Productivity, prices and market shares in multiproduct firms*

Gianluca Antonecchia
Erasmus School of Economics and Tinbergen Institute †

Ajay Bhaskarabhatla
Erasmus School of Economics ‡

January 16, 2020

Abstract

Pharmaceutical pricing is a contentious issue globally and poses unique challenges for emerging economies as it balances industry, trade, and health policies. We examine the relationship between product market shares and prices in the Indian pharmaceutical industry featuring multiproduct firms. Using detailed data on product-level sales and prices for 8000 narrowly defined markets (drug formulations), we observe that product wholesale and retail prices are positively correlated with product market shares. We decompose product market share into the contributions of wholesale price, retail markup, firm scope and product appeal, and identify the causal effect of a price change on product market share instrumenting for wholesale price with product-level productivity. We find that a one percent increase in wholesale price reduces market share by approximately one percent and retail markups have a small but significant negative effect on market shares. For the market leaders, on the contrary, an increase in wholesale price has a positive effect on product market share. Leaders also benefit from offering higher retail margins and face a relatively more inelastic demand. We also find that wholesale prices are correlated negatively with product-level productivity, suggesting that, although productivity differences induce price competition, they do not necessarily improve access to medicines in the presence of manufacturer market power and retailer buyer power.

*We thank Serguey Braguinsky, Bruno Cassiman, Benoit Crutzen, Florin Maican, Gloria Moroni, Melissa Newham, Nina Pavcnik, Amil Petrin, Mark Roberts, Jo Van Biesebroeck, and the seminar participants at the University of Miami (NAPW), ISI New Delhi (ACEGD), University of Ghent (Workshop on Firm Heterogeneity), University of Mannheim (MaCCI), University of Warwick (RES), University of Michigan (CAED), University of Barcelona Pompeu Fabra (EARIE) and Erasmus University of Rotterdam (ESE and TI). 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 one of India's best-selling medicines with more than 200 brands competing in the market. The leading brand in the market, *Augmentin*, charges a wholesale price of 206 Rupees for ten 625 mg tablets. By contrast, for the same number of tablets and dosage strength, the brand with the second-highest market share charges a relatively lower wholesale price of 63 Rupees. An analysis of more than 8000 narrowly defined medicine markets in India, consisting of the combination of an active ingredient and dosage form, during April 2011 and March 2016 confirms the relationship observed in the Amoxicillin example: on average, products with a higher wholesale price have a higher market share. As shown in Table 1, this pattern holds even after controlling for the number of rival products, and firm, market, and year fixed effects.

The positive correlation between prices and market shares might not seem surprising, given the high costs of innovation, the presence of patent monopolies, the prevalence of non-price competition through marketing investments, the inelastic nature of demand for pharmaceutical products, and the mechanical correlation between prices and revenues (Lakdawalla, 2018; Gaynor et al., 2015; Chernew et al., 2018; Berndt, 2002). However, the Indian pharmaceutical context characterized by scores of competing brands, selling mostly off-patent medicines, and consumers directly purchasing medicines as out-of-pocket expenses, suggests intense competition in these markets. Indeed, many scholars suggest low prices and high levels of competition as reasons for not regulating these markets (Berndt and Cockburn, 2014). Nevertheless, if confirmed, a positive relationship between prices and market shares, to the degree indicated by the Amoxicillin example, would point to the high cost of purchasing medicines, an overwhelming majority of which are out of patent protection, in one of the world's most populous countries characterized by negligible levels of health insurance and limited public provision of healthcare.

This paper examines the relationship between prices and market shares in the Indian pharmaceutical industry. Estimating the relationship between prices and market shares poses several challenges. First, the positive correlation observed in Table 1 is likely to be biased because prices and demand are simultaneously determined and other factors that might influence the relationship are omitted. Second, firms in our context produce multiple products, often more than a hundred, complicating efforts to infer product-level costs that influence pricing decisions. In addition, firms producing multiple products adopt pricing strategies that leverage their scope further complicating the assessment of the relationship between prices and market shares. Third, differences in prices can stem either from differences in productivity or markups, and isolating these effects poses additional challenges. While higher productivity leads to lower prices, higher demand, and a higher market share, with inelastic demand, more productive products may not compete on prices, resulting in a positive correlation between prices and market shares. Fourth, the use of vertical price restraints can alter the relationship between prices and market shares (Asker and Bar-Isaac, 2014). The presence of retailer buyer power can create incentives to promote the sale of brands manufactured by less-productive products that offer higher retail markups at the expense of more productive products with lower retail markups. This practice would imply an allocative inefficiency in production. On the other hand, more productive manufacturers may gain market share by offering higher retail markups, indicating that the inefficiency on the supply side lies with the buyer power of the downstream intermediaries rather than the manufacturers' market power. Fifth, if higher prices signal higher quality then unobserved quality differences are a possible explanation for why higher prices can be correlated with higher market shares but quality is not easily unobserved (Bronnenberg et al., 2015).

Using detailed data on product-level sales and wholesale prices coupled with firm-level financial information, we identify the effect of price on product market shares in multiproduct firms. Building on Hottman et al. (2016), we model the contributions to product market share of the following drivers: i) wholesale price; ii) retail markup; iii) firm scope; and iv) firm and product appeal. We address the simultaneity bias by using product-level productivity as an instrument for prices. Following Foster et al. (2008), we expect productivity, a proxy for technical efficiency in production, to be correlated with prices but not with the product market share except through its impact on prices. Building on recent advances in productivity estimation for multiproduct firms, we overcome challenges posed by the multiproduct nature of our data. Since our data record wholesale and retail prices separately, we observe retail markups for each product, allowing us to address isolate the effects of wholesale prices from markups. We exploit variation in brand launch dates and firm fixed effects to control for quality differences.

We find that the positive relationship between prices and market shares described in Table 1 is robust to the inclusion of other determinants of market shares. However, once correctly identified, the sign of this relationship turns negative: a one percent increase in relative wholesale prices reduces market share by approximately one percent. This finding does not hold for the market leaders, however, whose price effect on market share is strongly positive. Retail markups have, on average, a small but significant negative effect on market shares but we observe a positive effect for the market leaders. Indeed, market share increases by 0.3 percent for market leaders if they offer a one-percent higher retail markup. Similar to Foster et al. (2008), we find that wholesale prices are correlated negatively with quantity-based productivity, suggesting that productivity induces price competition.[1] These findings indicate that market leaders are insulated from price competition and that selection on productivity does not always result in lower prices, particularly in the presence of manufacturer market power and significant retailer incentives.

Our paper makes two main contributions. First, we contribute to the literature on estimating the effect of a price change on market shares at the product level by building on the framework of Hottman et al. (2016). We also build on Foster et al. (2008) for identification by instrumenting for wholesale prices with quantity-based product-level productivity. More broadly our paper contributes to the empirical literature on productivity estimation. Adapting traditional methods (Levinsohn and Petrin, 2003) and recent advances (De Loecker et al., 2016; Dhyne et al., 2017) to the pharmaceutical context, we estimate productivity at the product-level in multiproduct firms. In our dataset a product is defined at the stock keeping unit (SKU) level, the most disaggregated level possible. Our sample of pharmaceutical products allows us to calculate product-level productivity for almost 41,000 SKUs.[2] Product-level productivity estimation can be biased due to product-level input allocation (Foster et al., 2008; De Loecker et al., 2016); input price differentials (Katayama et al., 2009; De Loecker and Goldberg, 2014), simultaneity of productivity with input choice (Olley and Pakes, 1996; Ackerberg et al., 2015), and with the product scope of the firm (Bernard et al., 2010; Dhyne et al., 2017). We address these biases in estimating quantity-based product-level productivity and compare it to other revenue- and quantity-based measures. The instrumental variable approach we use to identify the effect of prices on market shares requires an analysis of the relationship between product price heterogeneity and productiivty (Foster et al., 2008; Goldberg and Hellerstein, 2012).[3] We find that prices are negatively correlated with productivity when quantity-based and positively correlated when revenue-based. This result confirms the findings in Foster et al. (2008), indicating that revenue-based productivity, used extensively in prior literature, is a distorted measure of technical efficiency. We also document differential effects of retail markups and margins on product demand and revenues. We show that retail margin affects the market share of the market leading products positively, helping to maintain the dominant position of the leaders in markets with highly

[1]Productivity can be distinguished between revenue-based and quantity-based, depending on whether the output variable in the production function is expressed in revenues or physical quantity.

[2]For 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.

[3]Smeets and Warzynski (2013) show that, when firms are multiproduct, the results of a firm-level analysis on prices can be largely biased as average firm-level price measures eclipse product-level heterogeneity in prices.

substitutable products. Our paper connects to the emerging literature on the rise in global (De Loecker and Eeckhout, 2018) and industry-specific markups (Berry et al., 2019). Our results also relate to recent studies documenting the role of intermediaries on competition and welfare (Hastings et al., 2017; Grennan et al., 2018; Craig et al., 2018; Starc and Swanson, 2018).

Second, our paper contributes to the broader literature on the pricing of pharmaceuticals. Pharmaceutical pricing is a contentious issue globally and poses unique challenges for India as it balances industry, trade, and health policies (Chaudhuri et al., 2006; Duggan et al., 2016). Indian pharmaceutical firms account for a significant share of exports of low-cost generic medicines to countries around the world. Yet, as our study highlights, despite the large number of competitors, Indian pharmaceutical markets exhibit limited price competition. Pharmaceutical manufacturers enjoy market power through branding of largely generic medicines coupled with laws that prevent substitution among them. Physicians prescribe brands rather than generics and retailers limit competition by agreeing to uniform margins of 24 to 30 percent (Bhaskarabhatla et al., 2016). Therefore, despite being more productive, larger pharmaceutical firms charge higher prices to sustain retailer incentives. Our paper offers evidence consistent with this narrative and motivates the need for a deeper examination of the role of the retailers. Responding to the need for more studies on firm productivity in emerging economies (Syverson, 2011) and pharmaceutical markets (Lakdawalla, 2018), we provide evidence on the relationship between market shares, prices and productivity in the Indian pharmaceutical industry. The Indian context has gained attention in the literature, as the country has undergone several important reforms over the last three decades. Our paper contributes an industry case study to the broader literature on India while previous studies examine the impact of trade reforms in India on price and markups (De Loecker et al., 2016; Alfaro and Chari, 2014), product scope (Goldberg et al., 2010a), productivity (Topalova and Khandelwal, 2011; Ahsan, 2013), multiproduct firms (Goldberg et al., 2010b).

The paper proceeds as follows. In Section 2 we present the theoretical model. In Section 3 we describe the dataset. In Section 4 we discuss the empirical strategy and introduce the methodology to estimate product-level productivity. In Section 5 we show and comment the main results. In Section 6 we present additional results with alternative methods. In Section 7 we discuss the robustness of our results and their policy implications. Section 8 concludes.

2 Theoretical framework

In this section, building on Hottman et al. (2016), we model product market share as a function of wholesale price, retail markup, firm scope and product appeal.[4]

We assume that utility at time t, U_t, is a Cobb-Douglas aggregate of physical quantities consumed, Q_{jt}, of a continuum of product markets:

$$U_t = \int_{j\in\Omega_j} \varphi_{jt}\, lnQ_{jt}\, dj \tag{1}$$

where j denotes each product market, φ_{jt} is the share of expenditure on product market j at time t, and Ω_j is the set of product markets.

Following Hottman et al. (2016), the aggregate physical quantities consumed across all products of market j, Q_{jt}, and across all products of firm f in market j, Q_{fjt}, are defined as follows:

$$Q_{jt} = \left[\sum_{f\in\Omega_{fj}} (\varphi_{fjt} Q_{fjt})^{\frac{\sigma_j - 1}{\sigma_j}}\right]^{\frac{\sigma_j}{\sigma_j - 1}} \quad , \quad Q_{fjt} = \left[\sum_{i\in\Omega_{ifj}} (\varphi_{it} Q_{it})^{\frac{\sigma_{fj} - 1}{\sigma_{fj}}}\right]^{\frac{\sigma_{fj}}{\sigma_{fj} - 1}}$$

where Q_{it} is the physical output of product i supplied by firm f in market j, Ω_{fj} is the set of firms operating in market j and Ω_{ifj} is the set of products of firm f operating in market j. The elasticity of substitution of the output supplied by the different firms within the market and the elasticity of substitution of the different product of a specific firm within the market are, respectively, σ_j and σ_{fj}. Product i's appeal is φ_{it} and φ_{fjt} is the average appeal of firm f's products within market j:

$$\varphi_{fjt} = \frac{1}{N_{fjt}} \sum_{i\in\Omega_{ifj}} \varphi_{it}$$

where N_{fjt} is the number of products of firm f for market j. The average appeal of firm f across all its products and markets is:

$$\varphi_{ft} = \frac{1}{N_{ft}} \sum_{j\in\Omega_j} \varphi_{fjt} \tag{2}$$

where N_{ft} is the number of markets in which firm f operates.

