Productivity, prices and market shares in multiproduct firms*

Gianluca Antonecchia
Erasmus School of Economics and Tinbergen Institute †

Ajay Bhaskarabhatla
Erasmus School of Economics and ERIM ‡

15 September 2018

Abstract

We examine the determinants of product market shares in the Indian pharmaceutical industry featuring multiproduct firms. Using detailed data on firm-level financial information coupled with product-level sales and wholesale prices, we find that in narrowly defined homogeneous product markets, products sold at higher wholesale prices have a higher market share. We estimate a reduced form model of product market share that controls for wholesaler prices as well as the role of consumers and intermediaries. To disentangle the demand- and supply-side effects of price on market shares we estimate product-level productivity, controlling for the biases related to input measurement, simultaneity, and product scope. We find that product prices are correlated negatively with quantity-based productivity and positively with revenue-based productivity. We use quantity-based productivity measures to instrument for product wholesale price and find that an increase in wholesale price stemming from the supply side affects market share negatively. Retail margins and product appeal drive market share upward, suggesting that the observed positive correlation between price and market share is mainly due to demand-related features, involving retailer buyer power and uninformed consumers.

*We are grateful to Serguey Braguinsky, Amil Petrin and the participants to our presentation at Erasmus School of Economics (OSE seminar) and Miami Business School (NAPW). We thank AIOCD for sharing the data.

†Contact: antonecchia@ese.eur.nl

‡Contact: bhaskarabhatla@ese.eur.nl

1 Introduction

Amoxicillin Clavulanic Acid, an antibiotic launched in October 1993, is India's best-selling medicine with more than 200 brands competing in the market. The leading brand in the market, *Augmentin*, with a market share of 22.6 percent in January 2016 has a wholesale price of Rs. 206 for ten 625 mg tablets. By contrast, the brand with the second-highest market share of 12 percent has a relatively lower wholesale price of Rs. 63 for ten tablets. The analysis of more than 7800 narrowly defined medicine markets in India from 2011 to 2014 shows that, on average, products with a higher wholesale price have a higher market share (Table 1). This pattern holds in a subsample of markets featuring ten or more firms and when we normalize the wholesale price with the highest wholesale price in the market in the month. The positive relationship between prices and market shares is not only puzzling but, if confirmed, it also points to a public health crisis in one of the world's most populous countries characterized by limited public health infrastructure, low levels of health insurance, and consumers directly purchasing and paying for medicines exclusively from specialist, government-licensed retailers.

This paper examines the relationship between prices and market shares in the Indian pharmaceutical industry featuring multiproduct firms. The positive relationship that we observe can be due to both supply- and demand-side factors. On the demand side, quality difference among products is a common explanation of why higher price can be correlated with higher market share provided that consumers prefer high quality products. However, our analysis compares drugs in very narrowly defined markets, where products have in common both molecules and dosage form (tablet, injections, etc.) and only differ by pack size and dosage strenght. Therefore the markets in our analysis can be considered as composed by homogeneous products. In addition, Bennett and Yin (2014) conducted a quality test on the most important antibiotics in India and showed that 96 percent of the drugs sampled comply with Indian Pharmacopoeia quality standards.[1] However, although *effective* quality difference is not responsible for the observed price dispersion, we are unable to control for *perceived* quality. Prior research shows that uninformed consumers pay a higher premium for purchasing branded medicines while similar generic medicines are available (Bronnenberg et al., 2015). Naturally, if there are more uninformed consumers in the market, brands with higher prices have higher market shares as well.

[1] In our final analysis we consider only traded companies which are supposed to be more keen (and controlled) on quality aspects.

On the supply side, differences in prices can stem from both differences in productivity (cost structure) and markups. The supply of medicines in India involves a vertical distribution chain that includes pharmaceutical manufacturers in the upstream and wholesalers and retailers of medicines in the downstream. The retailers are represented by a trade association with considerable buyer power, and they employ restrictive vertical trade practices such as regulating entry into local markets and boycotting the sales of medicine brands that offer low retailer markups (Bhaskarabhatla et al., 2016b). The presence of such retailer buyer power can create incentives for the retailer trade association to promote the sale of brands manufactured by less-productive pharmaceutical manufacturers that offer higher markups while neglecting the more productive manufacturers with lower retailer markups, resulting in allocative inefficiency in production. On the other hand, more productive manufacturers may gain market share by offering higher retailer markups, indicating that the inefficiency on the supply side lies with the buyer power of the downstream intermediaries rather than the manufacturers' market power. Our paper attempts to disentangle these effects.

Using detailed data on firm-level financial information coupled with product-level sales and wholesale prices, we investigate the drivers of market shares among multiproduct firms at the product level. Following Hottman et al. (2016) we identify the drivers of product market share: cost, markup, appeal and firm scope. We adopt a reduced form approach and proxy: i) cost with product *wholesale price*; ii) markup with *retailers' margin* on the product; iii) appeal with *brand and firm age*; iv) firm scope with the *number of markets* where it operates. To identify the demand- and supply-side effects of prices on market shares we adopt an instrumental variable approach, using product-level productivity to instrument wholesale prices, as suggested by Foster et al. (2008). Productivity proxies for technical efficiency in the production of a single good, proving to be a relevant driver for prices and an exogenous variable for demand.

We estimate product-level productivity, controlling for the main biases highlighted by the prior literature: i) product-level input allocation; ii) input price differentials; iii) simultaneity bias between inputs and TFP; iv) product scope bias. Building on the recent advanced in the literature, we propose a strategy that addresses all these biases (De Loecker et al., 2016; Dhyne et al., 2017). By adopting a production function suitable for analysing the pharmaceutical industry, we estimate a quantity-based measure of product-level productivity and compare it to other measures proposed by the literature.[2]

[2]Productivity can be distinguished between revenue-based or quantity-based, depending on whether

Our findings can be summarized as follows. Market shares and wholesale prices are positively correlated with each other in narrowly defined homogeneous markets. Similar to Foster et al. (2008), we find that wholesale prices are correlated negatively with quantity-based productivity and positively with revenue-based productivity. Instrumenting wholesale price with quantity-based productivity shows that an increase in the price stemming from the supply side (manufacturer cost structure) affects market share negatively. The other determinants associated with retailers and consumers, instead, affect market share positively. This suggests that the observed positive correlation between price and market share is mainly due to demand-related features involving retailer buyer power (*perceived* quality) and uninformed consumers (*perceived* quality or product/firm appeal).

This research contributes to the extant literature in several ways. First, we contribute to the empirical literature on the determinants of product demand in imperfectly competitive markets where the power of brand, retailers or trade associations can heavily determine market shares of products. We disentangle the effect of demand and supply-side on market share and find that the observed positive correlation between price and market share is driven by demand-related factors. This implies that product-level technical efficiency is not sufficient to determine market shares in the pharmaceutical industry, generating distortions that affect product (and firm) profitability and survival.

Second, we contribute to the empirical literature on estimating productivity for multi-product firms. Using traditional methods (Levinsohn and Petrin, 2003), as well as recent advances (De Loecker et al., 2016; Dhyne et al., 2017), we address the specific biases arising in multiproduct firms' productivity estimation - i.e. input allocation across products and product scope bias - proposing an original estimation strategy that better suits with the industry we examine. We estimate productivity at the product-level, where each product is defined at the stock keeping unit (SKU). By comparison, Foster et al. (2008) select eleven seven-digit products manufactured in the US. Bernard et al. (2010) collect 1500 five-digit SIC codes products of the US manufacturing. De Loecker et al. (2016) use data for the entire Indian manufacturing and observe around 2400 products. Our sample of pharmaceutical products allows us to calculate product-level productivity for more than 41,000 stock keeping units.

