STEG WORKING PAPER

DEMAND FOR QUALITY, VARIABLE MARKUPS AND MISALLOCATION: EVIDENCE FROM INDIA

Apoorv Gupta

MAY 2023
STEG WP064

Demand for Quality, Variable Markups and Misallocation: Evidence from India

Apoorv Gupta
Department of Economics, Dartmouth College *

Abstract

Markups vary systematically across firms and generate misallocation, yet empirical evidence on sources driving markup variation is limited. I study how demand-side factors affect markups. Using detailed firm-product-level data from India, I document that *both* marginal costs and markups are increasing in firm-size. Changes in markups across the firm-size distribution in response to exogenous demand shocks to poor households lend support to the demand-based markup channel: producing better quality and selling to wealthier, less demand elastic households leads larger firms to incur higher costs and charge higher markups. Accounting for the demand-based channel reduces estimated misallocation losses by 30 percent.

Keywords: Productivity, Incomplete Pass-through, Market Power, Income Distribution.
JEL Classification: D24, D61, E20, L1, O47

*Contact: apoorv.gupta@dartmouth.edu.

This version: September, 2022. I am indebted to Effi Benmelech, Nicolas Crouzet, Dean Karlan, Dimitris Papanikolaou and Jacopo Ponticelli for their invaluable inputs. The paper has benefited from insightful discussions with and comments by David Atkin, David Berger, Vivek Bhattacharya, Diego Comin, Adrien Matray, Martí Mestieri, Terry Moon, Nina Pavcnik, Simon Quinn, Chris Snyder, Chad Syverson, Eric Verhoogen, Jonathan Zinman and as well as participants at Chicago Booth, Rochester Simon, Dartmouth, Berkeley Haas, Imperial College London, Maryland, Maryland Smith, University of Hong Kong, Bocconi, CAFRAL (Reserve Bank of India), Indian School of Business, Kellogg School of Management, Northwestern University (Macro Lunch, Development Lunch, Applied Micro), PEDL Young Scholars Conference, STEG Annual Workshops, EMCON, and NEUDC for helpful comments and suggestions. All mistakes are my own. The data used in this paper come from the Ministry of Statistics and Programme Implementation, Government of India. A previous version of this paper was circulated under the title "Firm Heterogeneity, Demand for Quality and Prices: Evidence from India".

1 Introduction

An important line of research in economics documents large differences in markups across firms. Such variation in markups is considered a significant, if not the leading, cause behind misallocation of resources across firms that potentially lowers the aggregate productivity in an economy (Hsieh and Klenow 2009, 2014; Bartelsman, Haltiwanger, and Scarpetta 2013; David and Venkateswaran 2019). Much of the literature, however, routinely treats firms' market power as exogenous and abstracts from the underlying sources driving differences in firms' markups.[1] If dispersion in markups stems largely from exogenous market frictions, then eliminating those frictions will improve resource allocation and increase aggregate productivity. However, if driven by firms choosing their optimal markups based on heterogeneous demand elasticities (i.e., demand-side factors), potential gains from reallocation will be limited.

I provide evidence on how demand-side features shape the relationship between firm size and markups. I show that segmentation in output product market coupled with differences in demand elasticities across consumers with different income levels can allow large systematic dispersion in markups to persist in equilibrium. An implication is that gains from policies that reallocate resources across firms will be limited because larger firms charge higher markups as they face low demand elasticities. Firms that face lower demand elasticities will respond to reallocation policies — such as subsidies intended to counteract markup distortions — by increasing their markup. This endogenous markup adjustment by firms lowers potential gains from reallocation. I document that estimates of the extent to which firms pass through changes to their costs into their prices is a sufficient statistic for correcting the bias in reallocation gains under endogenous markups.

I begin by using data from Indian Annual Survey of Industries (ASI) — a highly detailed panel on Indian manufacturing firms from 1998 to 2009 — to document a systematic relation between firms' size, and their marginal costs and markups. To derive markups at the firm-product level, I apply the methodology pioneered by De Loecker and Warzynski (2012). An important feature of the data is that it provides information separately on revenues as well as physical units for each firm-product, allowing to calculate product prices (unit values). Dividing the prices by corresponding markups yields marginal costs at the firm-product level. The ability to observe both marginal costs and markups allows to differentiate between existing models of firm heterogeneity that differ in their predictions on how costs and markups varies with firm size. I show that, first, marginal costs within a product group are increasing in firm size. This finding is consistent with the literature on product quality (Kugler and Verhoogen 2011; Atkin, Chaudhry, Chaudhry, Khandelwal, and Verhoogen 2015), and in line with the findings in these papers, I also find that prices for inputs are

[1] Notable exceptions include Atkeson and Burstein (2008); Haltiwanger, Kulick, and Syverson (2018); Dhingra and Morrow (2019); Edmond, Midrigan, and Xu (2019); Peters (2020).

higher for larger firms. Second, and more importantly, I find that the markups are also increasing in firm size. These relationships are more pronounced in sectors with greater scope for quality differentiation, as proxied by the Rauch (1999) classification of non-homogeneous goods.