Using the properties of CES demand function, Hottman et al. (2016) show that the demand for the physical output of product i supplied by firm f within product

[4] Hottman et al. (2016) distinguish the contribution to firm revenues from firm cost, markups, appeal and firm's product scope.

market j, Q_{it} is:

$$Q_{it} = (\varphi_{fjt})^{\sigma_j - 1}(\varphi_{it})^{\sigma_{fj}-1} Y^r_{jt}(P^r_{jt})^{\sigma_j - 1}(P^r_{fjt})^{\sigma_{fj}-\sigma_j}(P^r_{it})^{-\sigma_{fj}} \tag{3}$$

where Y^r_{jt} is total consumer expenditure in market j, P^r_{jt} is an index of retail price in market j, P^r_{fjt} is an index of retail price of firm f's products within market j and P^r_{it} retail price of product i. Market and firm-market retail price indexes are, respectively:

$$P^r_{jt} = \left[\sum_{f \in \Omega_{fj}} \left(\frac{P^r_{fj}}{\varphi_{fjt}}\right)^{1-\sigma_j}\right]^{\frac{1}{1-\sigma_j}} \quad , \quad P^r_{fjt} = \left[\sum_{f \in \Omega_{ifj}} \left(\frac{P^r_{it}}{\varphi_{it}}\right)^{1-\sigma_{fj}}\right]^{\frac{1}{1-\sigma_{fj}}}$$

where P^r_{it} is retail price of product i

We use the model to define the different sources of market shares heterogeneity across products. Firm revenues from product i, Y^w_{it}, are given by:

$$Y^w_{it} = P^w_{it} Q_{it}$$

where P^w_{it} is the wholesale price of product i. Using Equation (3) firm revenues from product i can be rewritten as:

$$Y^w_{it} = P^w_{it}(\varphi_{fjt})^{\sigma_j - 1}(\varphi_{it})^{\sigma_{fj}-1} Y^r_{jt}(P^r_{jt})^{\sigma_j - 1}(P^r_{fjt})^{\sigma_{fj}-\sigma_j}(P^r_{it})^{-\sigma_{fj}}$$

Retail prices can be divided into two components: wholesale prices P^w and retail markups μ. We can, therefore, define the market share of product i as:

$$\frac{Y^w_{it}}{Y^w_{jt}} = \left(\frac{P^w_{jt}}{P^w_{fjt}}\right)^{\sigma_j}\left(\frac{P^w_{fjt}}{P^w_{it}}\right)^{\sigma_{fj}}\left(\frac{P^w_{it}}{P^w_{jt}}\right)\left(\frac{\mu^w_{jt}}{\mu^w_{fjt}}\right)^{\sigma_j}\left(\frac{\mu_{fjt}}{\mu_{it}}\right)^{\sigma_{fj}}(\varphi_{fjt})^{\sigma_j - 1}(\varphi_{it})^{\sigma_{fj}-1}$$

In narrowly defined markets, if we assume the elasticity of substitution of the output supplied by the different firms within the market is equal to the elasticity of substitution of the different products of a specific firm within the market are the same, namely $\sigma_j = \sigma_{fj} = \sigma$.[5] In such a case, multiplying and dividing by $(N_{ft})^{\sigma-1}$ we obtain:

$$\frac{Y^w_{it}}{Y^w_{jt}} = \left(\frac{P^w_{it}}{P^w_{jt}}\right)^{1-\sigma}\left(\frac{\mu_{it}}{\mu_{jt}}\right)^{-\sigma}(N_{ft})^{\sigma-1}\left(\frac{\varphi_{fjt}}{N_{ft}}\right)^{\sigma-1}(\varphi_{it})^{\sigma-1} \tag{4}$$

[5]The narrower the definition of market, the closer the substitutability of goods between firms and within firms in the market. This is indeed the case of our definition of market that includes drugs with the same chemical formulation and dosage form. Hottman et al. (2016) prescribe σ_j to be equal or higher than σ_{fj}. In case of equality the model replicates a standard CES system at the product level, but it has no implications for our purpose.

Revenue-based market share is therefore a function of the relative wholesale price of product i, $\frac{P^w_{it}}{P^w_{jt}}$, the relative retail markup of product i, $\frac{\mu it}{\mu jt}$, the scope (number of markets) of firm f, N_{ft}, product i's appeal, φ_{it}, and the contribution of market j to firm f's appeal, $\frac{\varphi_{fjt}}{N_{ft}}$, as from Equation (2).

3 Data

We use product-level data of all the pharmaceutical firms operating in India during April 2011 and March 2016. Since April marks the beginning of a new financial year in India, we have data for five financial years. Data are compiled by the All India Organization of Chemists and Druggists (AIOCD), the union of Indian pharmacies. For each *product*, identified by a unique stock keeping unit (SKU), we observe both wholesale and retail prices, as well as the number of units sold on India's pharmaceutical market, on a monthly basis. Since we employ these data at the product-year level, for every product we aggregate sales revenues and units across months, and define wholesale and retail prices as the weighted averages of the year.

AIOCD data contain nearly 3.1 million SKU-month observations spanning 917 firms, and around 95,000 different SKUs. For each product, we observe its *formulation* (the molecular composition of the active ingredient of the drug) and *dosage form* (e.g., injection, tablet, etc.), *dosage strength* (e.g., 10 mg, 100 ml, etc.) and *pack size* (number of tablets, syringes, etc.). We define a *market* as the combination of a formulation and dosage form, aggregated over different strengths and pack size. We identify more than 8,100 product markets for 2900 formulations. Products within a firm often share a common *brand*, as medicines with common formulation but differing dosage forms, strength, and pack size are assigned the same brand. In our data there are more than 51,000 brands.

We merge the AIOCD data with Prowess data on firm financials compiled by the Centre for Monitoring Indian Economy (CMIE).[6] The Prowess data contain annual financial information for publicly listed firms traded on the National and the Bombay Stock Exchanges in India. We identify the sample of firms in the category "Manufacture of pharmaceuticals, medicinal chemical and botanical products" (division 21) of the National Industry Classification (NIC) 2008. We manually match firm

[6] The CMIE Prowess data are used in other works regarding Indian multiproduct firms, such as Goldberg et al. (2010a); Topalova and Khandelwal (2011); Ahsan (2013); De Loecker et al. (2016) among the others. Another dataset that contains the same financial variables required to conduct our of analysis is the ASI dataset. The latter has the advantage of gathering also small and non-traded Indian firms. However, it does not include the name of the firm, which is the only identifier we have for the firms in the AIOCD dataset. Merging the two datasets is, therefore, not possible.

names between the Prowess and AIOCD data during 2011 and 2015. Among the 917 firms in the AIOCD dataset, we successfully merge 122 public firms accounting for over 130,000 product-year observations, spanning almost 43,000 SKUs. We present descriptive statistics for the merged dataset in Table 2. These merged firms represent 46 percent of all SKUs, their sales account for 60 percent of the total industry sales, and produce nearly 2,500 formulations. All the firms in the merged sample are multiproduct firms, manufacturing, on average, 315 different SKUs. The firms in the merged sample are also multiscope firms, operating, on average, in 164 markets of 70 medicines, often changing their product mix. In our data, the rate of entry of firms is 15 percent while the exit rate is 6 percent, consistent with Goldberg et al. (2010b) who show that entry is more common than exit in this industry, because high sunk costs make firms reluctant to withdraw production lines.

Within a market the products with the same formulation and dosage form can differ by dosage strength (e.g., 500 mg of Metformin vs 1000 mg of Metformin) and pack size (e.g., 10 tablets vs. 15 tablets). To facilitate comparison among prices and quantities of products, we normalize prices (P) and physical units of product (Q) by dosage strength (DS) and pack size (PS) using the formulae:

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

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

Note that multiplying the normalized price per unit (NP) by normalized quantity (NQ) gives product sales revenue. Since each market sells medicines with specific chemical components and therapeutic use, it is not feasible to compare prices and quantities across products between markets using absolute value measures. Instead, relative measures like market share (defined as the ratio between the sales of a product and the sum of all sales in the market) and relative price (defined as the ratio between normalized price of a product and the highest normalized price observed in the same market) are comparable. Using relative measures we observe significant dispersion in market shares and prices within and across the markets. In Table 2, Panel b, we notice that the distribution of market shares is heavily right-skewed, suggesting the presence of market leaders. The distribution of wholesale prices standardized at the market level and plotted in Figure 1 shows that price dispersion within the market is substantial.

As shown in Equation (4), market shares depend not only on prices but also, among other factors, on retail markups. To illustrate this, we show in Table 3 top ten markets by revenue in our data. Although many firms compete in these markets, competition is limited by the presence of products with dominant positions. Indeed, in each of these markets more than a hundred different products compete, but the revenue share of the market-leading product is more than ten percent in all markets except Atorvastatin. The dominant position of the market leaders is more evident when we compare their wholesale prices and retail markups and margins with those of the other products in the market. Market leaders charge a higher (normalized) wholesale price and lower *retail markup*, measured as the ratio between retail and wholesale price. Despite the lower retail markups, charging higher wholesale prices allows the leaders to provide higher *retail margin*, measured as the difference between retail and wholesale price.

4 Methodology

4.1 Empirical strategy

We estimate the relation between product market share and wholesale prices in the pharmaceutical industry, including the other determinants of market share pointed out by the model solution in Equation (4). The estimation equation is the following:

$$share_{ifjt} = \alpha_0 + \alpha_1 price_{ifjt} + \alpha_2 markup_{ifjt} + \alpha_3 scope_{ft} + \alpha_4 age_{ifjt} + \alpha_5 X_{fjt} + \epsilon_{ifjt} \tag{5}$$

where $share_{ifjt}$ is the share of revenues of product i of firm f in market j. The variable $price_{ifjt}$ is product i's relative wholesale price, defined as the ratio between normalized price of a product and the highest normalized price observed in the same market. Retail markup is captured by $markup_{ifjt}$, the ratio between retail and wholesale price of product i. The variable $scope_{ft}$ represents firm scope, measured as the log of the number of markets in which firm f operates. Brand appeal is captured by the log of the number of years since the product brand was launched, age_{ifjt}. The vector X_{fjt} controls for market competition, using the log of the number of competing products in the market, and other demand shifters including market, year and firm fixed effects. The error term, ϵ_{ifjt}, is product-year specific.

Estimating (5) using OLS introduces an upward bias in α_1, as an idiosyncratic shock in demand might stimulate an increase in price by the manufacturers. To identify price elasticity in a standard demand function for single-product firms, Foster et al. (2008) instrument for prices with a product-level measure of productivity, a supply-side driver of prices containing information on a firm's cost of production. Indeed, productivity is a measure of technical efficiency, directly comparable across products within and across markets. Moreover, it is unlikely to be correlated with idiosyncratic product-specific demand shocks in the short run. We adopt the same methodology and instrument for relative wholesale price with a measure of quantity-based productivity. Unlike Foster et al. (2008), our dependent variable is market share and our firms produce multiple products.

The use of productivity as an instrument can be questioned for two reasons (Foster et al., 2008). First, if higher productivity leads to higher probability of survival when drawing a negative demand shock. The first issue is mitigated by the fact that our data span for five years, and a bad draw in such a short period can be amortized by cross subsidization of other products of the firm. Second, measurement error in estimating productivity can undermine the usefulness of the instrument. Such measurement error can arise if quantities are not directly observed but instead calculated by dividing sales with prices. Since we observe quantities separately from prices and sales, measurement is not a concern in our study.