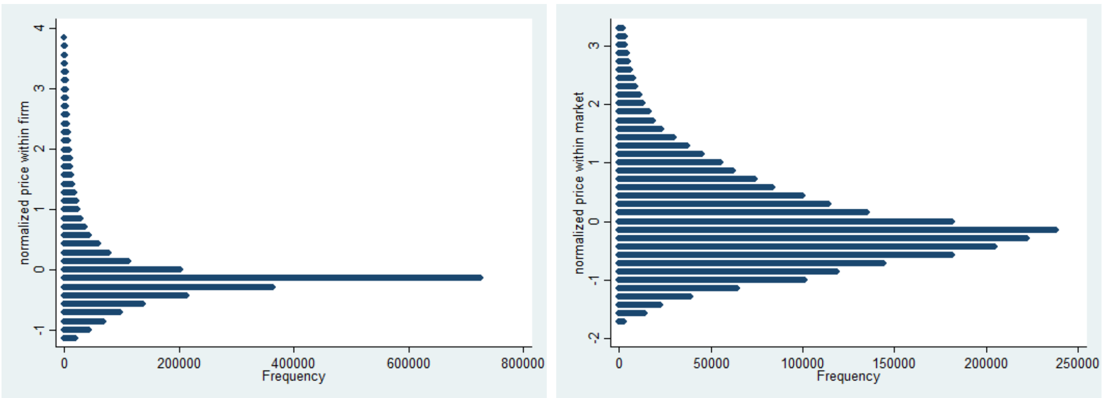
Third, the instrumental variable approach we use to identify the supply-side effect on

revenue or physical units sold is used as the output variable in the production function.

market share includes an analysis of product price heterogeneity. Figure 1 shows product price distribution within the firm (panel a) and within the market (panel b). Studying prices using an average firm-level price measure rather than individual product prices set across hundreds of product markets is a compromising solution which eclipses product-level heterogeneity in prices (Smeets and Warzynski, 2013). We contribute to this area of research, studying the relationship between product-level productivity and wholesale prices. We find the negative correlation between quantity-based productivity and product wholesale price turns into positive when revenue-based productivity is considered. This result confirms the findings in Foster et al. (2008), pointing at revenue-based productivity, used by the large majority of the literature, as a distorted measure of technical efficiency.

Figure 1: Distribution of product prices

(a) within-firm prices (b) within-market prices



Source: AIOCD dataset

Note: Panel (a): percentiles 1 to 99 of the normalized distribution of product prices to the retailer within the firm (the normalized mean is the average firm price) . Panel (b): percentiles 1 to 99 of the normalized distribution of product prices to the retailer within the market (the normalized mean is the average market price).

Fourth, our paper contributes to an emerging literature that examines the rise in markups in the US in the recent decades and points to an increased market concentration (De Loecker and Eeckhout, 2017). Although our context features markets that are seemingly competitive, we examine the role of price margins, that is the difference between retail and wholesale prices, in maintaining the dominant position of market leaders who are also the price leaders. Our findings show that margins positively affect the market share of the product, suggesting a determinant (and potentially distortive) role of the intermediaries.

Fifth, we provide evidence about the functioning of an imperfectly competitive industry in an emerging country. Indian firms have gained attention recently because India has undergone several important reforms over the last three decades.[3] Using a sample of multiproduct firms from India, Goldberg et al. (2010b) show that multiproduct firms are larger, more productive, and perform better. Recent studies examine the impact of trade reforms in India on price and markups (De Loecker et al., 2016), product scope (Goldberg et al., 2010a), productivity (Topalova and Khandelwal, 2011; Ahsan, 2013). Other scholars focus on the deregulation process occurred in India and found non-significant effects on productivity growth (Bollard et al., 2013) but significant effects on the reallocation of market shares towards the bigger firms (Alfaro and Chari, 2014). The imposition of price control regulation on several essential medicines has recently received attention (Bhaskarabhatla et al., 2016b,a).

The paper proceeds as follows: in Section 2 we present the dataset and in Section 3 we identify the determinants of market share and discuss the empirical strategy. In Section 4 we introduce the methodology to estimate product-level productivity in multiproduct firms. Section 5 shows and comments the main results and Section 6 concludes.

2 Data

We use new product-level data of the universe of multiproduct firms operating in the Indian pharmaceutical industry. Data are provided by the All India Organization of Chemists and Druggists (AIOCD). A unique feature of our data is that we observe both wholesale and retail prices for each stock keeping unit (SKU) of the medicine, as well as the sales of each SKU of each firm collected for 48 consecutive months starting in January 2011. The data contain nearly 2.5 million SKU-month observations spanning 901 firms and almost 89 thousand SKUs. The sample allows defining product *markets* based on the molecular composition (four-digit therapeutic category) of the drug. We can identify more than 2900 markets. Within a market we further distinguish *submarkets* based on the dosage form of the drug (like injection, tablet, etc.). We can identify more than 7800 submarkets. Within a firm we also group the products into *brands* based on

[3]Firms in developing countries are, in general, the target of academic studies because they face political constraints or innovations which are particularly useful for policy evaluation analysis (Pavcnik, 2002; Van Biesebroeck, 2005; Lileeva and Trefler, 2010). Moreover, countries at different stages of development exhibit nonnegligible differences in firm size, firm performance and resource allocation (Hsieh and Klenow, 2009; Bartelsman et al., 2013; De and Nagaraj, 2014).

the commercial name of the drug. One brand can gather many products having different dosage form, strength and pack size. In the Indian pharmaceutical industry there are more than 51000 brands. We present an overview of AIOCD dataset in Table 2. Almost 99% of the firms in the Indian pharmaceutical industry are multiproduct firms, manufacturing, on average, 101 different SKUs. Indian pharmaceutical firms are mainly multiscope firms, operating, on average, in 47 markets and 61 submarkets, showing a large propensity to change their product mix - introducing or dropping a (new) product - or market mix - entering or quitting a (new) market.

We merge AIOCD product-level data with Prowess data on firm financials compiled by the Centre for Monitoring Indian Economy (CMIE).[4] The CMIE Prowess data include yearly information on the standalone financial statement and other identifying information on the firms (name, industry, location, age) of all Indian companies traded on the National Stock Exchange and the Bombay Stock Exchange. The merged dataset contains over one million product-month observations, spanning over 41 thousand SKUs produced by 125 firms (Table 2, second column). These firms are larger than the average and represent 14% of the initial AIOCD sample, their products cover 46% of all SKUs, and their sales account for 60% of the total sales. These firms operate in more than 2400 markets. Every firm in the merged sample produces on average almost 300 different SKUs, and operates in 124 different markets.

The additional information provided by the CMIE dataset implies a cost in the representativeness of the sample that we use for our estimations. Traded companies are typically larger and show different entry and exit dynamics. The entry rate of firms in the pharmaceutical industry is high (15%) and relevantly above the exit rate (6%), giving us the idea of a very dynamic environment.[5] Moreover, traded firms show a higher propensity for product and market switching than AIOCD companies. These features can result in selection biases. In our estimation we employ specific techniques to control for this problem.

In the pharmaceutical industry the products sold in every submarket are homogeneous in the sense that they have all the same chemical composition. Drugs in every submarket also have the same quality as the bulk drugs used to produce the final goods are standard

[4]The CMIE data was used in other works of the field: Ahsan (2013); De Loecker et al. (2016); Goldberg et al. (2010a); Topalova and Khandelwal (2011) among the others.

[5]Goldberg et al. (2010b) show that in Indian manufacturing industry, entering in a product-level market is more common than exiting, because high sunk costs make firms reluctant to withdraw production lines.

chemical elements. The only sources of differentiation among products in the same submarket are the pack size (measured in units, e.g., the number of single tablets in the pack) and the dosage strength (measured in weight or capacity, e.g., Metformin has dosages of 500, 850 or 1000 mg). To facilitate comparison among product prices per units (P) and physical units of product (Q) sold within the submarket, we normalize these measures (NI for $P, Q \in I$) for dosage strength (DS) and pack size (PS) using the formula:

$$NP = \frac{P}{DS \times PS}$$

$$NQ = Q \times DS \times PS$$

Note that multiplying the normalized price per unit (NP) by normalized quantity (NQ) gives product sales revenues. Comparability of quantities and prices between submarkets are not possible, but relative measures like quantity-based market share, intended as the ratio between normalized quantity of a product and the sum of all normalized quantities sold in the same submarket, and relative price, intended as the ratio between normalized price of a product and the highest normalized price for a product in the same submarket. We observe a large dispersion of market shares and prices within and across the submarkets (Table 2, Panel b).