Why are markups increasing in firm-size? I posit that larger firms have higher markups because of *assortative matching* between firms producing higher quality goods and wealthier consumers. The approach is motivated by two theoretical ideas. First, following Linder (1961), consumers are asymmetric in income and their willingness to pay for product quality; and firms producing higher quality goods cater to the demand of wealthier households. Second, firm productivity and input quality are complements in determining output quality, and in equilibrium higher quality is produced by more productive and larger firms (Kugler and Verhoogen 2011). Together, this implies that wealthier households source larger share of their consumption from goods produced by larger firms, particularly in the quality differentiated sector. Because wealthier households are less price sensitive (as I document), this matching on product quality implies that larger firms charge higher markups. I term this as the *demand-based* markup channel.

Next, I provide evidence on the demand-based markup channel by tracing how firms change their markups to changes in their demand (composition). I propose an empirical strategy that uses quasi-exogenous income shocks to poor households, both across districts and over time, as a source of variation in their demand. The majority of the poor households in India are employed in agricultural sector and face substantial productivity risk — even today, less than one-third of the agricultural land is irrigated —, making their incomes significantly dependent on the local rainfall variation. Because consumers across income levels differ in the shares of their consumption basket sourced from large-, mid-, and small-sized firms, these changes to poor households' income affect the demand faced by small- and mid-sized firms more than large firms.

I find that firms lower their markups by 0.5 percent in response to an increase in poor households' income driven by positive rain shocks. I find no affect of demand shocks on firms' marginal costs or its underlying components: physical productivity (TFPQ) and input prices. The demand-based markup channel above posits that any changes in demand from lower income groups should affect the weighted demand elasticity, and hence markups, only for firms selling to both rich and poor households. These firms are proxied in my data by firms in the middle of the size distribution. To test for this channel, I examine how firms across the size distribution change their markups in response to the rain shocks. I find a non-monotonic effect on markups across the firm-size distribution. Specifically, mid-sized firms lower their markups by 1 percent in response to higher demand from the poor. In contrast, markups for firms in the lower and upper ranges of the distribution remain unchanged. More importantly, these responses are only present in quality differentiated sectors, where markup variation is expected to be a consequence of differences in demand composition. I argue that this non-monotonic markup response to demand shocks to the poor is unique to the demand-based markup channel, and provide empirical evidence inconsistent with alternative mechanisms.

These results suggest that differences in demand are an important source of dispersion in markups — and hence marginal revenue product (MRP) of inputs — across firms. What are the consequences for aggregate productivity arising from *demand-based* markup dispersion? Since demand factors are not prone to reallocation, the aggregate productivity gains that could be attained from a reallocation exercise are correspondingly smaller. To quantify the gains from reallocation, I use the aggregate productivity growth decomposition from Petrin and Levinsohn (2012), and consider a tax-subsidy policy that subsidizes (taxes) firms with high (low) markups while assuming markup distortion to be exogenous. Specifically, I consider a policy that serves a planner's objective to equalize MRP within industries under a fixed aggregate supply of resources, while (erroneously) assuming that any variation in MRP across firms is generated only from exogenous distortions.[2]

The main result from the exercise is that when markups are endogenous, firms could adjust their markups in response to tax-subsidy policies. I propose a sufficient statistic — the estimate of pass-through of changes to firm's costs into its prices — for firms' markup adjustment to the reallocation policy. The knowledge of the pass-through rate, however, is insufficient to separate the contribution of market conduct and consumer demand. This is particularly relevant in oligopolistic settings, as under imperfect competition both the curvature of demand and its elasticity affect firms' pass-through. I propose a methodology, motivated by the framework in Weyl and Fabinger (2013) and Atkin and Donaldson (2015), to separately identify the underlying determinants of the pass-through rate. The framework applies the results presented above — that is, firms in homogeneous sector face the same slope of demand, while it can vary across firms in differentiated sector due to differences in their demand composition — to separately identify firm conduct and demand parameters from estimated markups and pass-through rates. This allows me to assess reallocation gains across different values of market conduct, while keep the demand parameters fixed.