4.2 Product-level productivity in multiproduct firms

Product-level productivity, intended as the product-specific efficiency with which a firm turn inputs into outputs, serves as an instrument for addressing endogeneity in estimating the impact of wholesale prices on market shares. We estimate product-level productivity controlling for the "product scope bias" that arises when estimating the production function in multiproduct firms: firm productivity is correlated with product switching, suggesting that firms endogenously select the goods to produce Bernard et al. (2010). De Loecker and Goldberg (2014) show that, given the wide heterogeneity in prices at the product level, the estimation of productivity turns out to be non-negligibly biased. De Loecker et al. (2016) control for the product mix when estimating the productivity for India's manufacturing firms.[7] Dhyne et al. (2017), propose a new approach to estimate productivity at the product level, which accounts for the firm product scope. Other works, consider the production function of a multiproduct firm as the sum of single product production functions. By contrast, Dhyne et al. (2017) estimate a multiproduct production function.

Dhyne et al. (2017) calculate product-level productivity as the residual of a production function, whose output elasticities are estimated using firm-level (and not product-level) inputs. They consider a quantity-based loglinear Cobb-Douglas pro-

[7] 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.

duction function, specified as follows:

$$q_{it} = \omega_{it} + \alpha k_{ft} + \beta_l l_{ft} + \beta_m m_{ft} + \gamma y_{-it} \tag{6}$$

where, for each product i, firm f and year t, q is log quantity sold in physical units, k is log capital employed, l is log salaries, m is log raw materials and y_{-i} is log revenues of all other products except from i produced by the firm. Adding this latter measure to the production inputs extends the single product setting by estimating 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" (Dhyne et al., 2017). Product-specific log productivity (ω) is Hicks-neutral and can be computed as a Solow residual.

Building on Dhyne et al. (2017) we estimate the output elasticities of 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, indeed, both raw materials and salaries can be considered as product-specific. Given that the chemical composition of each drug is fixed, a marginal increase in real raw materials expenditure for product i affects the output of product i only, and not also the output of other products of the firm. The same can be assumed for salaries. Given the highly automated production process of the pharmaceutical industry, a marginal increase in real salaries of the workers producing product i affects the output of product i only, and not also the output of other products of the firm. An increase in real capital expenditure, instead, being related to machinery, software, or plant infrastructure, is more likely to affect more than one product of the firm, and can enter the production function at the firm level, as in Dhyne et al. (2017). We propose the following production function, in which variable inputs enter at the product level and capital enters at the firm level:

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

To estimate Equation (7) we need product-specific raw materials and salaries, which we do not observe. This issue is tackled by assuming that in the pharmaceutical industry, variable inputs within a market have the same quality across products. Consequently, we assume that the unit cost of a variable input is the same across all the products of a market. Exploiting this and other commonly employed assumptions for the purpose, we apportion the amount of firm-level variable inputs into firm-product-level inputs. We provide details related to input allocation in Appendix A.

To address the simultaneity bias, we adopt the estimator proposed by Levinsohn and Petrin (2003) (henceforth, LP) using materials as a proxy.[8] We estimate the output elasticities at the formulation level and obtain a quantity-based measure of product-level productivity for multiproduct firms (*TFP-QEM*). In Figure 2 we show the distribution of product-level productivity and the central moments of the distributions of the output elasticities. On average the output elasticity with respect to capital is 0.57, with respect to labor is 0.20, with respect to materials is 0.52. The coefficient γ is negative on average, -0.06, as expected, since an increase in firm revenues, holding product variable inputs and firm-level capital constant, would result in a decrease in the quantity of the focal product.

5 Results

5.1 The determinants of market share

The OLS estimates of Equation (5), reported in Table 4, confirm the positive relation between revenue-based market share and prices, that remains robust to the inclusion of the other determinants of market share (Column 1). The results remain similar when we restrict the sample to more competitive markets with five or more products (Column 2) and those with ten or more products (Column 3). Retail markup is negatively correlated with market share in columns 1-3, consistent with our expectation. The coefficient estimate of firm scope is not significant, as the inclusion on firm-specific fixed effects likely absorbs much of the variation. Brand age, a proxy for brand appeal, is positively and significantly related to market share.

Although the 122 pharmaceutical firms in the merged sample account for 60 percent of the total sales in the industry, there are 795 firms in the original AIOCD dataset that are not observed in the Prowess dataset. Since the Prowess sample consists of publicly traded firms, the non-merged firms are on average smaller in size and scope. Moreover, there are markets which are not covered by the Prowess firms, potentially implying that our results are not representative for the informal and fringe firms that populate India's pharmaceutical industry. However, when we compare the OLS estimates on the merged sample in Table 4 with those obtained from the full AIOCD sample in Table B.1 in the Appendix, we notice that the results are very similar.

We expect the OLS estimates to be biased. Following Foster et al. (2008), we instrument for relative wholesale price with the measure of quantity-based productivity

[8] Since the introduction of y_{-it} causes problems of endogeneity, we include its lagged value among the conditioning variables of the GMM estimation in the second stage of LP procedure, as suggested by Dhyne et al. (2017). Find the adjustment operated to the LP estimator in Appendix A.

that addresses the product scope bias (*TFP-QEM*). Instrumenting for wholesale prices with quantity-based productivity reverses the sign of its effect on market share (Column 4). This result suggests that the initially-observed positive relationship between prices and revenue-based market shares is not due to distortions on the supply side, as higher technical efficiency drives prices downwards. The reported F-statistics confirms that the instrument chosen is strongly relevant. The other explanatory variables retain their sign as in the OLS estimates. The results are similar in markets including more than five and more than ten products (Column 5 and 6).

5.2 Market power and buyer power

The results reported in Table 4 constitute an average effect in the industry of a change in product price on market share. However, this effect can be different for products with different levels of market power (Frank and Salkever, 1997; Berndt and Aitken, 2011; Reiffen and Ward, 2005). Next, we study how prices and market shares are related differently for market-leading products and the others.

We identify the market-leading product as the product with the highest share of sales within each market. We restrict the sample to markets with five or more products competing and distinguish between market leaders and followers using the dummy variables *Leader*, which takes value one when the product is market leader and zero otherwise. We interact the dummy variable with the wholesale price.

Wholesale price, however, is not the only determinant that can present asymmetric effects for leaders and followers. Higher retail markups or margins can incentivize the retailer to provide a larger availability of a certain product. In such a case, higher retail markups (margins) can boost the product market share, despite they imply also higher retail prices. If the decrease in sales due to higher retail prices is negligible, then increasing retail markups (margins) is a profitable strategy for the firm. We consider whether products with higher levels of retail markups (margins) affect market shares positively and if there is any asymmetry in this relationship across products with different market power. To test this we interact retail markups and margins with the market-leader dummy defined above.

The introduction of retail markups and margins into the equation potentially introduces a source of bias, as it can respond to a product-specific demand increase, likewise the wholesale prices. To identify the effect of retail markups and margins we instrument them with firm retail markup across all its products. This variable is likely to be a relevant driver of product retail markup and margin, as the firm might strategically set an average retail markup across all its products, before deciding product-specific markup or margin. This instrument is also supposed to be uncorrelated with idiosyncratic product-specific demand shocks in the short run, because of the multiproduct nature of the firm.

Table 5 shows that the coefficient estimates for prices of leaders and followers have opposite signs in both OLS (Column 1) and IV estimates (Column 2). An increase in price of the leading product affects market share positively, but for a follower, the same increase in price leads to a decrease in market share. The IV estimates strengthen the magnitude of the coefficients without changing the significance. These results are in line with the findings of Berndt and Aitken (2011) indicating that market leaders are shielded from price competition in the Indian pharmaceutical industry. Column 3-6 show the results for retail markups and margin. The coefficient estimate of the interaction between retail markups and market-leader is significant and positive in both OLS and IV estimates (Columns 3 and 4). The same occurs to retail margins (Columns 5 and 6). In theory, an increase in retail markup (margin) should diminish the demand for the product as its retail price, ceteris paribus, is higher. However, higher markups (margin) incentivize the retailer to stock the product and allocate greater selling effort.

The results in Table 5 provide evidence of how market power drives product demand in narrowly defined markets. If the product has some market power, a price increase affects market share positively. However, the results also suggest that pharmaceutical firms can benefit from retailer buyer power. If the retailer buyer power is high, a higher retail price can increase the demand for the leading products. On the consumer side, both market power and buyer power contribute to a welfare loss. The manufacturers of best selling drugs appear not to lower their prices and the retailers are more willing to provide products that yield them a higher margin. Moreover, consumer-specific factors such as brand appeal is a relevant determinant of market shares in all our estimates. Demand-side factors involving consumers and retailers, and the presence of asymmetries for market leaders seem to be primarily responsible for the positive relation between prices and revenue-based market shares observed.

6 Alternative methods and additional results

6.1 The relationship between prices and productivity

6.1.1 Alternative measures of productivity

In this section, we examine the robustness of our results to alternative measures of productivity. To estimate productivity at the product level, prior literature usually considers a log-additive production function (e.g. Cobb-Douglas) whose coefficients remain constant over the sample period:

$$x_{it} = \omega_{it} + \alpha k_{it} + \boldsymbol{\beta} \mathbf{v}_{it} \tag{8}$$

where, for each product i and year t, x is log output, k is log capital and v is a vector of variable inputs in logs. Product-specific log productivity (ω) is Hicks-neutral. The production functions is either *revenue-based*, if output is measured in sales revenues y, or *quantity-based*, if output is measured in quantity of physical units sold q.

Estimating productivity 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; and b) we do not observe productivity (neither at the firm nor at the product level). We discuss our approach to addressing (a) in Appendix A. Concerns related to (b) can lead to two potential biases: i) a *product scope bias*, as the number of products is decided by the firm according to the observed productivity; and ii) a *simultaneity bias*, as the amount of inputs is chosen based on firm or product productivity. We previously discussed the product scope bias and the measure of productivity we proposed in Section 4.2 addresses it.

The simultaneity bias concerns the computation of output elasticities, α and $\boldsymbol{\beta}$. This can be addressed in two different ways: a) equalling elasticities to average input cost share over the sample (cost share-based method); or b) estimating the elasticities econometrically (estimation-based method). The first method follows the theoretical framework of cost minimization of the firm and 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. The estimation-based method, instead, addresses the simultaneity bias, which arises as input quantities are chosen according to observed or expected (by the firm) productivity (Olley and Pakes, 1996). We use the LP estimator for our estimation-based method.

The productivity measure used for obtaining the results in the previous section, *TFP-QEM*, addresses all the biases that might occur when estimating productivity. Other measures of productivity could be adopted, although they fail to address at least one of the aforementioned biases. We compute five additional measures of product-level productivity, and compare them to our preferred measure in Table 6, distinguishing between revenue- or quantity-based and cost share- or estimation based. All input elasticities are calculated at the formulation level and their industry-level average is reported. The two cost-share-based measures of productivity *TFP-RC* and *TFP-QC* are computed using the same equation (same input elasticities), but they differ in terms of the output variable: revenues for *TFP-RC* and physical units for *TFP-QC*, as in Foster et al. (2008). The two revenue-estimation-based measures of productivity differ by either including raw materials in the output (value added-based), *TFP-VE*, as in Ahsan (2013) or in the inputs, *TFP-RE*, as in Topalova and Khandelwal (2011). The quantity-based version of *TFP-RE* is *TFP-QES*, suitable for single-product firms, as in De Loecker et al. (2016).

6.1.2 Product prices and selection on productivity

In Section 5.1 we used *TFP-QEM* to instrument for relative prices. We repeat the procedure using all six measures of product-level productivity to instrument for relative price in Equation (5). The results are shown in Table 7. As before, we include retail markups, firm scope, brand appeal, competition, and a battery of fixed effects in these regressions. We find that relative prices are positively correlated with the three revenue-based measures, but negatively correlated with the three quantity-based ones. Foster et al. (2008) find a similar pattern of results. They note that the results are driven by the fact that revenue-based productivity incorporates prices by definition although revenue-based productivity is calculated with deflated values. This suggests that a revenue-based measure is not a suitable instrument for prices to identify elasticity in Equation (5). Quantity-based productivity, instead, is more appropriate to satisfy the exogeneity condition.

The negative relation between quantity-based productivity and prices is not surprising: more efficient products have lower marginal costs, allowing firms to charge lower prices for them. Indeed, 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).[9] Also empirical studies of specific industries 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 homogeneous product, competitive selection on costs lowers product prices. Our analysis on multiproduct firms shows that higher productivity is associated with lower prices even at the product level.