3 Methodology

In Table 1 we show that in the Indian pharmaceutical industry the relation between submarket shares and prices is positive and significant. The economic theory predicts this relation to be negative, implying a trade-off in firm's profit-maximizing price-setting decision: increasing prices but losing market shares, or lowering prices but gaining market shares (Klemperer, 1995).[6] We aim to understand the forces that drive this relation, taking into account the others determinants of market shares that can affect the estimates. In this section we identify and discuss the determinants of product market share, distinguishing the role of the agents interacting in the pharmaceutical industry: manufacturing firms, retailers and consumers. We then propose a reduced form model to empirically disentangle the effect of demand- and supply-side factors, using instrumental variables to solve for possible endogeneity.

[6] In the Appendix we show that if price elasticity is higher than one in absolute value a marginal increase in the product price results in a decrease in revenue-based market share.

3.1 Market shares in multiproduct firms

A correct estimation of the effect of a price change on product market share needs to control for other product-, firm- and market-level variables that influence the relationship. Hottman et al. (2016) propose an equilibrium model that distinguishes the determinants of the market share of multiproduct firms into four main components: costs, markups, appeal and scope. The same variables can be considered at the product level and identified as the determinants of the market share of the product.

To introduce these controls in the specification for product market share we proxy them with product-level variables: i) cost is product *wholesale price*; ii) markup is *retailers' margin* on the product; iii) appeal is *brand and firm age*; iv) firm scope is the *number of markets* where the firm operates.

These variables can be seen as the contribution to product market share of the three agents operating in the pharmaceutical industry: manufacturing firms, retailers and consumers. The behaviour of both consumers and intermediaries, indeed, can help explain the positive relationship between manufacturer prices and market share. On the consumer side, the literature documented cases of lack of information in the pharmaceutical industry (Bronnenberg et al., 2015) implying higher and inelastic demand for branded and higher priced medicines. The theory of switching costs Klemperer (1995) can help understanding why consumers might have a strict preference on a good regardless of what the price is. This mechanism can help clarifying why we don't find a loss in market share when consumers face an increase in prices.

Intermediaries can play a role in maintaining the dominance of a product in its market, despite a relatively higher price. Indeed, manufacturers may incentivize intermediaries to promote their products offering higher retailer margins. Since pharmaceutical retailers in India are represented by a trade association, it is possible for them to coordinate the distribution of a specific product across the country. Linking the availability of a product to its retail margin may also affect the demand for that product as consumers' switching costs may depend on product availability.

In theory, products with a higher market share are those charging lower prices as relatively more productive. The presence of uninformed consumers and coordinated retailers, however, may affect the selection process of a product in the market, distorting the 'true' demand for the product. Therefore, it is also possible that the market-leading products selected through this process are relatively more expensive because relatively less productive.

We can, then, divide the determinants of product market share by the manufacturer into demand-side ones (brand and firm appeal and retailer's margin) and supply-side ones (product wholesale price and firm scope).

3.2 Empirical strategy

We estimate the relation between product market share and its determinants in the pharmaceutical industry extending the usual demand function, that relates quantities and prices, to incorporate the other variables pointed out in the previous section. The estimation equation is the following:

$$\begin{aligned} share_{ifjt} = \alpha_0 + \alpha_1 price_{ifjt} + \alpha_2 margin_{ifjt} + \alpha_3 scope_{ft} + \\ + \alpha_4 age_{ifjt} + \alpha_5 age_{ft} + \alpha_6 competition_{jt} + \delta_t + \eta_j + \epsilon_{it} \end{aligned} \tag{1}$$

where $share_{ifjt}$ is submarket j shares of revenues of product i produced by firm f, or submarket j shares of normalized physical units of product i produced by firm f The variable $price_{ifjt}$ can be either product i's normalized price to retailer or product i's normalized relative price to retailer. Retailer's margin is the difference between normalized wholesaler price and normalized retailer price for product i ($margin_{ifjt}$). Firm scope is the number of products produced by firm f in logs ($scope_{ft}$). Firm and brand appeal are captured by the years since the firm foundation (age_{ft}) and the years since the brand was launched (age_{ifjt}), respectively. We also control for market competition using the number of competing products in the market in logs ($competition_{jt}$), and for demand shift, including year (δ_t) and submarket (η_j) dummies, which adjust for any industry-wide and therapeutic category-wide demand variation, respectively. The error term (ϵ_{it}) is product-year specific.

Estimating (1) using OLS might give a positive bias to price coefficient, as an idiosyncratic shock in demand would imply an increase in price by the producers. To identify price elasticity of physical quantity sold Foster et al. (2008) instrument prices with a measure of physical (quantity-based) productivity, a supply-side price driver which contains information on firm's cost. Indeed, productivity is a measure of technical efficiency, directly comparable across plants in the same submarket. Moreover, it should not be correlated with idiosyncratic product-specific demand shocks in the short run. We adopt the same idea and instrument our measures of wholesale prices with a measure of quantity-based productivity, calculated addressing all the potential biases concerning productivity in multiproduct firms. In the following section we document how we calculate product-level productivity in the pharmaceutical industry.

4 Product-level productivity in multiproduct firms

4.1 Instrumenting prices: product-level productivity

Product-level productivity, intended as the efficiency with which firms turn inputs into outputs for each good they produce, is a key element of our analysis as it serves as instrument for avoiding endogeneity between market share and wholesale prices. Although productivity differences among firms within an industry - even narrowly defined industries composed of a homogeneous product - are well-documented (Bartelsman and Doms, 2000; Syverson, 2011), yet little empirical work examines the relation between prices and productivity and its impact on market shares (Goldberg and Hellerstein, 2012). Our identification strategy, relies specifically on the possibility that productivity is exogenous with respect to market shares and correlated with prices, and that this statement holds at the product level.

For this purpose we estimate product-level productivity in multiproduct firms. Indeed, although the choice of single-product industries simplifies the estimation of productivity and modeling competition, there is a growing recognition of the heterogeneity in products within standard industry classifications and their implications for market structure (Klepper and Thompson, 2006). The share of manufacturing firms that produce more than one good is large (Bernard et al., 2010; Goldberg et al., 2010b; Dhyne et al., 2017). Recent studies find that the choice of products mix is endogenous to the firm's productivity and, as a consequence, the usual firm-level and industry-level aggregate productivity measures that do not account for the product mix are biased (Bernard et al., 2009, 2010). With few exceptions, however, the literature on productivity treats firms as single-product producers, leaving aside issues like the choice of the firm product mix and price strategy in every market where the firm competes.[7] We compute six measures of product level productivity

[7] Among the exceptions: (a) theoretical model incorporating the product mix of multiproduct firms can be found in Bernard et al. (2009, 2010); Mayer et al. (2014); (b) applied works are Foster et al. (2008); De Loecker et al. (2016); Mayer et al. (2016).

for multiproduct firms building on the main and most up-to-date methodologies.

We consider a product-level log-additive production function (e.g. Cobb-Douglas) whose coefficients remain constant over the sample period:

$$x_{it} = tfp_{it} + \alpha k_{it} + \beta \mathbf{v}_{it} \tag{2}$$

where, for each product i and each year t, x is log output, k is log capital employed and v is a vector of variable inputs in logs. Product-specific log productivity (tfp) is Hicks-neutral and can be computed as a Solow residual. According to the nature of the output, the literature divides production functions in *revenue-based*, if output is measured in sales revenues y, and *quantity-based*, if output is measured in quantity of physical units sold q.

Estimating TFP of multiproduct firms at the product level encounters specific problems of feasibility involving variable existence, selection and identification. As discussed in De Loecker et al. (2016), the estimation of a product-level log-additive production function needs to take into consideration two main aspects: a) we do not observe product-level inputs, but only firm-level ones; b) we do not observe TFP (neither at the firm nor at the product level). The first aspect can lead to two potential biases: a1) an *input allocation bias*, related to the possible mismeasurement in the process of addressing shares of firm-level inputs to each product; a2) an *input price bias*, related to the differences in purchase prices of the same input across different markets and qualities. Unobserved (to the econometrician) TFP, instead, can lead to two other potential biases: b1) a *simultaneity bias*, as the amount of inputs is chosen based on firm or product TFP; b2) a *product scope bias*, as the number of products is decided by the firm according to observed TFP. In the next two subsections, we discuss how we address these potential biases.