I estimate firm-level pass-through rates and find them to be decreasing in firm size, with the relationship stronger in differentiated sectors. Because high markup firms also have lower pass-through rates, productivity gains are substantially lower from the reallocation policy. I estimate that the reallocation gains are substantially large (about 47 percent) when pass-through is assumed

[2]The presence of firm-specific distortions or markups is not the only source of dispersion in MRP of inputs. Few studies use detailed micro-data to attribute the observed dispersion in marginal products into underlying economic forces unrelated to misallocation. These studies attribute this dispersion to unobserved heterogeneity in physical productivity (Gollin and Udry 2021), or adjustment costs (Asker, Collard-Wexler, and De Loecker 2014), or model mis-specification (Haltiwanger, Kulick, and Syverson 2018), or measurement error (Bils, Klenow, and Ruane 2018; Rotemberg and White 2020). I use dispersion in marginal revenue product of materials which is less susceptible to adjustment costs. Alternatively, I also consider a policy of equating markups across firms. Markup dispersion, unlike dispersion in productivity, is less likely to be driven by unobserved heterogeneity. Similar to Sraer and Thesmar (2020) and Bau and Matray (2020), I also rely on using natural experiment to address issues associated with measurement error and adjustment costs.

to be complete. However, once I account for incomplete pass-through across firms estimated from the data, the reallocation gains are 15 percent. This substantial decrease in productivity gains arises because firms endogenously adjust markups in response to policies enacted to lower their markups in the first place. Part of this endogenous markup adjustment arises from high markup firms facing less-price elastic consumers. I show that the contribution of heterogeneous consumer demand is large — holding demand parameters faced by firms fixed to their estimated levels, reallocation gains are about 33 percent if all firms operated in the most competitive environment observed in the data. This implies that the demand-based markup channel lowers productivity gains from reallocation by 30 percent (14 percentage points).

Contribution to the literature. These findings relate to two distinct, yet related, literatures. First, a recent and important empirical literature shows that markups vary systematically across firms and that they respond to changes in firms' operating environment. In particular, markups are high for exporters (De Loecker and Warzynski 2012; Atkin, Chaudhry, Chaudhry, Khandelwal, and Verhoogen 2015), low for entering firms (Foster, Haltiwanger, and Syverson 2008), and decrease in response to an increase in trade-induced competition (De Loecker, Goldberg, Khandelwal, and Pavcnik 2016). Relative to these papers, I empirically document how firm heterogeneity interacts with differences in consumer preferences and their demand elasticities to generate systematic markup dispersion and endogenous misallocation.[3]

By assessing the role of variable markups for misallocation losses, I also relate to the work on variable markups by Peters (2020); Edmond, Midrigan, and Xu (2019); Haltiwanger, Kulick, and Syverson (2018); Dhingra and Morrow (2019); Behrens, Mion, Murata, and Suedekum (2020). Unlike Peters (2020), this paper studies variable markups driven by differences in consumer preferences (i.e. demand). The demand-based source of variable markups is, in spirit, similar to Edmond, Midrigan, and Xu (2019) with few important distinctions. First, while Edmond, Midrigan, and Xu (2019) rely on a representative consumer with Kimball (1995) demand, this paper documents the important role of consumer heterogeneity. Second, while they quantify the aggregate welfare losses from markups, the focus of this paper is primarily on implications of markup dispersion for allocative efficiency. Third, the approach in this paper is fundamentally different from theirs, and can be applied to data from other settings and countries with relatively minor assumptions on

[3]Recent work by Faber and Fally (2020) have documented similar patterns on assortative matching in the US using retail scanner data. There are few important differences between their work and this paper. First, unlike Faber and Fally (2020) which uses retail-level data on households consumption basket to compute markups with assumptions on market conduct, I use firm-level production data to estimate markup by relying on 'cost-side' approach used in recent literature. This allows me to back out the true markup distribution without making assumptions on nature of competition, or on consumer demand. Second, I provide an identification strategy to isolate the role of demand composition for markup dispersion. Third, while Faber and Fally (2020) assess the impact of trade liberalization policies on real income inequality, I study the importance of demand-driven markup dispersion for allocative efficiency among manufacturing firms.

the demand-side as well as nature of market structure faced by firms. This approach is similar to the one employed in Haltiwanger, Kulick, and Syverson (2018). However, relative to their work, I provide a sufficient statistic approach to estimate productivity gains from reallocation. I also provide a micro-foundation for variable elasticity of demand preferences arising from differences in the composition of demand faced by firms.