6.2 Quantity-based market share and price elasticity

In Table 1 and Table 5 we observe that revenue-based market share is positively correlated with wholesale prices when estimated using OLS. Since revenues include prices by definition, there is a mechanical relation between the two variables. In Table (8) we address this issue by estimating the relationship between*quantity-based* market share and relative prices. Quantity-based market share is calculated as the product share of normalized quantity sold in the market. Within the market, the products are homogeneous in chemical composition, and therefore we can sum their normalized quantities sold and obtain a measure of quantity-based market share comparable between different markets. Quantity-based market share is negatively correlated with wholesale prices when estimated with OLS. IV estimates confirm the direction of the relation, but the size of the price coefficient increases in absolute value, as expected. A comparison with the estimates reported in Table 5, whose dependent variable is revenue-based market share, shows that the IV coefficients are very close. These findings suggest that the positive correlation between revenue-based market share and prices does not, on average, imply a higher demand (quantity) for the product. As consumers might perceive higher prices as higher product quality, this result implies that perceived quality is not the main driver of the upward relationship between revenue-based market share and prices.

In Table 9 we show the results of the analysis on the asymmetric effects of market leaders and followers, using quantity-based market share as the dependent variable. Table 9 confirms the findings presented in Table 6: a price increase of the leading product not only increases its revenue-based market share, but also its quantity-based market share. This is consistent with Bronnenberg et al. (2015), suggesting that the higher prices of the leaders can be perceived as higher quality by the consumers, boosting the product demand.

[9] 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).

Demand elasticity is, therefore, another important component of the puzzle. We investigate price elasticity of demand using the same controls as in Equation (5) and test for asymmetric effects across market leaders and others. The dependent variable is the log of normalized units of product sold and the main explanatory variable is log normalized wholesale prices. Table 10 presents the estimates of price elasticity. The OLS estimates are negative and the elasticity is less than one (Column 1 and 2). However, with market share as the dependent variable OLS estimation can be upward biased. Therefore, we instrument normalized prices with quantity-based productivity and obtain a higher price elasticity.[10] The IV results show that a one-percent increase in prices leads to a 5.6 percent decrease in drug units sold (Column 3). The results remain similar in a subsample of markets with 5 or more competitors (Column 4-6).

Next we allow price elasticity to differ for market leaders and followers. The results, shown in Table 11, indicate that price elasticity is always significantly higher for market followers in both OLS and IV estimates. Similarly, we also allow for differential effects of retailer margin. The marginal effect of an increase in retail margin (i.e., retail margin elasticity of demand) has opposite signs, depending on whether the product is market leader or follower. A one percent increase in the level of retail margin decreases demand by 0.8 percent for the market followers and increases it by 0.4 percent for the market leaders.

7 Robustness analysis and policy implications

In this section, we examine the robustness of our results. Our objective is to investigate whether the pattern of results we observed previously are due to the idiosyncrasies of the Indian pharmaceutical industry or the unusual manner in which health care is organized and consumed in India. We divide the observations in our sample across firm, market and product characteristics and estimate Equation (5) using the IV method described in Section 4.1.

We begin with robustness checks in subsamples of our data based on firm and market characteristics. We examine whether the results are driven by a perceived quality effect of multinational firms. That is, if more expensive medicine brands also have the largest market share, is such an effect driven by the multinational firms in our dataset, whose brands are perceived to be of a higher quality? If so, then our findings are explained by quality differences rather than the incentives of the intermediaries such as pharmacists. We examine this by distinguishing domestic and multinational firms in our dataset and conducting a split-sample analysis of the rela-

[10] First-stage regression estimates reported in Table B.2 in Appendix.

tionship between relative wholesale price and market share in Table 12. We find the coefficient estimate of relative wholesale price in columns 1 and 2 to be similar, indicating that a higher price is associated with a lower market share, after instrumenting for prices with product-level productivity for both domestic and multinational firms.

Next, we examine, in columns 3-5 of Table 12 whether the results are specific to some dosage forms than others. Unlike medicines in solid form (such as tablets), injectibles are often consumed with the help of a healthcare professional. It is likely that the relationship between prices and market shares differ from medicines in their solid, liquid, and injectable forms. In Table 12 we examine these effects. Overall, the results remain qualitatively similar but the effect size is smaller for injectables. The results show that an increase in relative price lowers market share of injectables by half as much as for solids or liquids.

We also examine the population of medicines in India, which are composed of medicines from different launch periods. It is possible that our results are driven by recently launched medicines that tend to be more expensive. In particular, India agreed to the WTO TRIPS agreement in 1995 and implemented it since 2005. In Table 12, we distinguish between medicines launched before 1995, between 1995 and 2004, and since 2005 and estimate the relationship between relative wholesale price and market share. We find broadly similar results in columns 6-8 in Table 12, indicating that the results are not driven by the vintage of the medicine.

We further examine the robustness of our results to several subsample analyses based on medicine characteristics. The results are reported in Table 13, the differential effects for acute and chronic medicines. Medicines for chronic diseases such as diabetes are consumed over extended periods of time and they are associated with repeat purchases. By contrast, medicines for acute conditions such as flu or fever are characterized by occasional use. The frequency of purchase of medicines can influence the relationship between relative wholesale price and market share. We find that the coefficient estimate of price for Chronic subsample (column 2) is significantly larger than for Acute (column 1), indicating that the markets for Chronic medicines are relatively more competitive.

Nearly half the medicines sold in India are fixed dose combination medicines or medicines composed of multiple ingredients. We examine, in column 3 and 4 of Table 13, whether there are differential effects for single-ingredient and combination medicines. The results indicate that combination medicine markets are relatively less competitive (see column 4). Nevertheless, the pattern of results we obtained previously remain qualitatively similar for single -ingredient and combination medicines.

India adopted a price ceiling regulation in mid-2013 for essential medicines, covering nearly a fifth of the market for pharmaceuticals in India. We estimate the relationship between relative wholesale price and market share for unregulated and regulated medicines in columns 5 and 6 of Table 13. We find that the regulated medicine markets are relatively more competitive but the relationship is qualitatively similar for both regulated and unregulated markets. We also distinguish between over-the-counter and prescription medicines and find broadly similar results (see columns 7 and 8).

8 Conclusion

We examine the relationship between prices and market shares in multiproduct firms. We exploit a rich dataset on Indian pharmaceutical firms containing detailed information on quantities, wholesale and retail price of every drug sold in the country. Using a naive OLS estimation that does not control for simultaneity and omitted variable bias, we find that market share is positively correlated with wholesale prices. In this paper we address the aforementioned biases and identify the relationship between wholesale prices and market shares.

Building on Hottman et al. (2016) model of heterogeneous multiproduct firms, we distinguish the principal drivers of product market share into product wholesale price, retail markup, product and firm appeal and firm scope. We control for these confounding factors and estimate the effect of a change in relative wholesale price on product market share. We instrument for wholesale prices with a measure of quantity-based, product-level productivity, a supply side driver of prices, uncorrelated with product-specific demand shocks. We find that a one percent change in wholesale price affects market share by almost the same amount, but with the opposed sign. This finding indicates that, on average, higher prices do not lead to higher market share.

Analysing market leaders and market followers separately, however, reveals the significant market power the leaders possess in the markets we study. A price increase of the market-leading product affects its market share positively, indicating that leaders are insulated from price competition in the market. One mechanism that helps leader's insulation from competition seems to be the incentive provided to the retailers in the form of higher markups or margins. These results reveal how retailer's buyer power can foster a win-win relationship between retailers and market leaders. In our findings the relatively higher demand for market leaders depends on both retailer incentives and perceived quality. Since we consider markets with close substitute goods, our findings imply a welfare loss for the less informed consumers. However, in our data we cannot distinguish if the observed results for the market leaders is the outcome of a retailer-driven strategy to reduce local availability of competing products. In such a case, even the well-informed consumers would have little opportunity to switch from the market-leading product to a cheaper alternative. Nevertheless, our results also show that the selection mechanism in the Indian pharmaceutical market does not reward the less productive products. Quantity-based productivity is negatively correlated with product wholesale price, implying that productivity triggers price competition, on average. Indeed, this consideration excludes the market leaders whose price setting decision leverages their interaction with the intermediaries and the consumers.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

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.

Asker, J. and Bar-Isaac, H. (2014). Raising retailers' profits: On vertical practices and the exclusion of rivals. *The American Economic Review*, 104(2):672—686.

Bate, R., Jin, G. Z., and Mathur, A. (2011). Does price reveal poor-quality drugs? evidence from 17 countries. *Journal of Health Economics*, 30(6):1150–1163.

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. (2010). Multiple-product firms and product switching. *The American Economic Review*, 100(1):70.

Berndt, E. R. (2002). Pharmaceuticals in us health care: determinants of quantity and price. *Journal of Economic perspectives*, 16(4):45–66.

Berndt, E. R. and Aitken, M. L. (2011). Brand loyalty, generic entry and price competition in pharmaceuticals in the quarter century after the 1984 waxman-hatch legislation. *International Journal of the Economics of Business*, 18(2):177–201.

Berndt, E. R. and Cockburn, I. M. (2014). The hidden cost of low prices: limited access to new drugs in india. *Health Affairs*, 33(9):1567–1575.

Berry, S. T., Gaynor, M., and Morton, F. S. (2019). Do increasing markups matter? lessons from empirical industrial organization. Technical report, National Bureau of Economic Research.

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

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.

Chaudhuri, S., Goldberg, P. K., and Gia, P. (2006). Estimating the effects of global patent protection in pharmaceuticals: a case study of quinolones in india. *American Economic Review*, 96(5):1477–1514.

Chernew, M., Cooper, Z., Larsen-Hallock, E., and Morton, F. S. (2018). Are health care services shoppable? evidence from the consumption of lower-limb mri scans. Technical report, National Bureau of Economic Research.

Craig, S. V., Ericson, K. M., and Starc, A. (2018). How important is price variation between health insurers? Technical report, National Bureau of Economic Research.

De Loecker, J. and Eeckhout, J. (2018). Global market power. 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. (2017). Multi product firms, import competition, and the evolution of firm-product technical efficiencies. *National Bureau of Economic Research*.

Duggan, M., Garthwaite, C., and Goyal, A. (2016). The market impacts of pharmaceutical product patents in developing countries: Evidence from india. *American Economic Review*, 106(1):99–135.

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.

Frank, R. G. and Salkever, D. S. (1997). Generic entry and the pricing of pharmaceuticals. *Journal of Economics & Management Strategy*, 6(1):75–90.

Gaynor, M., Ho, K., and Town, R. J. (2015). The industrial organization of health-care markets. *Journal of Economic Literature*, 53(2):235–84.

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.

Grennan, M., Myers, K., Swanson, A., and Chatterji, A. (2018). Physician-industry interactions: Persuasion and welfare. Technical report, National Bureau of Economic Research.

Hastings, J., Hortaçsu, A., and Syverson, C. (2017). Sales force and competition in financial product markets: the case of mexico's social security privatization. *Econometrica*, 85(6):1723–1761.

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.

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.

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

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.

Lakdawalla, D. N. (2018). Economics of the pharmaceutical industry. *Journal of Economic Literature*, 56(2):397–449.

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

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.

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.

Reiffen, D. and Ward, M. R. (2005). Generic drug industry dynamics. *Review of Economics and statistics*, 87(1):37–49.

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.

Starc, A. and Swanson, A. (2018). Preferred pharmacy networks and drug costs. Technical report, National Bureau of Economic Research.

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.