4.2 Product's input measurement

4.2.1 Product's input allocation: the 'reference firm'

The most influential papers in the literature deal with input allocation either apportioning firm-level input values or introducing a method to avoid the problem. Foster et al. (2008) apportion product's share of plant inputs using product's share of plant sales.[8] De Loecker

[8]The method is valid under perfect competition or assuming constant markups across firm products. Since they select 11 four-digit industries producing homogeneous goods (concrete, gasoline, coffee among them) and highly product-specialized plants (at least 50 percent of plant's revenues are obtained from

et al. (2016) avoid product's input allocation by estimating TFP using only single product firms.[9] Dhyne et al. (2017) implement a technique to estimate product-level TFP using only firm-level inputs. We exploit specific features of the pharmaceutical industry to make some assumptions and reliably impute the values of product variable inputs.

The pharmaceutical industry is composed of a large number of submarkets selling homogeneous goods. The unit cost of variable inputs - that is, raw materials, labor, and energy - can be assumed to be the same in the single submarket. Since the chemical composition of each drug is unique, we can assume that the cost of raw materials (bulk-drugs) used to produce a unit of the drug does not vary across all firms of the same submarket. The Indian pharmaceutical industry, which overwhelmingly produces out-of-patent medicines, is arguably more labor intensive than its counterparts in the developed world, where R&D and innovation-related staff play an important role. Since the skills to be employed in the pharmaceutical industry is standardized at the submarket level, and given the highly automated production process, we assume that the cost of labor used to produce a unit of the drug does not vary across all firms of the same submarket. We make the same assumption for energy, whose unit cost is the same for all firms of the same submarket. To identify the cost per unit produced of each variable input we select a firm for each submarket, the one charging the lowest (normalized) price for the drug, which we assume to produce at the marginal cost. We refer to it as the 'reference firm' of the submarket.

The imputation of unit variable input cost for all products in each submarket leverages on the reference firms ($\bar{f}$). First, we calculate its input expenditure in the refereed submarket ($\bar{j}$) using the submarket's revenue shares of the firm ($s_{\bar{j}\bar{f}t}$):[10]

$$\mathbf{v}_{\bar{j}\bar{f}t} = \mathbf{v}_{\bar{f}t} s_{\bar{j}\bar{f}t} \tag{3}$$

Second, we split reference firm's input expenditure in the refereed submarket ($\mathbf{v}_{\bar{j}\bar{f}t}$) across all its products (i) using product's share of physical units produced in the submarket by

the product of interest), these assumptions are appropriate.

[9]They assume that a single-product firm uses the same technology of a multi-product firm to produce the same good. In a second stage, they use a system of equations based on firm-level TFP to allocate the inputs of multiproduct firms across products. They assume product share of firm's input to be the same across all different inputs.

[10]To do so we have to assume that the reference firm has constant markup over all inputs in the refereed submarket.

the firm ($s^q_{i\bar{j}\bar{f}t}$):[11]

$$\mathbf{v}_{i\bar{j}\bar{f}t} = \mathbf{v}_{\bar{j}\bar{f}t} s^q_{i\bar{j}\bar{f}t} \tag{4}$$

This makes possible to obtain a measure of unit input cost for all products of the reference firm in the submarket ($\mathbf{v}^*_{i\bar{j}\bar{f}t}$), dividing the input cost of each reference firm's product in the refereed submarket by the physical quantity of product sold ($q_{i\bar{j}\bar{f}t}$).

$$\mathbf{v}^*_{i\bar{j}\bar{f}t} = \frac{\mathbf{v}_{i\bar{j}\bar{f}t}}{q_{i\bar{j}\bar{f}t}} \tag{5}$$

Since we assumed the unit cost of variable inputs to be the same in the submarket, we can impute the input cost for all products of all other firms (f) in the submarket ($\bar{j}$) multiplying reference firm's unit input cost to every product's physical units produced in the submarket ($q_{i\bar{j}ft}$).

$$\mathbf{v}_{i\bar{j}ft} = \mathbf{v}^*_{i\bar{j}\bar{f}t} q_{i\bar{j}ft} \tag{6}$$

To impute product-level capital we simply apportion firm-level employed capital among the different products of the firm using product's share of firm sales.

4.2.2 Product's input price

In general, product price variation within an industry may depend on the difference in the quality of the goods, which in turn may imply different inputs quality, and different input unit costs. The pharmaceutical industry produces homogeneous goods at the same submarket level. Since every drug sold in the submarket is composed of a fixed chemical composition, product quality variation can be excluded as well as difference in variable input quality within the same submarket. The bulk-drugs to obtain the final drugs are standardized; the workers in the chain of one product do not need to be more skilled than the other workers in the same submarket; energy has no quality variation within the submarket.

Another cause of input prices variation across comparable products is the existence of local input markets with different prices for the same good. We can exclude how this source of price variation affects inputs of products of the same submarket. The market for intermediate inputs such as bulk drugs is organized at the international level as it involves

[11] Units produced are normalized to take into account both the selling size of the good (quantity of drugs in the pack) and the dosage strength.

chemical elements. In addition, workers for the same role have the same bargaining power in India, the energy market is nationwide (at least).

Specific features of Indian pharmaceutical industry, however, do not help to infer about the difference in price of the capital goods employed for producing each single product. We stick to the O-Ring theory by (Kremer (1993)) and to Kugler and Verhoogen (2011), which model and show, respectively, that more expensive inputs lead to more expensive products. Product's share of firm sales, that we use for apportioning firm-level employed capital among the different products incorporates this information.

An important assumption we make on input prices is that they do not depend on input quantities.[12] If this assumption is violated because the input market power of the reference firm - from which we calculate the unit cost of inputs - is high thanks to a high share of input purchased, our imputation method can generate problems. To help to validate this assumption, we verified that only 13% of the reference firms have the highest sales share in the refereed submarket, implying that less than 13% of the reference firms are top purchasers on their input markets.

4.3 Estimating product-level TFP

4.3.1 Addressing (or avoiding) the simultaneity bias

To calculate TFP as a Solow residual using equation (2) we must be able to measure the output elasticities, α and β. This can be done in two different ways: a) equaling elasticities to average input cost share over the sample (cost share-based method); b) estimating the elasticities by estimating the production function (estimation-based method). The first method follows the theoretical framework of cost minimization of the firm; the second one follows assumptions on the nature of productivity shocks and firm's information set.

While the cost share-based method is easy to construct, it is only valid under the assumption of perfect competition and constant returns to scale. Respective to the output indicator, we calculate two measures of cost share-based productivity, one revenue-based ($TFPt$) and one quantity-based ($TFPQt$) as in Foster et al. (2008). The first row of Table 3 shows the elasticities of the cost share-based production functions.

The estimation-based method, instead, faces the problem of simultaneity bias, which

[12] The same assumption is also maintained by De Loecker et al. (2016).

arises as input quantities are decided according to observed or expected TFP (Olley and Pakes, 1996). Using the estimator proposed by Levinsohn and Petrin (2003), we estimate four measures of estimation-based productivity, two revenue-based (*TFPVAlp* and *TFPRlp*) and two quantity-based (*TFPQRlp* and *TFPQKlp*). This method is commonly used in the empirical literature on Indian firms (Ahsan, 2013; Topalova and Khandelwal, 2011). The second and third rows of Table 3 show the elasticities of the estimation-based product-level production functions.

4.3.2 Addressing the product scope bias

Although widely used, the method proposed by Levinsohn and Petrin (2003) is not sufficient to eliminate all the sources of bias as it does not consider the product mix of the firm. Bernard et al. (2010) show that product switching is correlated to firm productivity and suggest that firms endogenously select the goods to produce. The product scope bias, thus, arises when estimating the production function in multiproduct firms. Given the wide heterogeneity in prices at the product level, the estimation of TFP turns out to be non-negligibly biased (De Loecker and Goldberg, 2014). Although the implications of these unsolved issues affect the measurement of firm-level TFP for all multiproduct firms, the empirical research on this field is still scant. Only recently, De Loecker et al. (2016) included a control for product-mix in the estimation of productivity in single-product firms. [13] Dhyne et al. (2014) and Dhyne et al. (2017), instead, offered some econometric approach to deal with price heterogeneity at the product level, proposing a specification which accounts for the firm product scope. As far as we are concerned, Dhyne et al. (2017) is the only empirical study addressing the product scope bias in estimating prodct-level productivity by multiproduct firms.[14].