To the best of my knowledge, this is the first paper to provide an analytical methodology that relies on pass-through rate to adjust for bias associated in estimating gains from reallocation under variable markups. The pass-through rate has been the focus of Nakamura and Zerom (2010); Goldberg and Hellerstein (2012); Atkin and Donaldson (2015); Bergquist and Dinerstein (2020) which emphasize the extent of exchange-rate, or other cost pass-through and its implications for estimating the market power. I instead apply similar logic to assess the sources behind market power for each firm, with the goal to infer how these sources can generate biases in potential gains from resource reallocation. In doing so, this paper proposes a methodology to separately estimate demand and supply factors behind incomplete pass-through. An advantage of the methodology is that it is based on a reduced-form approach rather than structural estimation.

Second, following the seminal work by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), an extensive literature has focused on factors driving misallocation (Baqaee and Farhi 2019b; David and Venkateswaran 2019). These papers, however, treat these firm-specific factors as exogenous distortions. I depart from this literature in two ways. First, I provide an endogenous source behind misallocation arising from firms' pricing decisions. Second, I assess misallocation losses from variable markups and highlight the importance of studying endogenous markup responses to policies aimed at improving allocative efficiency.[4] Given that pass-through rates have generally been found to be lower than one across multiple settings,[5] failure to adjust for markup responses generates an upward bias in estimated losses from misallocation. This provides a potential explanation for the observation that productivity losses from misallocation reported using the indirect approach — that commonly relies on exogenous wedges — are typically larger than the losses reported using the direct approach — that relies on studying responses to specific enacted policies (Restuccia and Rogerson 2017).

The rest of the paper is organized as follows. Section 2 provides the empirical framework,

[4] By doing this I relate to Rodrik (1987) who argues that when distortions are endogenous, the design of appropriate policy will need to focus on the specific mechanism generating that distortion. As Rodrik (1987) points out: a policy broadly aimed towards addressing observed market wedges will not work because different mechanisms — even if they lead to similar divergence between social and private costs — could behave differently to policies.

[5] For example, in the coffee industry, Nakamura and Zerom (2010) find a pass-through elasticity of commodity prices of 25 percent. Similarly, in beer industry, Goldberg and Hellerstein (2012) find a pass-through of only 5 percent of an exchange rate change into final prices of traded goods. More relevant to this paper, using data on US manufacturing firms, Ganapati, Shapiro, and Walker (2020) find a average pass-through of 70 percent and Haltiwanger, Kulick, and Syverson (2018) find an average pass-through rate of 50 percent across 11 industries.

and describes the methodology to estimate markups and marginal costs from the firm-product-level production data. Section 3 describes the data. Section 4 presents stylized facts on markups and costs variation across firms, and on variation in price elasticities across consumer income groups. Section 5 describes the empirical strategy to isolate the role of consumer demand behind markup variation, and presents the results. This endogenous markup dispersion matters, in turn, for understanding the productivity losses from resource misallocation. Section 6 assess the gains from a reallocation policy under various scenarios, and Section 7 concludes.

2 Empirical Framework

In this section, I derive a general expression for markups from firms' profit maximization problem. I show that the level of firm markup depends on both demand- and supply-side factors, and relate markups to resource misallocation across firms. I then describe the methodology to compute markups and marginal costs from the firm-level production data.

2.1 Model

Consider a firm i that uses a variable input factor X to produce output Q_i. The price of the input is $W_i^X = W^X \, \tau_i^X$, where W^X is the market price for input and $\tau_i^X > 0$ is an input distortion faced by firms on the price of that input. The output tax is modeled as markup μ_i charged by the firm. The firm's profit function is given by:

$$\Pi_i = P_i Q_i - MC(\tau_i^X, W^X, \Omega_i, Q_i) = \left[P_i - MC(\tau_i^X, W^X, \Omega_i)\right] Q_i$$

where $Q_i = \Omega_i F_i(X_i)$, X_i is the input demand, Ω_i is firms' exogenously given physical productivity, and $F_i(.)$ is firm's production function which exhibits constant returns to scale.[6] A profit maximizing firms will choose quantity to equate its marginal revenue to marginal costs MC_i:

$$\underbrace{\frac{\partial (P_i Q_i)}{\partial Q_i}}_{\text{Marginal Revenue}} = \frac{\partial P(Q_d)}{\partial Q_i} Q_i + P_i = MC_i$$