Tables

Table 1: Market share and wholesale prices
Naive specification: OLS estimates

Dependent variable	Market Share					
Column	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	All	N5	N5	N10	N10
Relative Wholesale Price	0.338***	0.029***	0.070***	0.006***	0.038***	0.007***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Market FE	No	Yes	No	Yes	No	Yes
Firm FE	No	Yes	No	Yes	No	Yes
Year FE	No	Yes	No	Yes	No	Yes
R-squared	0.233	0.678	0.035	0.233	0.017	0.173
Observations	329280	328746	289177	289164	257112	257099

Source: AIOCD.
Notes: The dependent variable is *Revenue-based Market Share*, calculated as the share of sales of the product in the market using wholesale prices. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. A *Market* includes all the drugs having the same formulation and dosage form (tablet, injection, syrup, etc.). Sample *All* includes all the markets, *N5* includes the markets with five or more products, *N10* includes the markets with ten or more products.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Summary statistics

Panel a - Observations

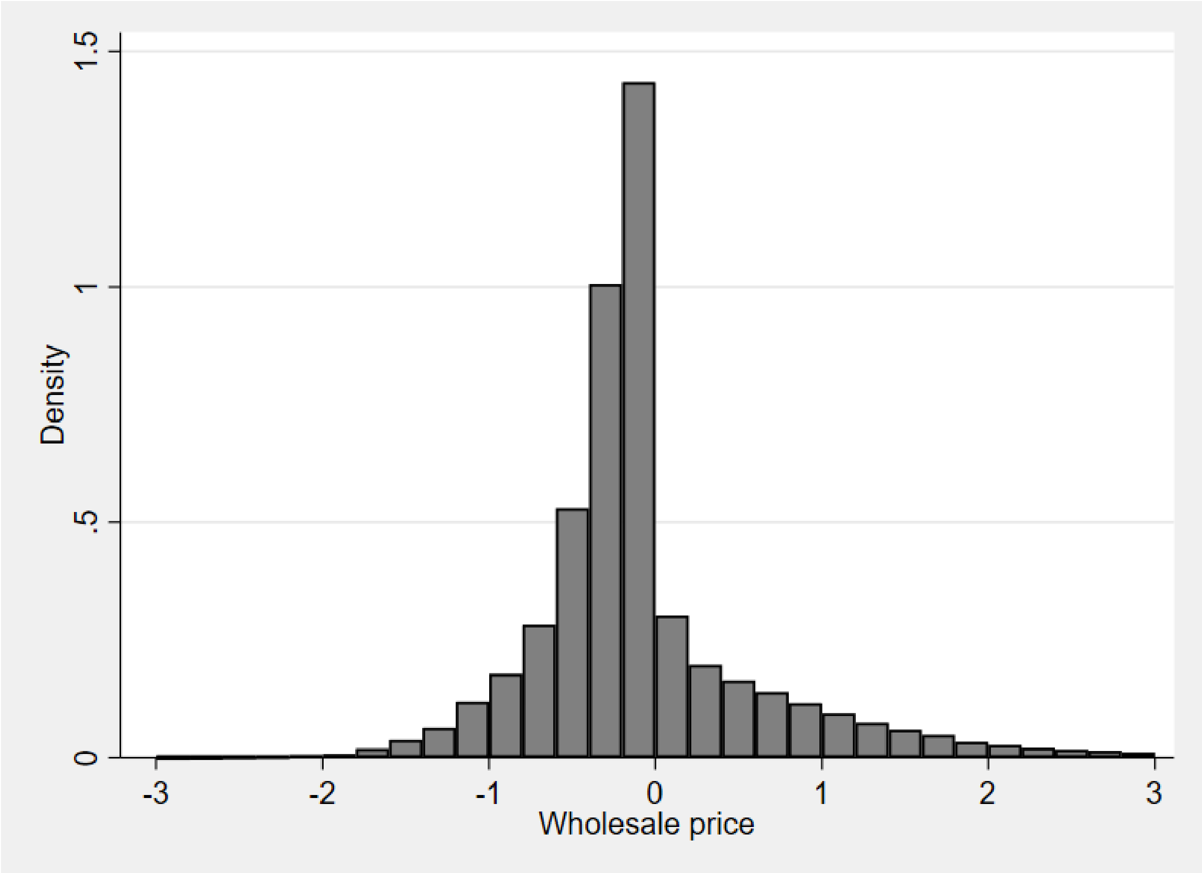
Firms (N)	122
Formulations (N)	2,323
Markets (N)	5,060
Brands (N)	20,796
Products (N)	40,784
Product-year Observations (N)	131,770
Firms changing product mix (%)	93.7
Firms changing market mix (%)	87.8

Panel b - Distributions

	mean	median	1 percentile	99 percentile
Products per firm (N)	316	150	3	2325
Formulations per firm (N)	70	122	2	691
Markets per firm (N)	166	90	2	1057
Sales per product (INR)	17.4 mn	1.8 mn	280	249 mn
Sales per firm (INR)	5.3 bn	1.1 bn	40.0 th	39.3 bn
Revenue-based Market Share (%)	11.4	0.7	5e-07	100
Quantity-based Market Share (%)	11.3	0.6	2e-07	100
Wholesale Price per Unit (INR)	265	51	3	5.0 th
Relative Wholesale Price (%)	31.6	15.1	0.01	100
Retail markup (%)	152.3	128.8	113.3	584.0
Retail margin (INR)	91	16	1	1.7

Source: AIOCD and Prowess, CMIE.
Notes: The dataset is obtained merging AIOCD and CMIE datasets. *Formulation* refers to the active pharmaceutical ingredient, *Market* refers to the combination between formulation and dosage form (tablet, injection, syrup, etc.), *Brand* refers to the retail name of the drug (irrespective of the strength or dosage), *product* refers to the SKU. *Product mix* changes when a firm starts producing or stops producing a stock keeping unit, *Market mix* changes when a firm enters a market with a new drug or exits a market ceasing producing drugs for that market. *Revenue-based market share* is the ratio between product sales and total market sales. *Quantity-based market share* is the ratio between normalized product units and the total normalized units sold on the market. *Wholesale Price per Unit* is wholesale price for one unit (pack) of product. *Relative Wholesale Price* is the ratio between normalized product wholesale price and the highest normalized product wholesale price in the market. *Retail Markup* is the ratio between retail and wholesale price of the product. *Retail Margin* is the difference between retail and wholesale price of the product. *Sales* and *Wholesale Price per Unit* are expressed in Indian Rupees (INR) where *bn* is billion, *mn* is million, *th* is thousand.

Figure 1: Distribution of wholesale prices within the market



Source: AIOCD dataset

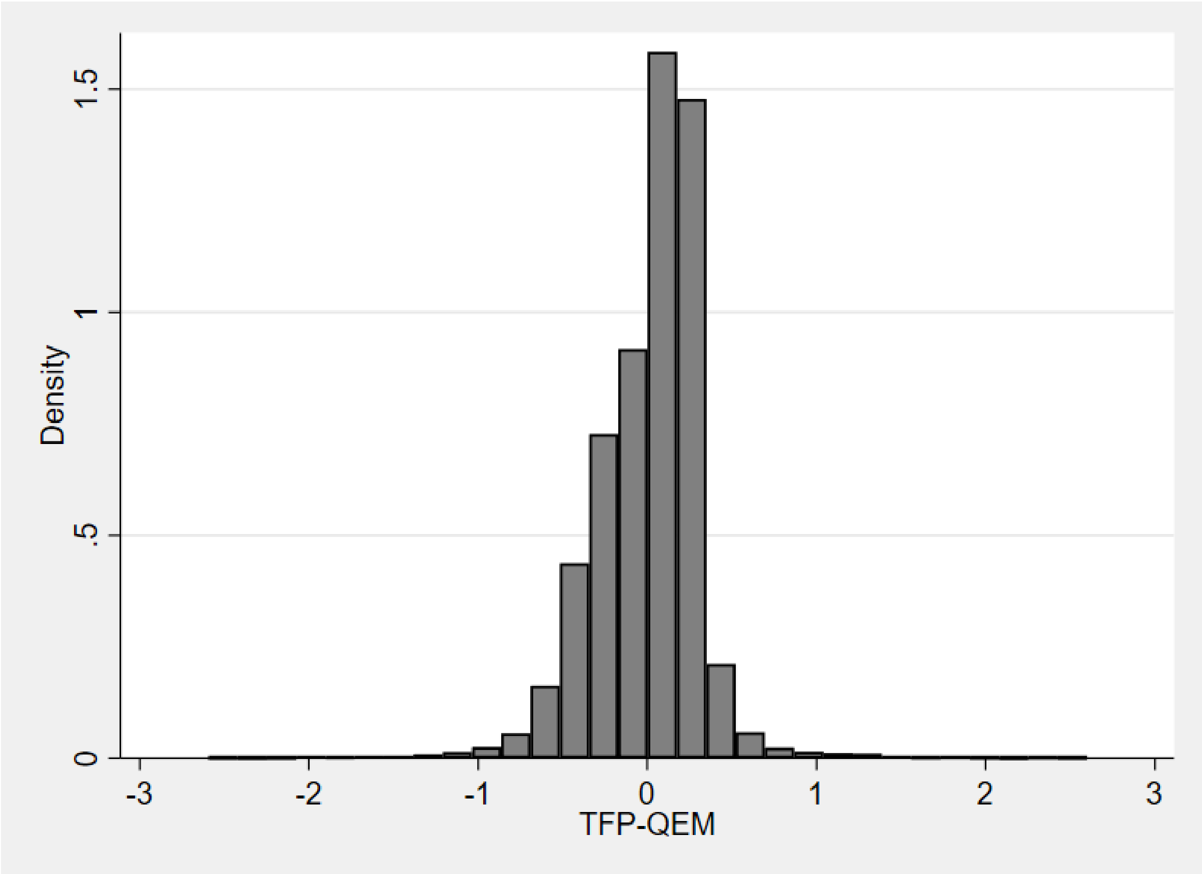
Note: Distribution of normalized product wholesale prices within the market, standardized at the market level (the average price of every market takes value zero).

Table 3: Top 10 largest markets. Leaders versus followers

Market	Competitors within market	Leader's market share	Leader's Wholesale Price / Follower's Wholesale Price	Leader's Retail Markup / Follower's Retail Markup	Leader's Retail Margin / Follower's Retail Margin
Amoxycillin & Clavulanic Acid (Tablet)	411	13.69	0.96	0.95	1.15
Atorvastatin (Tablet)	332	5.17	1.85	0.81	0.92
Azithromycin (Syrup)	503	18.12	1.23	0.95	1.05
Cefixime (Tablet)	336	21.53	1.03	0.87	1.07
Cefuroxime (Tablet)	246	25.90	1.54	0.89	1.30
Chlorpheniramine & Codeine (Tablet)	101	35.82	1.23	0.94	1.25
Glimepiride & Metformin (Tablet)	274	12.83	1.17	1.01	1.19
Pantoprazole (Tablet)	249	27.88	1.33	0.85	1.22
Rosuvastatin (Tablet)	196	10.44	1.37	1.00	1.49
Telmisartan (Tablet)	224	18.56	1.24	0.98	1.62

Source: AIOCD.
Notes: The 10 largest markets in terms of sales are listed in alphabetical order. *Market* refers to the combination between formulation and dosage form. *Competitors within markets* is the number of different products competing in the market. *Leader's market share* is revenue-based market share of the product with the highest level of sales within the market. *Leader's Wholesale Price / Follower's Wholesale Price* is the ratio between the normalized wholesale price of the market leader and the average normalized price of all the market followers. *Leader's Retail Markup / Follower's Retail Markup* is the ratio between the retail markup of the market leader and the average retail markup of all the market followers, where *Retail Markup* is the ratio between retail and wholesale price of the product. *Leader's Retail Margin / Follower's Retail Margin* is the ratio between the retail margin of the market leader and the average retail margin of all the market followers, where *Retail Margin* is the difference between retail and wholesale price of the product.

Figure 2: Product-level productivity estimates
Using a quantity-based production function for multiproduct firms (TFP-QEM)



	α	β_l	β_m	γ
Mean	0.571	0.205	0.524	-0.059
	(0.417)	(0.357)	(0.336)	(0.399)
Median	0.600	0.163	0.550	-0.016

Source: AIOCD and Prowess, CMIE.
Notes: Figure plots the standardized distribution of the product-level productivity estimated as from Equation (7). The production function is quantity-based and estimated using the LP estimator, adjusted to control for firm's product scope. Output elasticities are estimated at the formulation level for 2,323 formulations. Table reports means, standard deviations (not standard errors, in brackets) and medians of the formulation-level output elasticities estimated. Column α reports the output elasticity to capital, Column β_l reports the output elasticity to labor, Column β_m reports the output elasticity to materials, Column γ reports the output elasticity to firm's product scope.