Dhyne et al. (2017) incorporate the product scope of the firm in an estimation-based method for measuring output elasticities at the product level which uses only firm-level input measures. In general the production function of the firm is considered as a sum of single product production function (De Loecker et al., 2016). Dhyne et al. (2017) implement a multiproduct production function re-elaborating the contribution of Diewert (1973). They consider a loglinear Cobb-Douglas production function, specified as follows:

[13] De Loecker et al. (2016) use a sample of firms that have been single-product at least for one year in the time span. Their purpose is, actually, not to control for the product scope bias, but for a selection bias regarding the nature of firms which decide to change their product-mix.

[14] Both Foster et al. (2008) and De Loecker et al. (2016) calculate product-level productivity using a sample of single-product firms

$$q_{it} = tfp_{it} + \alpha k_{ft} + \beta_l l_{ft} + \beta_m m_{ft} + \gamma y_{-it} \tag{7}$$

where, for each product i, each firm f and each year t, q is log quantity sold in physical units, k is log capital, l is log labour, m is log raw materials and y_{-i} is log revenues of all other products produced by the firm. Adding this latter measure to the production inputs Dhyne et al. (2017) "extend the single product setting" calculating a production function which gives "the maximal amount of output achievable of one of the goods the firm produces holding inputs and the levels of other goods produced constant." Moreover, their method avoids the procedure of apportioning firm-level inputs to each product. We follow their idea estimating a hybrid production function, which is single-product with respect to the variable inputs and multiproduct with respect to the capital.

For the pharmaceutical industry variable inputs can be considered to be product-related as raw materials and labour can only be used to produce a specific (good or at least be reallocated in the short run to other product of the same submarket). Firm capital expenditure, instead, is more likely to involve many products, complicating the exercise of identifying the exact contribution of firm capital to each product. We propose the following hybrid production function, in which variable inputs enter at the product level and capital enters at the firm level:

$$q_{it} = tfp_{it} + \alpha k_{ft} + \beta_l l_{it} + \beta_m m_{it} + \gamma y_{-it} \tag{8}$$

To contrast the simultaneity bias, we adopt the standard Levinsohn and Petrin estimator using materials as a proxy. Since the introduction of y_{-it} causes problems of endogeneity, we instrument it with its lagged value, as suggested by Dhyne et al. (2017). We expect coefficient γ to be negative, as an increase in firm revenues, holding product variable inputs and firm-level capital constant, would result into a decrease in the quantity of the product under exam.

5 Results

5.1 First stage: quantity-based vs revenue-based productivity

The measures of product-level TFP calculated are collected in Table 3, distinguished by revenue- or quantity-based and cost share- or estimation based. All coefficients of the production function are calculated at the market level. The two cost share-based measures of productivity *TFPt* and *TFPQt* have the same equation (same input elasticities), they just differ for the output variable: revenues for *TFPt* and physical units for *TFPQt*. The two revenue-estimation-based measures of TFP differ according to the inclusion of raw materials in the output (subtracted to the revenues) or among the inputs. For *TFP-VAlp* we estimate a product value-added function, where energy costs serve as the proxy variable, as in (Ahsan, 2013). For *TFPRlp* we estimate a product sales function where raw material costs serve as the proxy variable, as in Dhyne et al. (2017), Table A3. The estimated elasticities of capital and labor are significantly lower for the latter measure. The quantity-based measures are both functions of product units, but differ from each other thanks to the control variables. *TFPQRlp* includes labour and raw materials among the inputs, where the latter also serves as a proxy variable, as in De Loecker et al. (2016). *TFPQKlp*, instead, uses firm-level capital input and addresses the product scope bias by including a measure of firm's scope among the explanatory variables, as in Dhyne et al. (2017), Table 3. The estimated elasticities of capital, labor and materials are very close for the latter measure. The negative coefficient for firm product scope (γ) confirms the expectations.

Table 4 shows a postitive and significant correlation between each of the revenue-based measures of TFP to another. The same happens for quantity-based measures. The correlation between revenue- and quantity based TFP, instead, depends on te type. Cost share-based measures are positively correlated each other, while estimation-quantity-based measures have a negative correlation with estimation-revenue-based TFP. A significant correlation can also be found between both (normalized) price measures and all six measures of product-level productivity. Interestingly, prices are positively correlated with the three revenue-based measures of TFP, but negatively correlated with the three quantity-based measures of TFP. Foster et al. (2008) find the same correlation signs and attribute the reason to the fact that, although calculated with deflated values, revenue-based TFP incorporates prices by definition. This implies that revenue-based TFP measures are not a suitable instrument for prices to identify elasticity in (1). Quantity-based TFP, instead, theoretically satisfies the exogeneity condition, necessary for that role. We estimate (1) using a physical measures of TFP to instrument product prices.

The first stage regression of the 2SLS estimator (equivalent to IV in results) regresses wholesale prices over productivity and gives a measure of relevance of the instrument. This relation, however, can also suggest some information about product price-setting of the firm and the role of technical efficiency in this decision. We calculate the first stage regression for all the six measure that we estimated (Table 5). Including the other determinants of market share in the equation - retailer margins, firm scope, firm and brand appeal and competition - the opposite direction of the coefficient of revenue- and quantity- based product-level productivity is confirmed.

The negative relation between quantity-based TFP and prices is not surprising: more efficient products have lower marginal costs allowing firms to charge lower prices for them. Yet many dynamic models predict that more productive firms set lower prices, forcing less productive firms to exit the industry and gaining market shares (Jovanovic, 1982; Hopenhayn, 1992; Jovanovic and MacDonald, 1994). The ensuing competitive process spurs the reallocation of production inputs from less productive plants and firms to more productive ones, fostering growth (Foster et al., 2016).[15] The industry-specific empirical analysis focused on the relation between productivity and prices show that on average an increase in the productivity levels within an industry leads to lower prices. For example, using data on ready-mixed concrete, Syverson (2007) shows that when producers have heterogeneous costs and sell one standardized product, competitive selection on costs lowers product prices. Our analysis on multiproduct firms shows that the relationship higher productivity-lower prices can be valid also at the product level. Important is to notice that price variation can only limitedly be imputed to quality differentials among products or input price variations within the submarket.

5.2 The determinants of market share

Running equation 1 with OLS (Table 6) still shows a positive relation between revenue-based market share and prices, remaining robust to the inclusion of other determinants

[15]Similarly, models of international trade predict that more productive firms enter into exporting, as they can cover transportation and other costs relative to less productive firms (Melitz, 2003; Mayer et al., 2014; Melitz and Redding, 2014).

of market share. Retailer' margin is negatively correlated with market share, in line to what one would expect in perfect competition, as an increase in margin increases price to retailers, i.e. the prices that makes the demand for the good. Firm scope's coefficient is also positive suggesting that market leadership is associated with the ability of the firm to produce more good (and then with productivity). Firm and brand age proxy product appeal and are both positive and significant. Hypothesis of uninformed consumers or switching costs may find an evidence in this result. A comparison among coefficients shows that the correlation of these determinants is similar for both quantity- and revenue-based market share. Only the coefficient of prices has opposed sign.

Given that the OLS estimates are supposed to be biased, following Foster et al. (2008), we instrument the price variable with a measure of quantity-based productivity calculated in Section 4. We decide to use the only measure that addresses product scope bias: *TFPQKlp*, but the results with the other two are very similar. Table 7 shows the results.

Instrumenting wholesaler prices with a physical measure of productivity turns its effect on revenue-based market share to be negative and significant. This result suggests that the initially observed positive relation between prices and revenue-based market shares is not due to distortions on the supply side, as a higher technical efficiency drives prices downwards. The coefficient related to retailer margin flips into positive, indicating a potential role of selection by retailers based on margins. Consumers-side factors, like firm and brand appeal remain positive determinants. Demand-side factors involving consumers and retailers are thus primarily responsible for the positive relation between prices and revenue-based market shares. This result remains consistent also in submarkets including more than 10 different products.

The coefficient between prices and quantity based market share was already negative using OLS. Estimating it using IV confirms the negative relation, but increases the size of the coefficient, as expected.

