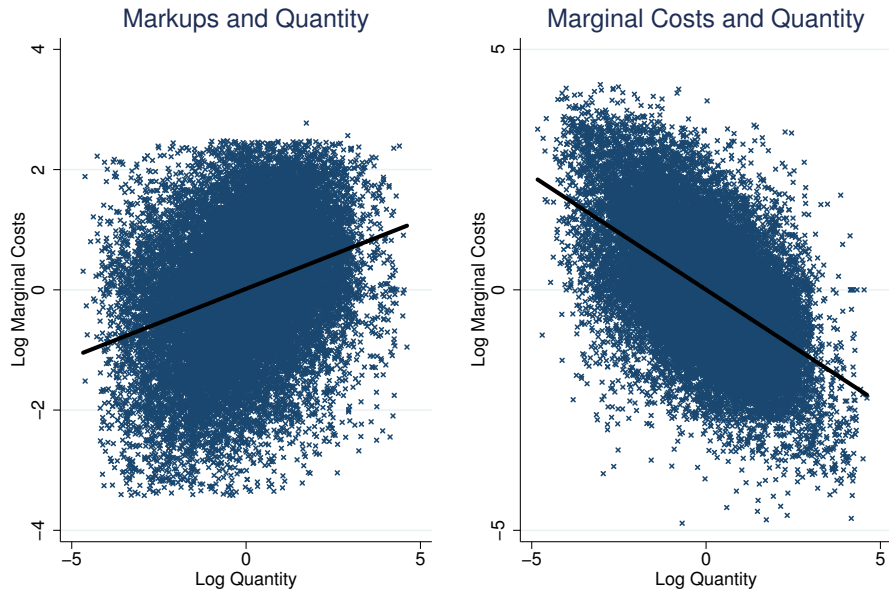
Variables de−meaned by product−year FEs.
Markups, cost and quantity outliers are trimmed below and above 3rd and 97th percentiles.

Figure 3: Product Scope and Productivity

Firm productivity is demeaned by the firm's main industry−year FE.
Productivity outliers are trimmed below and above 3rd and 97th percentiles.

Figure 4: Markups, Costs and Product Rank

Markups and marginal costs are de-meaned by product-year and firm-year FEs.
Markup and marginal cost outliers are trimmed below and above 3rd and 97th percentiles.

Figure 5: Markups, Costs and Product Sales Share
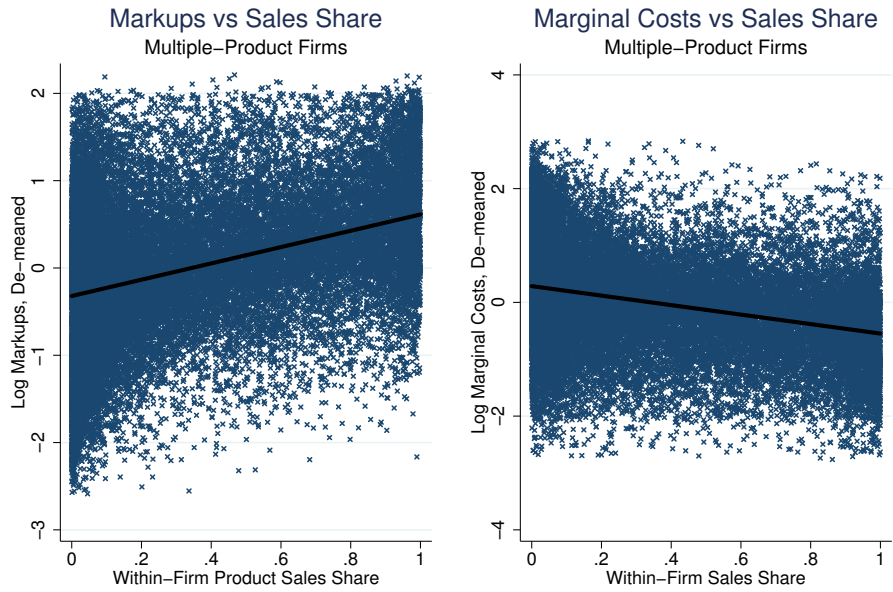
Markups and marginal costs are de-meaned by product-year and firm-year FEs.
Markup and marginal cost outliers are trimmed below and above 3rd and 97th percentiles.

Figure 6: Distribution of Prices in 1989 and 1997

Distribution of Prices
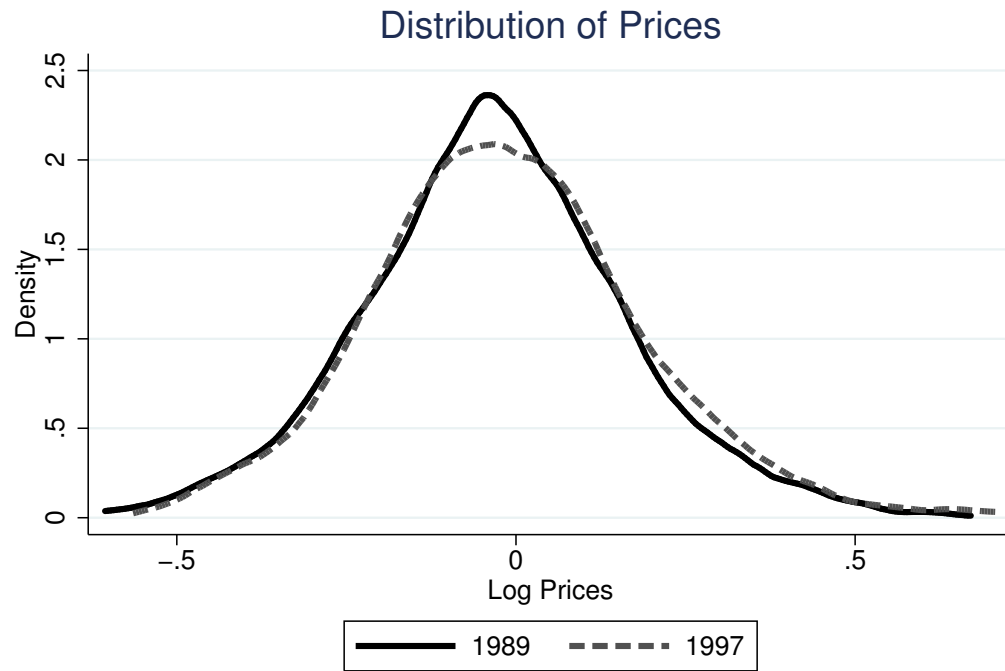
Sample only includes firm–product pairs present in 1989 and 1997.
Observations are de–meaned by their time average, and outliers above and below the 3rd and 97th percentiles are trimmed.

Figure 7: Distribution of Markups and Marginal Costs in 1989 and 1997