Table 4: The determinants of market shares
Baseline specification: OLS and IV estimates

Dependent variable	Market share					
Estimator	OLS			IV		
Column	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	N5	N10	All	N5	N10
Relative Wholesale Price	0.013***	0.011***	0.014***	-1.074***	-0.880***	-0.701***
	(0.004)	(0.003)	(0.003)	(0.050)	(0.039)	(0.035)
Retail Markup	-0.001***	-0.001***	-0.001***	-0.003***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Firm Scope (log)	-0.004	-0.003	-0.002	-0.005	-0.002	0.003
	(0.004)	(0.003)	(0.002)	(0.010)	(0.008)	(0.006)
Brand Age (log)	0.033***	0.029***	0.024***	0.017***	0.017***	0.016***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Rivals Brands (log)	-0.149***	-0.050***	-0.032***	-0.361***	-0.187***	-0.130***
	(0.004)	(0.002)	(0.002)	(0.013)	(0.008)	(0.007)
Observations	131770	115505	100867	131770	115505	100867
R-squared	0.687	0.306	0.226			
F-stat				798.5	1700.4	1717.4

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Revenue-based Market Share*, calculated as the share of sales of the product in the market using wholesale prices. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. A *Market* includes all the drugs having the same formulation and dosage form (tablet, injection, syrup, etc.). Time, firm and market fixed effects are included. Sample *All* includes all the markets, *N5* includes the markets with five or more products, *N10* includes the markets with ten or more products. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*). *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: The determinants of market shares
Market leaders and market followers: OLS and IV estimates

Dependent variable	Market Share					
Sample	N5					
Column	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	IV	OLS	IV	OLS	IV
Relative Wholesale Price	-0.037***	-0.490***	0.008***	-0.436***	0.004**	-0.483***
	(0.002)	(0.031)	(0.002)	(0.020)	(0.002)	(0.023)
Relative Wholesale Price X Leader	0.554***	1.282***				
	(0.008)	(0.092)				
Retail Markup	-0.000***	-0.001***	-0.001**	-0.007*	-0.000***	-0.001**
	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)
Retail Markup X Leader			0.202***	0.295***		
			(0.039)	(0.007)		
Retail Margin (log)					0.002***	-0.008
					(0.000)	(0.005)
Retail Margin X Leader					0.101***	0.130***
					(0.002)	(0.002)
Firm Scope (log)	-0.001	0.002	-0.006	-0.008	-0.001	-0.000
	(0.003)	(0.005)	(0.007)	(0.012)	(0.002)	(0.005)
Brand Age (log)	0.022***	0.007***	0.018***	0.007***	0.017***	0.008***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Rivals Brands (log)	-0.037***	-0.081***	-0.034***	-0.096***	-0.035***	-0.106***
	(0.002)	(0.008)	(0.003)	(0.004)	(0.002)	(0.005)
Observations	115505	115505	115505	115505	115489	115489
R-squared	0.546		0.603		0.630	
F-stat		34.7		11.1		81.5

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Revenue-based Market Share*, calculated as the share of sales of the product in the market using wholesale prices. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. *Relative Wholesale Price X Leader* is the interaction between *Relative Wholesale Price* and a dummy taking value 1 when the product has the highest market share. *Retail Markup* is the ratio between retail and wholesale price of the product. *Retail Markup X Leader* is the interaction between *Retail Markup* and a dummy taking value 1 when the product has the highest market share of sales. *Retail Margin (log)* is the difference between retail and wholesale price of the product in logs. *Retail Margin X Leader* is the interaction between *Retail margin (log)* and a dummy taking value 1 when the product has the highest market share of sales. Time, firm and market fixed effects are included. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*), *Retail Markup* and *Retail margin (log)* with firm average retail markup. All estimates are conducted on markets with five or more products. *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Comparing alternative production function estimates

		Revenue-based	Quantity-based
Cost share -based	Label	TFP-RC	TFP-QC
	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	$\alpha = .75; \beta_l = .08; \beta_m = .15; \beta_e = .02$	$\alpha = .75; \beta_l = .08; \beta_m = .15; \beta_e = .02$
	Literature	Foster et al. (2008)	Foster et al. (2008)
Estimation -based	Label	TFP-VE	TFP-QES
	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	$\alpha = .69; \beta_l = .32$	$\alpha = .52; \beta_l = -.04; \beta_m = .43$
	Literature	Ahsan (2013)	De Loecker et al. (2016)
	Label	TFP-RE	TFP-QEM
	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	$\alpha = .71; \beta_l = .11; \beta_m = .23$	$\alpha = .57; \beta_l = .20; \beta_m = .52; \gamma = -.06$
	Literature	Topalova and Khandelwal (2011)	Dhyne et al. (2017)

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 formulation level. The variables indicated in the production functions are expressed in logs: y is sales revenue, 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, respectively. The estimated elasticities reported are the industry average of formulation-level estimates.

Table 7: Relative wholesale price and product-level productivity
First stage of the baseline specification: OLS estimates

Dependent variable	Relative Wholesale Price					
Column	(1)	(2)	(3)	(4)	(5)	(6)
Revenue-based TFP						
TFP-RC	0.043***					
	(0.001)					
TFP-VE		0.006***				
		(0.001)				
TFP-RE			0.006***			
			(0.002)			
Quantity-based TFP						
TFP-QC				-0.099***		
				(0.001)		
TFP-QES					-0.017***	
					(0.002)	
TFP-QEM						-0.012***
						(0.001)
Retail markup	-0.000***	-0.000***	-0.000***	-0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Firm scope (log)	0.002	-0.002	-0.000	0.040***	0.003	-0.001
	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)	(0.008)
Brand age (log)	-0.013***	-0.016***	-0.016***	-0.009***	-0.009***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Rivals brands (log)	-0.194***	-0.194***	-0.194***	-0.206***	-0.196***	-0.198***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	131764	121417	131775	131764	131775	131770
R-squared	0.650	0.646	0.644	0.740	0.655	0.651

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Relative Wholesale Price*, computed as the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. The TFP measures (in logs) are defined as follows. *TFP-RC*: revenue-based productivity calculated using the share of costs following Foster et al. (2008); *TFP-VE*: value added-based productivity, based on Ahsan (2013); *TFP-RE*: revenue-based productivity calculated using Topalova and Khandelwal (2011); *TFP-QC*: quantity-based productivity calculated using the share of costs following Foster et al. (2008); *TFP-QES*: quantity-based productivity for single-product firms, based on De Loecker et al. (2016); *TFP-QEM*: quantity-based productivity for multiproduct firms based on Dhyne et al. (2017). Time, firm and market fixed effects are included.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: The determinants of quantity-based market shares
Baseline specification: OLS and IV estimates

Dependent variable	Quantity-based Market share					
Estimator	OLS			IV		
Column	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	N5	N10	All	N5	N10
Relative Wholesale Price	-0.107***	-0.073***	-0.046***	-1.269***	-1.042***	-0.828***
	(0.004)	(0.003)	(0.003)	(0.055)	(0.043)	(0.038)
Retail Markup	-0.000	0.000	-0.000***	-0.002***	-0.002**	-0.002***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Firm Scope (log)	-0.003	-0.001	-0.001	-0.004	-0.001	0.005
	(0.004)	(0.004)	(0.002)	(0.011)	(0.009)	(0.007)
Brand Age (log)	0.031***	0.028***	0.023***	0.014***	0.015***	0.014***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Rivals Brands (log)	-0.171***	-0.062***	-0.040***	-0.397***	-0.211***	-0.147***
	(0.004)	(0.003)	(0.002)	(0.014)	(0.009)	(0.007)
Observations	131770	115505	100867	131770	115505	100867
R-squared	0.673	0.297	0.212			
F-stat				798.5	1700.4	1717.4

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Quantity-based Market Share*, calculated as the product share of units sold in the market. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. A *Market* includes all the drugs having the same formulation and dosage form (tablet, injection, syrup, etc.). Time, firm and market fixed effects are included. Sample *All* includes all the markets, *N5* includes the markets with five or more products, *N10* includes the markets with ten or more products. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*). *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: The determinants of quantity-based market shares
Market leaders and market followers: OLS and IV estimates

Dependent variable	Quantity-based Market Share					
Sample	N5					
Column	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	IV	OLS	IV	OLS	IV
Relative Wholesale Price	-0.110***	-0.695***	-0.075***	-0.652***	-0.079***	-0.698***
	(0.003)	(0.035)	(0.003)	(0.027)	(0.003)	(0.030)
Relative Wholesale Price X Leader	0.425***	1.142***				
	(0.009)	(0.087)				
Retail Markup	0.000	-0.001	-0.000	-0.007**	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.003)	(0.000)	(0.001)
Retail Markup X Leader			0.175***	0.258***		
			(0.034)	(0.006)		
Retail Margin (log)					0.002***	-0.011
					(0.001)	(0.007)
Retail Margin X Leader					0.087***	0.114***
					(0.002)	(0.002)
Firm Scope (log)	0.000	0.003	-0.004	-0.006	0.000	0.001
	(0.004)	(0.007)	(0.006)	(0.011)	(0.003)	(0.007)
Brand Age (log)	0.022***	0.006***	0.018***	0.006***	0.018***	0.006***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Rivals Brands (log)	-0.052***	-0.116***	-0.048***	-0.131***	-0.049***	-0.140***
	(0.002)	(0.009)	(0.003)	(0.006)	(0.002)	(0.006)
Observations	115505	115505	115505	115505	115489	115489
R-squared	0.425		0.501		0.515	
F-stat		34.7		11.1		81.5

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Quantity-based Market Share*, calculated as the product share of units sold in the market. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. *Relative Wholesale Price X Leader* is the interaction between *Relative Wholesale Price* and a dummy taking value 1 when the product has the highest market share. *Retail Markup* is the ratio between retail and wholesale price of the product. *Retail Markup X Leader* is the interaction between *Retail Markup* and a dummy taking value 1 when the product has the highest market share of sales. *Retail Margin (log)* is the difference between retail and wholesale price of the product in logs. *Retail Margin X Leader* is the interaction between *Retail margin (log)* and a dummy taking value 1 when the product has the highest market share of sales. Time, firm and market fixed effects are included. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*), *Retail Markup* and *Retail margin (log)* with firm average retail markup. All estimates are conducted on markets with five or more products. *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Demand function and price elasticity
Baseline specification: OLS and IV estimates

Dependent variable	Units (log)					
Sample	All			N5		
Column	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	OLS	IV	OLS	OLS	IV
Wholesale Price (log)	-0.911***	-0.882***	-5.624***	-0.903***	-0.876***	-5.663***
	(0.012)	(0.012)	(0.146)	(0.012)	(0.012)	(0.146)
Retail Markup		-0.038***	-0.162***		-0.038***	-0.160***
		(0.011)	(0.036)		(0.011)	(0.036)
Firm Scope (log)		0.191**	0.341		0.193*	0.335
		(0.096)	(0.220)		(0.102)	(0.239)
Brand Age (log)		0.993***	0.556***		1.006***	0.622***
		(0.022)	(0.058)		(0.022)	(0.059)
Rivals Brands (log)		-0.326***	-0.107		-0.253***	-0.127
		(0.045)	(0.114)		(0.054)	(0.149)
Observations	138238	131770	131770	120879	115505	115505
R-squared	0.567	0.587		0.528	0.550	
F-stat			1173.0			3276.2

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Units (log)*, calculated as the logarithm of normalized product units sold. *Wholesale Price (log)* is the logarithm of normalized product wholesale price. Time, firm and market fixed effects are included. Sample *All* includes all the markets, *N5* includes the markets with 5 or more products. In IV estimates *Wholesale Price (log)* is instrumented with quantity-based productivity (*TFP-QEM*). *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Demand function and price elasticity
Market leaders and market followers: OLS and IV estimates

Dependent variable	Units (log)					
Sample	N5					
Column	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	IV	OLS	IV	OLS	IV
Wholesale Price (log)	-0.902***	-5.233***	-0.879***	-5.235***	-0.924***	-5.381***
	(0.012)	(0.131)	(0.012)	(0.131)	(0.012)	(0.141)
Wholesale Price X Leader	0.477***	0.772***				
	(0.008)	(0.078)				
Retail Markup	-0.036***	-0.145***	-0.040***	-0.372***	-0.059***	-0.119***
	(0.011)	(0.033)	(0.012)	(0.087)	(0.015)	(0.032)
Retail Markup X Leader			1.852***	2.799***		
			(0.371)	(0.104)		
Retail Margin (log)					0.554***	-0.807***
					(0.020)	(0.245)
Retail Margin X Leader					0.884***	1.274***
					(0.017)	(0.045)
Firm Scope (log)	0.227**	0.377*	0.163	0.261	0.201**	0.365
	(0.100)	(0.220)	(0.119)	(0.266)	(0.099)	(0.233)
Brand Age (log)	0.899***	0.487***	0.907***	0.497***	0.909***	0.483***
	(0.021)	(0.056)	(0.028)	(0.055)	(0.021)	(0.058)
Rivals Brands (log)	-0.099*	0.111	-0.104*	0.042	-0.126**	0.086
	(0.053)	(0.138)	(0.056)	(0.139)	(0.051)	(0.145)
Observations	115505	115505	115505	115505	115489	115489
R-squared	0.578		0.576		0.587	
F-stat		29.6		11.1		90.9