6 Conclusion

This paper examines the relationship between prices and market shares in multiproduct firms. We exploit a rich dataset on Indian pharmaceutical firms containing detailed information on quantities, wholesaler and retailer price of every drug sold in India. Without imposing assumptions on the market structure and competition among products, we estimate the demand curve that pharmaceutical products face in terms of market share in their therapeutic category market. We observe that market share is positively and robustly correlated with wholesaler prices. This puzzling direction of the relationship suffers, however, from estimation biases stemming from the simultaneous formation process of prices and quantities sold. Our approach, then, uses a measure of quantity-based product-level TFP to identify the effect on market share triggered by the supply side (an increase in wholesaler price) from the effect activated by the demand side, like an increase in retailer markups, or firm appeal or brand appeal.

To measure quantity-based product-level TFP in the multiproduct firms of the pharmaceutical industry, we address all the biases arising when estimating the production function at the product level. Building on the most up-to-dated advances in the literature and exploiting peculiar characteristics of the pharmaceutical industry, we introduce a new method for estimating the input elasticities of the production function which accounts also for the product scope of the firm.

We compare the TFP measure that we obtain to other measure of product-level productivity (both quantity- and revenue-based) calculated using the most recently introduced methodologies, and find that it is negatively correlated with prices as well all other quantity-based measures. Conversely the measures of revenue-based productivity are positively correlated with prices: the same evidence found by Foster et al. (2008).

We show that product wholesaler price, once instrumented with a quantity-based measure of productivity, turns to have a negative and significant effect on revenue-based market share. The positive relation between prices and revenue-based market shares that we observe in the OLS estimation is, thus, not an outcome of a supply-side distortions: within narrowly defined market, more productive products have lower wholesaler prices. Demand-side factors like brand and firm appeal, and especially retailer margins are, then, the upward drivers of the market shares. The IV analysis is based on data of listed firms, which are larger and more productive than the average. Therefore, these results are not representative of the entire pharmaceutical industry. However, they suggest the relevance of the buyer power in the pharmaceutical industry. The market power of the manufacturer is, thus, mainly related to its ability of keeping the brand appeal and possibility of allowing the retailers to obtain higher margins.

References

Ahsan, R. N. (2013). Input tariffs, speed of contract enforcement, and the productivity of firms in india. *Journal of International Economics*, 90(1):181–192.

Alfaro, L. and Chari, A. (2014). Deregulation, misallocation, and size: Evidence from india. *The Journal of Law and Economics*, 57(4):897–936.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *The American Economic Review*, 103(1):305–334.

Bartelsman, E. J. and Doms, M. (2000). Understanding productivity: Lessons from longitudinal micro-data. *Journal of Economic literature*, 38(3):569–594.

Bennett, D. and Yin, W. (2014). The market for high-quality medicine. Technical report, National Bureau of Economic Research.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2009). Products and productivity. *The Scandinavian Journal of Economics*, 111(4):681–709.

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

Bhaskarabhatla, A., Chatterjee, C., Anurag, P., and Pennings, E. (2016a). Mitigating regulatory impact: the case of partial price controls on metformin in india. *Health policy and planning*, 32(2):194–204.

Bhaskarabhatla, A., Chatterjee, C., and Karreman, B. (2016b). Hit where it hurts: Cartel policing using targeted sales and supply embargoes. *The Journal of Law and Economics*, 59(4):805–846.

Bollard, A., Klenow, P. J., and Sharma, G. (2013). India's mysterious manufacturing miracle. *Review of Economic Dynamics*, 16(1):59–85.

Bronnenberg, B., Dube, J.-P., Gentzkow, M., and Shapiro, J. (2015). Do pharmacists buy bayer? informed shoppers and the brand premium. *Quarterly Journal of Economics*, 130(4):1669–1726.

De, P. K. and Nagaraj, P. (2014). Productivity and firm size in india. *Small Business Economics*, 42(4):891–907.

De Loecker, J. and Eeckhout, J. (2017). The rise of market power and the macroeconomic implications. Technical report, National Bureau of Economic Research.

De Loecker, J. and Goldberg, P. K. (2014). Firm performance in a global market. *Annu. Rev. Econ.*, 6(1):201–227.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

Dhyne, E., Petrin, A., Smeets, V., and Warzynski, F. (2014). Import competition, productivity and multi-product firms. Technical report, Working Paper Research.

Dhyne, E., Petrin, A., Smeets, V., and Warzynski, F. (2017). Multi product firms, import competition, and the evolution of firm-product technical efficiencies. *National Bureau of Economic Research*.

Diewert, W. E. (1973). Functional forms for profit and transformation functions. *Journal of Economic Theory*, 6(3):284–316.

Foster, L., Grim, C., and Haltiwanger, J. (2016). Reallocation in the great recession: cleansing or not? *Journal of Labor Economics*, 34(S1):S293–S331.

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

Goldberg, P. and Hellerstein, R. (2012). A structural approach to identifying the sources of local currency price stability. *Review of Economic Studies*, 80(1):175–210.

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

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

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly journal of economics*, 124(4):1403–1448.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, pages 649–670.

Jovanovic, B. and MacDonald, G. M. (1994). The life cycle of a competitive industry. *Journal of Political Economy*, 102(2):322–347.

Klemperer, P. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *The review of economic studies*, 62(4):515–539.

Klepper, S. and Thompson, P. (2006). Submarkets and the evolution of market structure. *The RAND Journal of Economics*, 37(4):861–886.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–575.

Kugler, M. and Verhoogen, E. (2011). Prices, plant size, and product quality. *The Review of Economic Studies*, 79(1):307–339.

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

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity... for some plants*. *The Quarterly Journal of Economics*, 125(3):1051.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. (2014). Market size, competition, and the product mix of exporters. *The American Economic Review*, 104(2):495–536.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. (2016). Product mix and firm productivity responses to trade competition. *National Bureau of Economic Research*.

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

Melitz, M. J. and Redding, S. J. (2014). Heterogeneous firms and trade. *Handbook of International Economics,*, 4:1–54.

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

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

Smeets, V. and Warzynski, F. (2013). Estimating productivity with multi-product firms, pricing heterogeneity and the role of international trade. *Journal of International Economics*, 90(2):237–244.

Syverson, C. (2007). Prices, spatial competition and heterogeneous producers: an empirical test. *The Journal of Industrial Economics*, 55(2):197–222.

Syverson, C. (2011). What determines productivity? *Journal of Economic literature*, 49(2):326–365.

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

Van Biesebroeck, J. (2005). Exporting raises productivity in sub-saharan african manufacturing firms. *Journal of International Economics*, 67:373–391.

Tables

Table 1: Market share and wholesale prices

	Revenue-based market share				Quantity-based market share			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	+10	+10	All	All	+10	+10
Normalized PTR	0.025***		0.019***		-0.116***		-0.057***	
	(0.006)		(0.004)		(0.007)		(0.005)	
Relative PTR		0.023***		0.007***		-0.086***		-0.042***
		(0.003)		(0.002)		(0.003)		(0.002)
Observations	253656	253656	198863	198863	253656	253656	198863	198863
R-squared	0.663	0.663	0.103	0.103	0.644	0.647	0.099	0.104

Source: AIOCD and Prowess,CMIE.
Notes: OLS estimates. Market share is the dependent variable. *Revenue-based market share* is calculated as the share of sales of the product in the submarket. *Quantity-based market share* is calculated as the product share of units sold in the submarket. *Normalized PTR* is calculated as the product price to retailers divided by the pack size (in units) and dosage strength (measured in weight or capacity). Observations belonging to top and bottom 1% values of normalized PTR are truncated. *Relative PTR* is the ratio between product's normalized PTR and the highest normalized PTR in the submarket. Time and submarket fixed effects are included. Set *All* includes all the submarkets, set *+10* includes the submarkets with 10 or more products.
$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Sample statistics

Panel a	AIOCD	AIOCD + CMIE
Firms (N)	901	125
Markets (N)	2,904	2,458
Submarkets (N)	7,842	5,562
Brands (N)	51,580	21,527
Products (N)	88,902	41,291
Product-year observations	259,556	115,300
Product-month observations	2,452,144	1,097,420
Exiting firms (%)	6.0	2.7
Entering firms (%)	15.6	4.9
Firms changing product mix (%)	87.3	93.7
Firms changing market mix (%)	77.0	87.8