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Units (log)*, calculated as the logarithm of normalized product units sold. *Wholesale Price (log)* is the logarithm of normalized product wholesale price. *Wholesale Price X Leader* is the interaction between *Wholesale Price* and a dummy taking value 1 when the product has the highest market share. *Retail Markup* is the ratio between retail and wholesale price of the product. *Retail Markup X Leader* is the interaction between *Retail Markup* and a dummy taking value 1 when the product has the highest market share of sales. *Retail Margin (log)* is the difference between retail and wholesale price of the product in logs. *Retail Margin X Leader* is the interaction between *Retail margin (log)* and a dummy taking value 1 when the product has the highest market share of sales. Time, firm and market fixed effects are included. In IV estimates *Wholesale Price (log)* is instrumented with quantity-based productivity (*TFP-QEM*), *Retail Markup* and *Retail margin (log)* with firm average retail markup. All estimates are conducted on markets with five or more products. *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: The determinants of market shares
Subsample analysis on firm and market characteristics: IV estimates

Dependent variable	Market share							
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subsample	Domestic	Multinational	Solids	Liquids	Injectables	Before 1995	1995-2004	After 2005
Relative Wholesale Price	-1.053***	-1.199***	-1.311***	-1.141***	-0.588***	-1.004***	-1.201***	-1.016***
	(0.054)	(0.131)	(0.098)	(0.106)	(0.044)	(0.062)	(0.102)	(0.130)
Retail Markup	-0.003***	-0.004***	-0.004***	-0.009**	-0.001***	-0.005	-0.002***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.000)	(0.003)	(0.001)	(0.001)
Firm Scope (log)	-0.001	-0.107***	-0.024*	0.052*	0.002	-0.003	0.017	-0.071**
	(0.010)	(0.034)	(0.013)	(0.031)	(0.014)	(0.015)	(0.015)	(0.031)
Brand Age (log)	0.013***	0.023***	0.019***	0.008	0.019***	0.018***	0.013***	0.036***
	(0.002)	(0.008)	(0.003)	(0.007)	(0.004)	(0.003)	(0.004)	(0.008)
Rivals Brands (log)	-0.350***	-0.380***	-0.370***	-0.410***	-0.322***	-0.352***	-0.368***	-0.427***
	(0.013)	(0.037)	(0.020)	(0.032)	(0.021)	(0.018)	(0.023)	(0.037)
Observations	117231	14434	89024	19275	19869	56683	56616	17578
F-stat	663.4	239.5	322.7	188.6	891.7	999.8	195.0	74.3

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Revenue-based Market Share*, calculated as the share of sales of the product in the market using wholesale prices. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. *Retail Markup* is the ratio between retail and wholesale price of the product. Time, firm and market fixed effects are included. The subsamples *Multinational* and *Domestic* include, respectively, multinational companies and domestic companies. The subsamples *Solid*, *Liquid* and *Injectables* include, respectively, drugs whose dosage form is solid (e.g. tablet), liquid (e.g. syrup) and injectable (e.g. syringe). The subsamples *Before 1995*, *1995-2004* and *Since 2005* include, respectively, drug formulation launched before 1995, between 1995 and 2004, from 2005 onwards. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*), *Retail Markup* and *Retail margin (log)* with firm average retail markup. *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: The determinants of market shares
Subsample analysis on drug characteristics: IV estimates

Dependent variable	Market share							
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Subsample	Acute	Chronic	Single	Combination	Non-regulated	Regulated	Over-the-counter	Prescription
Relative Wholesale Price	-0.995***	-1.409***	-1.155***	-0.962***	-1.004***	-3.864***	-1.002***	-1.169***
	(0.054)	(0.131)	(0.073)	(0.066)	(0.049)	(0.968)	(0.064)	(0.076)
Retail Markup	-0.003***	-0.005	-0.002***	-0.008***	-0.003***	-0.008*	-0.003***	-0.002***
	(0.001)	(0.004)	(0.001)	(0.002)	(0.001)	(0.005)	(0.001)	(0.001)
Firm Scope (log)	-0.002	-0.021	-0.002	-0.009	-0.007	-0.107	-0.013	0.005
	(0.011)	(0.020)	(0.013)	(0.014)	(0.010)	(0.088)	(0.013)	(0.015)
Brand Age (log)	0.016***	0.021***	0.023***	0.007*	0.014***	0.065***	0.010***	0.028***
	(0.002)	(0.005)	(0.003)	(0.004)	(0.002)	(0.016)	(0.003)	(0.003)
Rivals Brands (log)	-0.343***	-0.455***	-0.405***	-0.312***	-0.346***	-1.039***	-0.349***	-0.383***
	(0.014)	(0.034)	(0.019)	(0.016)	(0.012)	(0.250)	(0.015)	(0.021)
Observations	93767	37998	84481	47282	116672	15080	76014	55752
F-stat	219.8	144.7	187.6	138.1	286.8	16.8	147.0	348.9

Source: AIOCD and Prowess, CMIE.
Notes: The dependent variable is *Revenue-based Market Share*, calculated as the share of sales of the product in the market using wholesale prices. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. *Retail Markup* is the ratio between retail and wholesale price of the product. Time, firm and market fixed effects are included. The subsamples *Acute* and *Chronic* include drugs used for treating, respectively, acute and chronic. The subsamples *Single* and *Combination* include drugs composed by, respectively, a single formulation and the combination of two or more formulations. The subsamples *Non-regulated* and *Regulated* include drugs whose price is, respectively, non-regulated and regulated by the Indian government. The subsamples *Over-the-counter* and *Prescription* include drugs that, respectively, do not and do need a medical prescription to be purchased. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*), *Retail Markup* and *Retail margin (log)* with firm average retail markup. *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A Product-level productivity in multiproduct firms

As discussed in De Loecker et al. (2016), to estimate product-level productivity correctly we must deal with the impossibility of measuring product-level inputs. The first aspect can lead to two potential biases: i) the *input allocation bias*, concerning the possible mismeasurement in the process of allocating the shares of firm-level inputs to each product; ii) the *input price bias*, concerning the differences in purchase prices of the same input across products of different quality. In the following two subsections we present and discuss how we address these biases. In the last subsection we describe the changes introduced to the standard LP estimator in order to estimate productivity using Equation (7).

A.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 using methods to avoiding the problem. Foster et al. (2008) apportion product's share of plant inputs using product's share of plant sales.[11] De Loecker et al. (2016) avoid product's input allocation by estimating productivity using only single product firms.[12] Dhyne et al. (2017) implement a technique to estimate product-level productivity using only firm-level inputs. We exploit specific features of the pharmaceutical industry to make assumptions and impute the values of product variable inputs.

The pharmaceutical industry is composed of a large number of markets within which drugs have the same therapeutic category, i.e. are used to treat the same diseases. The unit cost of variable inputs - raw materials and labor - can be assumed to be the same across all products within the market. Since the chemical composition of the drugs within a market is unique, we assume that the cost of raw materials (bulk drugs) used to produce one unit of the drug does not vary across firms. 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. Given the highly automated production process, the working skills required to produce a drug are common across firms. Therefore, we assume that the cost of labor used to produce one unit of drug does not vary across firms within the market. To identify the cost per unit produced of each variable input, for each market we select the firm 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 market.

To impute the expenditure in variable input for all the products of a market, we leverage on the reference firms ($\bar{f}$). First, we calculate its input expenditure in the referenced market ($\bar{j}$) using the

[11] The method is valid under perfect competition or assuming constant markups across firm products. Since Foster et al. (2008) 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 the product of interest), these assumptions are appropriate.

[12] 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 productivity 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.

market's revenue shares of the reference firm, $\frac{y_{\bar{f}\bar{j}t}}{y_{\bar{f}t}}$:[13]

$$\mathbf{v}_{\bar{f}\bar{j}t} = \mathbf{v}_{\bar{f}t} \frac{y_{\bar{j}\bar{f}\bar{j}t}}{y_{\bar{f}t}} \tag{9}$$

Second, we split reference firm's input expenditure in the referenced market ($\mathbf{v}_{\bar{j}\bar{f}t}$) across all its products (i) using product's share of physical units produced in the market by the firm, $\frac{q_{i\bar{f}\bar{j}t}}{q_{\bar{f}\bar{j}t}}$:[14]

$$\mathbf{v}_{i\bar{f}\bar{j}t} = \mathbf{v}_{\bar{f}\bar{j}t} \frac{q_{i\bar{f}\bar{j}t}}{q_{\bar{f}\bar{j}t}} \tag{10}$$

Since we assumed the unit cost of variable inputs to be the same for all the products within the market, we can impute the input cost for all products of all other firms (f) in the referenced market ($\bar{j}$) by proportionally rescaling reference firm's product input cost for every product's physical units produced in the market ($q_{if\bar{j}t}$).

$$\mathbf{v}_{if\bar{j}t} = \mathbf{v}_{i\bar{f}\bar{j}t} \frac{q_{if\bar{j}t}}{q_{i\bar{f}\bar{j}t}} \tag{11}$$

We use firm-level input data from the Prowess dataset. The measure of capital that we adopt is the variable "capital employed" included in the data. It is measured as the sum of equity capital, non-reevaluated reserves and borrowings. We use this measure of capital as the fixed asset variables in Prowess have many missing values. Labor is defined as the amount of salaries and wages of the firm, as employment variables are not reliable enough. Materials are measured as the raw material expenditure of the firm, excluding consumption of stores and spares. Variable inputs are deflated by pharmaceutical 4-digit NIC wholesale price index. Following Ahsan (2013), capital is deflated using an investment deflator, computed as the average of the wholesale price index for two industries: "manufacture of general purpose machinery" and "manufacture of special purpose machinery".

A.2 Product's input price

Product price dispersion within an industry may depend on the difference in quality among the products, which in turn may stem from different input quality, and different input costs. Since the bulk drugs used to obtain the final drugs have the same chemical composition and the workers in the chain of one product do not need to be more skilled than the other workers in the same market, we assume that input quality and input prices are the same across all products within a market. In principle in the pharmaceutical industry within the market, products should be materially and qualitatively homogeneous, as every drug has the same formulation and dosage form. In a cross country study, Bate et al. (2011) test the quality of drug samples and observe the drugs failing the test are priced lower than those which comply with standardized quality measures. However, they also show that price differences alone is insufficient to identify the quality of drugs. Bennett and Yin (2014) conduct a quality test on the most important antibiotics in India and show that 96 percent of the drugs sampled comply with Indian Pharmacopoeia quality standards. Yet, in the narrowly defined medicine markets that we compare, the magnitude of *actual* quality differences documented in previous studies

[13]To do so we have to assume that the reference firm has constant markup over all products in the referenced market.

[14]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.

alone cannot explain the sizeable dispersion in prices observed in our data (Figure 1). Moreover, in our estimation sample we consider only traded firms which are supposed to be more observant (and controlled) about quality aspects. We, therefore, consider product quality dispersion within the market a limited problem for our input price assumption.

The productivity measure we adopt to instrument for the prices in Equation (5) does not require to allocate firm-level capital across the products. However, in Section 6.1 we propose five other measures of productivity. No specific features of Indian pharmaceutical industry, help us make assumptions about the difference in price of the capital goods employed for a product. In that case, to impute product-level capital we simply apportion firm-level capital among the different products of the firm using product's share of firm sales as in Foster et al. (2008). We stick to the O-Ring theory by Kremer (1993) and to Kugler and Verhoogen (2011), which model and show that more expensive inputs lead to more expensive products. Product's share of firm sales, that we use for apportioning firm-level capital among the products, embeds this information.

An important assumption we make on input prices is that they do not depend on input quantities.[15] 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 percent of the reference firms have the highest sales share in the referenced market, implying that less than 13 percent of the reference firms are top purchasers on their input markets.