Panel b	AIOCD				AIOCD + CMIE			
	mean	med	min	max	mean	med	min	max
Products per firm (N)	101	44	1	2451	298	136	1	2451
Markets per firm (N)	47	24	1	731	47	123	1	731
Submarkets per firm (N)	61	31	1	1103	163	89	1	1103
Products per market (N)	24	7	1	3422	13	5	1	395
Products per submarket (N)	10	2	1	1792	6	2	1	185
Sales per product (IRs)	11.3 m	665 t	1	3.2 b	15.6 b	1.5 m	1	2.3 b
Sales per firm (IRs)	1.2 b	24.2 m	23.2	45.2 b	4.6 b	890 m	156	45.2 b
Sales per market (IRs)	272 m	45.9 m	3	15.0 b	174 m	32.5 m	3	6.7 b
Sales per submarket (IRs)	109 m	9.6 m	3	9.5 b	78.0 m	9.9 m	1	5.9 b
Product revenue submarket share (%)	10.4	0.3	1e-07	100	11.7	0.7	1e-07	100
Product unit submarket share (%)	10.4	0.4	6e-07	100	11.7	0.7	7e-07	100
PTR per unit (IRs)	191	45.2	0.01	264 t	269	50	0.01	264 t
Relative PTR (%)	30.8	13.8	3e-06	100	32.6	16.4	3e-06	100

Source: AIOCD and Prowess,CMIE.
Notes: The dataset AIOCD + CMIE is the data obtained merging AIOCD and CMIE datasets. *market* refers to the therapeutic category, *submarket* refers to the combination between therapeutic category and format of the drug (tablet, injection, syrup, etc.), *brand* refers to the retail name of the drug (irrespective of the strength or dosage), *product* refers to the SKU. *Exiting* and *entering firms* are the firms which appear or disappear in the dataset. *Product mix* changes when a firm starts producing or stops producing a stock keeping unit, *market mix* changes when a firm enters a therapeutic category with a new drug or exits a therapeutic category ceasing producing drugs for that market. *Sales* and *PTR* are expressed in Indian Rupees (IRs) where *b* is billion, *m* is million, *t* is thousand. *Product revenue submarket share* is the ratio between product sales and total submarket sales. *Product unit submarket share* is the ratio between normalized product units and the total normalized units sold on the submarket. *PTR per unit* is price to retailer for one unit of product. *Relative PTR* is the ratio between normalized product PTR and the highest normalized product PTR in the submarket.

Table 3: TFP estimates

		Revenue-based	Quantity-based
Cost share -based	Label	TFPt	TFPQt
	Function	$y_{it} = tfp_{it} + \alpha k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_e e_{it}$	$q_{it} = tfp_{it} + \alpha k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_e e_{it}$
	Elasticities	$\alpha = .75; \beta_l = .08; \beta_m = .16; \beta_e = .02$	$\alpha = .75; \beta_l = .08; \beta_m = .16; \beta_e = .02$
	References	Foster et al. (2008)	Foster et al. (2008)
Estimation -based	Label	TFPVAlp	TFPQRlp
	Function	$va_{it} = tfp_{it} + \alpha k_{it} + \beta_l l_{it}$	$q_{it} = tfp_{it} + \alpha k_{it} + \beta_l l_{it} + \beta_m m_{it}$
	Proxy	e_{it}	m_{it}
	Conditioning	-	$k_{it}; k_{it-1}; l_{it-1}; m_{it-1}; m_{it-2}$
	Elasticities	$\alpha = .79; \beta_l = .23$	$\alpha = .39; \beta_l = .09; \beta_m = .34$
	References	Ahsan (2013)	De Loecker et al. (2016)
	Label	TFPRlp	TFPQKlp
	Function	$y_{it} = tfp_{it} + \alpha k_{it} + \beta_l l_{it} + \beta_m m_{it}$	$q_{it} = tfp_{it} + \alpha k_{ft} + \beta_l l_{it} + \beta_m m_{it} + \gamma y_{-it}$
	Proxy	m_{it}	m_{it}
	Conditioning	$k_{it}; k_{it-1}; l_{it-1}; m_{it-1}; m_{it-2}$	$k_{ft}; k_{ft-1}; l_{it-1}; m_{it-1}; m_{it-2}; y_{-it-1}$
	Elasticities	$\alpha = .55; \beta_l = .13; \beta_m = .20$	$\alpha = .35; \beta_l = .15; \beta_m = .34; \gamma = -.19$
	References	Dhyne et al. (2017), Tab.A3	Dhyne et al. (2017), Tab.3

Source: AIOCD and Prowess,CMIE.
Notes: *Revenue-based* and *Quantity-based* refer to the output measure of the production function: product sales and product physical units sold, respectively. *Cost share-based* and *Estimation-based* refer to the methodology of calculation of output elasticities. The first equals elasticities to the average cost share of the market, the second estimates them at the market level using Levinsohn and Petrin method. The variables indicated in the production functions are expressed in logs: y is revenue sales, q is physical units sold, va is value added calculated as the difference between revenue sales and raw materials, k is capital employed, l is salary, m is raw materials, e is power and fuel expenses. Subscripts i, $-i$, f and t indicate product, all other products but product i, firm and year. The estimated elasticities reported are the industry average of market-wise estimates.

Table 4: Correlation table

	NPTR	RPTR	Q	TFPt	TFPVAlp	TFPRlp	TFPQt	TFPQlp	TFPQKlp
NPTR	1.000								
RPTR	0.195	1.000							
	(0.000)								
Q	-0.017	-0.033	1.000						
	(0.000)	(0.000)							
TFPt	0.014	0.019	0.005	1.000					
	(0.000)	(0.000)	(0.100)						
TFPVAlp	0.018	0.005	-0.000	0.019	1.000				
	(0.000)	(0.095)	(0.914)	(0.000)					
TFPRlp	0.035	0.083	0.023	0.047	0.045	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
TFPQt	-0.460	-0.287	0.116	0.207	-0.018	-0.055	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
TFPQlp	-0.010	0.011	0.002	0.003	-0.004	-0.016	0.017	1.000	
	(0.001)	(0.000)	(0.496)	(0.333)	(0.248)	(0.000)	(0.000)		
TFPQKlp	-0.011	0.015	0.004	-0.003	-0.002	-0.009	0.031	0.255	1.000
	(0.000)	(0.000)	(0.174)	(0.307)	(0.585)	(0.003)	(0.000)	(0.000)	

Source: AIOCD and Prowess,CMIE.
Notes: *NPTR* is normalized price to retailers (PTR), top and bottom 1% values are truncated. *RPTR* is normalized relative PTR: the ratio between normalized product PTR and the highest normalized product PTR in the submarket. *Q* is a normalized measure of product units sold. The TFP measures (in logs) are defined as follows. *TFPt*: revenue-based TFP calculated using the share of costs (Foster et al. (2008) call it "traditional TFP"); *TFPVAlp*: value added-based TFP, based on Ahsan (2013); *TFPRlp*: revenue-based TFP calculated using Levinsohn and Petrin (2003); *TFPQt*: quantity-based TFP calculated using the share of costs (Foster et al. (2008) call it "quantity TFP"); *TFPQlp*: quantity-based TFP, based on De Loecker et al. (2016); *TFPQKlp*: quantity-based TFP calculated, based on Dhyne et al. (2017). See Table 3 for details.
$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: First stage: wholesale price and productivity

	Revenue-based productivity			Quantity-based productivity		
	(1) TFPt	(2) TFPVAlp	(3) TFPRlp	(4) TFPQt	(5) TFPQlp	(6) TFPQKlp
TFP	0.009***	0.001***	0.002***	-0.017***	-0.000	-0.001**
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Margin	0.478***	0.466***	0.638***	0.448***	0.556***	0.512***
	(0.108)	(0.110)	(0.110)	(0.103)	(0.146)	(0.136)
N rivals brands	0.004*	0.007***	0.002	0.003	0.004	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Firm scope	0.001	-0.001*	-0.000	0.000	-0.001	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Firm age	0.003*	0.004**	0.003*	0.022***	0.003	0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Brand age	-0.000	-0.001	-0.001*	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	105463	95181	87224	105463	86692	80917
R-squared	0.730	0.728	0.735	0.744	0.730	0.719

Source: AIOCD and Prowess,CMIE.
Notes: OLS estimates. The dependent variable is normalized price to retailers. The TFP measures (in logs) are defined as follows. *TFPt*: revenue-based TFP calculated using the share of costs (Foster et al., 2008); *TFPVAlp*: value added-based TFP calculated using Levinsohn and Petrin (2003) like in (Ahsan, 2013); *TFPRlp*: revenue-based TFP calculated using Levinsohn and Petrin (2003); *TFPQt*: quantity-based TFP calculated using the share of costs (Foster et al., 2008); *TFPQlp*: quantity-based TFP calculated using Levinsohn and Petrin (2003) like in De Loecker et al. (2016); *TFPQKlp*: quantity-based TFP calculated using Levinsohn and Petrin (2003) like in Dhyne et al. (2017). See Table 3 for details. $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: The determinants of submarket shares

	Revenue-based market share				Quantity-based market share			
	(1) All	(2) All	(3) +10	(4) +10	(5) All	(6) All	(7) +10	(8) +10
Normalized PTR	0.030***		0.021***		-0.115***		-0.058***	
	(0.006)		(0.005)		(0.007)		(0.005)	
Relative PTR		0.018***		0.012***		-0.095***		-0.038***
		(0.003)		(0.002)		(0.003)		(0.002)
Margin	-0.010*	-0.006	-0.005	-0.002	-0.012**	-0.022**	-0.006*	-0.012
	(0.005)	(0.004)	(0.003)	(0.001)	(0.005)	(0.010)	(0.003)	(0.008)
N rivals brands	-0.158***	-0.155***	-0.023***	-0.022***	-0.157***	-0.174***	-0.023***	-0.028***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)
Firm scope	0.004***	0.004***	0.003***	0.003***	0.005***	0.003***	0.004***	0.003***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Firm age	0.013***	0.013***	0.009***	0.009***	0.011***	0.011***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Brand age	0.029***	0.029***	0.019***	0.019***	0.029***	0.028***	0.019***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	239862	239862	188259	188259	239862	239862	188259	188259
R-squared	0.681	0.681	0.150	0.150	0.662	0.667	0.139	0.143

Source: AIOCD and Prowess,CMIE.
Notes: OLS estimates. Market share is the dependent variable. *Revenue-based market share* is calculated as the share of sales of the product in the submarket. *Quantity-based market share* is calculated as the product share of units sold in the submarket. *Normalized PTR* is calculated as the product price to retailers divided by the pack size (in units) and dosage strength (measured in weight or capacity). Observations belonging to top and bottom 1% values of normalized PTR are truncated. *Relative PTR* is the ratio between product's normalized PTR and the highest normalized PTR in the submarket. Time and submarket fixed effects are included. Set *All* includes all the submarkets, set *+10* includes the submarkets with 10 or more products.
$p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: The determinants of submerket shares
Instrumenting wholesale prices with quantity-based TFP (TFPQKlp)

	Revenue-based market share				Quantity-based market share			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	+10	+10	All	All	+10	+10
Normalized PTR	-3.198***		-2.125***		-3.744***		-2.506***	
	(0.353)		(0.237)		(0.392)		(0.268)	
Relative PTR		-0.978***		-0.629***		-1.145***		-0.742***
		(0.071)		(0.044)		(0.073)		(0.047)
Margin	1.641***	0.335***	1.168***	0.222***	1.842***	0.313***	1.332***	0.216***
	(0.547)	(0.091)	(0.348)	(0.052)	(0.623)	(0.094)	(0.397)	(0.053)
N rivals brands	-0.128***	-0.336***	-0.033***	-0.134***	-0.124***	-0.367***	-0.032***	-0.150***
	(0.009)	(0.017)	(0.005)	(0.009)	(0.010)	(0.018)	(0.006)	(0.010)
Firm scope	-0.006*	-0.011***	-0.004	-0.006***	-0.006	-0.012***	-0.003	-0.006***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.004)	(0.002)	(0.002)	(0.002)
Firm age	0.032***	0.030***	0.025***	0.021***	0.032***	0.029***	0.025***	0.020***
	(0.007)	(0.005)	(0.005)	(0.003)	(0.008)	(0.005)	(0.006)	(0.003)
Brand age	0.027***	0.019***	0.022***	0.018***	0.027***	0.017***	0.021***	0.016***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.004)	(0.002)	(0.002)	(0.001)
Observations	80917	80917	64776	64776	80917	80917	64776	64776

Source: AIOCD and Prowess,CMIE.
Notes: IV estimates. Price variables are instrumented with the quantity-based TFP measures based on Dhyne et al. (2017) (*TFPQKlp*),. See Table 3 for more details. *Revenue-based market share* is calculated as the share of sales of the product in the submarket. *Quantity-based market share* is calculated as the product share of units sold in the submarket. *Normalized PTR* is calculated as the product price to retailers divided by the pack size (in units) and dosage strength (measured in weight or capacity). Observations belonging to top and bottom 1% values of normalized PTR are truncated. *Relative PTR* is the ratio between product's normalized PTR and the highest normalized PTR in the submarket. *All* includes all the submarkets, *+10* includes the submarkets with 10 or more products. $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A The effect of a price change on market share

Define revenue-based market share of product k in market j as:

$$MS_{kfjt} = \frac{y_{kfjt}}{y_{jt}} = \frac{p_{kfjt}q_{kfjt}}{\sum_{i\epsilon j} p_{ifjt}q_{ifjt}}$$

with $i \epsilon [1, ..., k, ..., n]$

Define relative price and quantity of product i in market j with respect to product n:

$$p^*_{ifjt} = \frac{p_{ifjt}}{p_{nfjt}}; q^*_{ifjt} = \frac{q_{ifjt}}{q_{nfjt}}$$

and rewrite the market share of product k as:

$$MS_{kfjt} = \frac{p^*_{kfjt}q^*_{kfjt}}{\sum_{i\epsilon j} p^*_{ifjt}q^*_{ifjt}}$$

Quantities sold of product k depend on a change in p^*_{kfjt}, but also quantities of product i, competing with product k, depend on p^*_{kfjt}: $q^*_{ifjt}(p^*_{kfjt})$; as well as price of product i: $p^*_{ifjt}(p^*_{kfjt})$. The change in market share of an increase in price for product k is then given by:

$$\frac{\partial MS_{kfjt}}{\partial p^*_{kfjt}} = \frac{\partial \frac{p^*_{kfjt}q^*_{kfjt}}{\sum_{i\epsilon j} p^*_{ifjt}q^*_{ifjt}}}{\partial p^*_{kfjt}} = \frac{\partial \frac{p^*_{kfjt}q^*_{kfjt}(p^*_{kfjt})}{\sum_{i\epsilon j} p^*_{ifjt}(p^*_{kfjt})q^*_{ifjt}(p^*_{kfjt})}}{\partial p^*_{kfjt}}$$

For N competitors in market j a change in price of product k introduces a change in market share of product k by:

$$\frac{\partial MS_{kfjt}}{\partial p^*_{kfjt}} = \sum_{i \neq k \epsilon j} q^*_{ifjt}p^*_{ifjt}(1 + \epsilon_{kk} - \epsilon_{ik}) - \sum_{i \neq k, n \epsilon j} q^*_{ifjt}p^*_{kfjt}\frac{\partial p^*_{ifjt}}{\partial p^*_{kfjt}}$$

where ϵ_{kk} is price elasticity of product k and ϵ_{ik} is cross price elasticity of product i to price of product k:

$$\epsilon_{kk} = \frac{\partial q_{kfjt}}{\partial p^*_{kfjt}}\frac{p^*_{kfjt}}{q_{kfjt}}; \epsilon_{ik} = \frac{\partial q_{ifjt}}{\partial p^*_{kfjt}}\frac{p^*_{kfjt}}{q_{ifjt}}$$

Market share is expected to be decreasing in prices as the latter term $(q^*_{ifjt}p^*_{kfjt}\frac{\partial p^*_{ifjt}}{\partial p^*_{kfjt}})$ is always non-negative, since the other producers will react to an increase in prices of k increasing or not changing price of i. The first term $(q^*_{ifjt}p^*_{ifjt}(1 + \epsilon_{kk} - \epsilon_{ik}))$ is expected to be negative as well since cross price elasticity ϵ_{ik}) must be positive and elasticity ϵ_{kk}) negative and higher than one in absolute value.[16]

[16] Foster et al. (2008) find that in all NACE2 manufacturing sectors the value of elasticity exceeds one in absolute value.