A.3 The LP estimator controlling for product scope

Dhyne et al. (2017) propose that all kind of inputs used by a multiproduct firm can create a synergy, allowing the firm to reach a higher point on the production possibility curve with the same amount of inputs. In the pharmaceutical industry, however, variable inputs can be considered product-specific. Firm capital expenditure, instead, is more likely to involve many products. To contrast the simultaneity bias, we estimate Equation (7):

$$q_{it} = \alpha k_{ft} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \gamma y_{-it} + \eta_{it}$$

adopting the LP technique, using materials as a proxy. Similar to Dhyne et al. (2017), we must modify the standard LP estimator as follows.

The same assumptions as LP must hold at the product level: (i) the demand for the intermediate input m is dependent on the two state variables and it is monotonically increasing in ω and, thus, can be inverted[16]

$$m_{it} = \theta(k_{ft}, \omega_{it}) \rightarrow \omega_{it} = \theta^{-1}(k_{ft}, m_{it}) \tag{12}$$

(ii) the law of motion of productivity, i.e. a first order Markov-chain process:

$$\omega_{it} = E[\omega_{it}|\omega_{it-1}] + \psi_{it} \tag{13}$$

where ψ_{it} is an innovation to productivity, uncorrelated with k_{ft} but not necessarily with l_{it}.

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

[16] Contrary to Dhyne et al. (2017) the equation is invertible as the materials are measured at the product level, creating a one-to-one relationship with product-level productivity.

We can rewrite the production function as:

$$q_{it} = \beta_l l_{it} + \phi(k_{ft}, m_{it}) + \gamma y_{-it} + \eta_{it} \tag{14}$$

where, as in the firm-level case:

$$\phi(k_{ft}, m_{it}) = \beta_0 + \beta_k k_{ft} + \beta_m m_{it} + \theta^{-1}(k_{ft}, m_{it}) \tag{15}$$

In the first stage of the LP estimation, Equation (15) can be substituted by a third-order polynomial approximation in k_{ft} and m_{it}. Equation (14) can, therefore, be rewritten as:

$$q_{it} = \delta_0 + \beta_l l_{it} + \sum_{g=0}^{3} \sum_{h=0}^{3-g} \delta_{gh} k_{ft}^{g} m_{it}^{h} + \gamma y_{-it} + \eta_{it} \tag{16}$$

and estimated consistently using OLS. From the first stage we obtain an estimation of $\widehat{\beta_l}$, $\widehat{\gamma}$, $\widehat{\delta_0}$ and $\widehat{\delta_{gh}}$. from the second stage we identify β_k and β_m. We first compute the predicted value of q_{it}, as:

$$\widehat{q_{it}} = \widehat{\delta_0} + \widehat{\beta_l} l_{it} + \sum_{g=0}^{3} \sum_{h=0}^{3-g} \widehat{\delta_{gh}} k_{ft}^{g} m_{it}^{h} + \widehat{\gamma} y_{-it}$$

and use it to obtain the prediction of $\phi(k_{ft}, m_{it})$:

$$\widehat{\phi}(k_{ft}, m_{it}) = \widehat{\delta_0} + \sum_{g=0}^{3} \sum_{h=0}^{3-g} \widehat{\delta_{gh}} k_{ft}^{g} m_{it}^{h} \tag{17}$$

Since $\phi(k_{ft}, m_{it}) = \beta_0 + \beta_k k_{ft} + \beta_m m_{it} + \theta^{-1}(k_{ft}, m_{it})$ for any candidate pair (β_k^*, β_m^*) we can compute a prediction for ω_{it} $(\widehat{\omega_{it}^*})$ in all periods t, as:

$$\widehat{\omega_{it}^*} = \widehat{\phi}(k_{ft}, m_{it}) - \beta_k^* k_{ft} - \beta_m^* m_{it} \tag{18}$$

The law of motion of productivity is a function of the productivity expectations $E[\omega_{it}|\omega_{it-1}]$, i.e. the expectations determine the prediction of ω_{it} $(\widehat{\omega_{it}})$. We can approximate the expectation function as:

$$E[\omega_{it}|\omega_{it-1}] = \gamma_0 + \gamma_1 \omega_{it-1} + \gamma_2 \omega_{it-1}^2 + \gamma_3 \omega_{it-1}^3 + \epsilon_{it} \tag{19}$$

So that:

$$\omega_{it} = \gamma_0 + \gamma_1 \omega_{it-1} + \gamma_2 \omega_{it-1}^2 + \gamma_3 \omega_{it-1}^3 + \varepsilon_{it}$$

where $\varepsilon_{it} = \epsilon_{it} + \psi_{it}$. Using every possible value of $\widehat{\omega_{it}^*}$ computed in all periods t, we can estimate:

$$\widehat{\omega_{it}^*} = \gamma_0 + \gamma_1 \widehat{\omega_{it-1}^*} + \gamma_2 \widehat{\omega_{it-1}^*}^2 + \gamma_3 \widehat{\omega_{it-1}^*}^3 + \varepsilon_{it} \tag{20}$$

Obtaining estimates for $\widehat{\gamma_0}, \widehat{\gamma_1}, \widehat{\gamma_2}, \widehat{\gamma_3}$. With this estimated coefficient we predict the value of the dependent variable $\widehat{\omega_{it}^*}$:

$$\begin{aligned}\widehat{\widehat{\omega_{it}^*}} &= \widehat{\gamma}_0 + \widehat{\gamma}_1 \widehat{\omega_{it-1}^*} + \widehat{\gamma}_2 \widehat{\omega_{it-1}^*}^2 + \widehat{\gamma}_3 \widehat{\omega_{it-1}^*}^3 \\ &= E[\widehat{\omega_{it}^* | \omega_{it-1}^*}]\end{aligned} \tag{21}$$

Remember Equation (12), we can rewrite Equation (7) as:

$$\begin{aligned} q_{it} &= \beta_l l_{it} + \beta_k k_{ft} + \beta_m m_{it} + E[\omega_{it}|\omega_{it-1}] + \psi_{it} + \gamma y_{-it} + \eta_{it} \\ \psi_{it} + \eta_{it} &= q_{it} - \beta_l l_{it} - \beta_k k_{ft} - \beta_m m_{it} - E[\omega_{it}|\omega_{it-1}] - \gamma y_{-it}\end{aligned}$$

Using $\widehat{\beta_l}$ and $\widehat{\gamma}$ estimated in the first stage, we can then compute $\widehat{\psi_{it} + \eta_{it}}$ for every candidate pair (β_k^*, β_m^*) and $E[\widehat{\omega_{it}^* | \omega_{it-1}^*}]$, as:

$$\widehat{\psi_{it} + \eta_{it}} = q_{it} - \widehat{\beta_l} l_{it} - \beta_k^* k_{ft} - \beta_m^* m_{it} - E[\widehat{\omega_{it}^* | \omega_{it-1}^*}] - \widehat{\gamma y_{-it}} \tag{22}$$

This residual must interact with at least two instrument to identify both β_k and β_m. Then, we need two moment conditions. If period t's capital stock is determined by the previous period's investment decision, it does not respond to shocks to this period's productivity innovation term ψ_{it}, providing the moment condition:

$$E[\psi_{it} + \eta_{it} | k_{ft}] = 0 \tag{23}$$

An additional moment condition is needed. Levinsohn and Petrin (2003) exploit the fact that previous period's materials are uncorrelated with present period's error term:

$$E[\psi_{it} + \eta_{it} | m_{it-1}] = 0 \tag{24}$$

Thus, with $\mathbf{Z}_{it} = (k_{ft}, m_{it-1})$, the GMM estimator solves:

$$\min_{(\beta_k^*, \beta_m^*)} \sum_h \left[\sum_i \sum_t (\widehat{\psi_{it} + \eta_{it}}) Z_{hit} \right]^2 \tag{25}$$

and identifies β_k^* and β_m^* separately.[17]

With the estimates of $\widehat{\beta_k}$ and $\widehat{\beta_m}$, and $\widehat{\beta_l}$ and $\widehat{\gamma}$ estimated in the first stage we calculate product-level productivity ($TFP - QEM$) as a residual:

$$\widehat{tfp_{it}} = q_{it} - \widehat{\beta_l} l_{it} - \widehat{\beta_k} k_{ft} - \widehat{\beta_m} m_{it} - \widehat{\gamma} y_{-it} \tag{26}$$

[17] Additional overidentification conditions define a conditioning vector $\mathbf{Z}_{it} = (k_{ft}, m_{it-1}, l_{it-1}, m_{it-2}, k_{ft-1}, y_{-it-1})$

B Additional Tables

Table B.1: The determinants of market shares
Baseline specification: OLS on full sample AIOCD

Dependent variable	Revenue-based Market Share			Quantity-based Market Share		
Column	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All	N5	N10	All	N5	N10
Relative Wholesale Price	0.018***	0.008***	0.009***	-0.102***	-0.070***	-0.044***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Retail Markup	-0.000**	-0.000**	-0.000*	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Firm Scope (log)	-0.002	-0.002	-0.001	-0.003	-0.002	-0.002*
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Brand Age (log)	0.029***	0.024***	0.019***	0.027***	0.023***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rivals Brands (log)	-0.145***	-0.039***	-0.022***	-0.165***	-0.048***	-0.028***
	(0.003)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
Observations	310627	273562	243132	310627	273562	243132
R-squared	0.691	0.256	0.198	0.674	0.238	0.173

Source: AIOCD.
Notes: The dependent variables are: *Revenue-based Market Share*, calculated as the share of sales of the product in the market using wholesale prices; and *Quantity-based Market Share*, calculated as the product share of units sold in the market. *Relative Wholesale Price* is the ratio between product's normalized wholesale price and the highest normalized wholesale price in the market. A *Market* includes all the drugs having the same formulation and dosage form (tablet, injection, syrup, etc.). Time, firm and market fixed effects are included. Sample *All* includes all the markets, *N5* includes the markets with five or more products, *N10* includes the markets with ten or more products. In IV estimates *Relative Wholesale Price* is instrumented with quantity-based productivity (*TFP-QEM*). *F-stat* is Kleibergen-Paap F statistic.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: Wholesale price and product-level productivity
First stage of the demand function estimation: OLS estimates

Dependent variable	Wholesale Price (log)					
Column	(1)	(2)	(3)	(4)	(5)	(6)
Revenue-based TFP						
TFP-RC	0.372***					
	(0.011)					
TFP-VE		0.036***				
		(0.005)				
TFP-RE			0.052***			
			(0.013)			
Quantity-based TFP						
TFP-QC				-0.873***		
				(0.007)		
TFP-QES					-0.158***	
					(0.016)	
TFP-QEM						-0.117***
						(0.006)
Retail Markup	-0.024***	-0.028***	-0.025***	-0.008***	-0.026***	-0.027***
	(0.005)	(0.007)	(0.005)	(0.002)	(0.006)	(0.006)
Firm Scope (log)	0.062	0.022	0.040	0.398***	0.072*	0.034
	(0.043)	(0.041)	(0.040)	(0.052)	(0.039)	(0.039)
Brand Age (log)	-0.078***	-0.102***	-0.104***	-0.042***	-0.035***	-0.019*
	(0.011)	(0.012)	(0.012)	(0.004)	(0.012)	(0.011)
Rivals Brands (log)	0.057**	0.062**	0.051**	-0.049***	0.033	0.022
	(0.023)	(0.024)	(0.023)	(0.011)	(0.021)	(0.022)
Observations	131764	121417	131775	131764	131775	131770
R-squared	0.844	0.844	0.837	0.958	0.852	0.848

Source: AIOCD and Prowess, CMIE.
Notes: *Wholesale Price (log)* is the logarithm of normalized product wholesale price. The productivity measures (in logs) are defined as follows. *TFP-RC*: revenue-based productivity calculated using the share of costs following Foster et al. (2008); *TFP-VE*: value added-based productivity, based on Ahsan (2013); *TFP-RE*: revenue-based productivity calculated using Topalova and Khandelwal (2011); *TFP-QC*: quantity-based productivity calculated using the share of costs following Foster et al. (2008); *TFP-QES*: quantity-based productivity for single-product firms, based on De Loecker et al. (2016); *TFP-QEM*: quantity-based productivity for multiproduct firms based on Dhyne et al. (2017). Time, firm and market fixed effects are included.
Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.