This yields the following expression of markups over marginal costs (P_i/MC_i):

$$\mu_i = \left(1 + \frac{\varphi_i}{N_d} \frac{\partial \log P_i}{\partial \log Q_d}\right)^{-1} \tag{1}$$

[6] The assumption of constant returns to scale technology is made for simplicity here. In estimation of markups and marginal costs, I allow for flexible returns to scale. As shown in Table A.1 in Appendix Section A.2, the estimates from the data show that firms across most sectors indeed closely exhibit constant returns to scale.

where Q_d is the total quantity of the good in the district d, N_d is the number of firms supplying that good in the district. $\varphi_i \equiv \frac{\partial Q_d}{\partial Q_i}$ is the supply conduct parameter and has the following interpretation: $\varphi_i = 0$ under perfect competition; $\varphi_i = 1$ under Cournot case ; and $\varphi_i = N_d$ in a collusive environment. I follow the 'conduct parameter' approach to model strategic interactions among firms from Atkin and Donaldson (2015) and define $\Phi_i \equiv \frac{N_d}{\varphi_i}$. This allows me to use a single parameter to proxy for competition, instead of separately identifying N_d and φ_i. The competitiveness parameter can be summarized under different models of competition: $\Phi_i \to \infty$ under perfect competition; $\Phi_i = N_d$ under Cournot case of monopolistic behavior; and $\Phi_i = 1$ in collusive environment. Markups expression 1 can be rewritten as:

$$\mu_i = \left(1 + \frac{1}{\Phi_i}\frac{\partial \log P_i}{\partial \log Q_d}\right)^{-1} \tag{2}$$

It is clear from equation 2 that markups are a function of the slope of the demand $\left(\sigma_i \equiv \frac{\partial \log Q_d}{\partial \log P_i}\right)$ and supply conduct parameter Φ_i faced by firms. That is, $\mu_i \equiv \mu(\Phi_i, \sigma_i)$.[7]

Marginal revenue product. A cost-minimizing firm selects X such that it equates it's marginal revenue product of input (MRPX) to its marginal cost.[8] The main assumption is that at least one input is variable, that is the input is chosen in the same time period as it is used. This rules out presence of adjustment costs for that input or inventories. This gives:

$$\text{MRPX}_i \equiv P_i \frac{\partial Q_i(.)}{\partial X_i} = \mu_i\, \tau_i^X\, W^X \tag{3}$$

The marginal revenue product for the input X is directly proportional to both markups μ_i (output wedge) and τ_i^X (input wedge). Thus markup dispersion generates resource misallocation: firms with higher markups μ_i will have higher marginal revenue product of input and demand lower X than their efficient input demand.

While markup variation causes misallocation, establishing the sources behind markup dispersion matters for assessing productivity losses from misallocation. When markups are exogenous, they are isomorphic to input distortions and providing subsidies to high marginal product firms

[7]A similar expression holds for markups if I assume a more general model of Bertrand competition (Nash in price) with differentiated products. Under symmetric model of monopolistic competition, markups over marginal costs are defined as $\mu_i \equiv P_i/MC_i = [1 + (\partial \log P_i/\partial \log Q_i)]^{-1}$. The second term in the bracket the inverse of firms' individual residual demand elasticity, which depends on the numbers of firms and elasticity of demand faced by the firm, firms' own- and cross-price elasticities, and degree of product differentiation. The estimates of markups, however, is not prone to assumption on the nature of underlying competition, or consumer demand.

[8]Under a cost-minimization objective firms choose inputs to minimize their costs, given a certain level of output. In presence of one variable and static input, changes in the demand of that input will be accompanied by changes in output. Therefore, cost-minimization will yield similar first-order conditions as obtained through profit-maximization where firms choose the output level that maximizes their current variable profits.

would counteract the effect of these wedges by reallocating production factors across firms and increasing aggregate productivity. However, if markup dispersion is driven by firms facing heterogeneous price elasticities of demand, then the underlying distortions are endogenous and generated by decisions of optimizing agents to begin with. The price sensitivity of consumers, which determines firms' elasticity of demand through σ_i, is guided by consumer preferences and is less addressable by policy because it is not susceptible to reallocation. Gains from reallocation would therefore be smaller, and their magnitude will depend on the extent of markup dispersion caused by demand factors. On the other hand, supply-side factors, which affect firms' markups through Φ_i, could be addressed by policy.

2.2 Estimating markups and marginal costs

Rearranging expression 3 provides the following expression for markups μ_{ij} for firm i producing product j:

$$\mu_{ij} \equiv \frac{P_{ij}}{MC_{ij}} = \underbrace{\left(\frac{\partial \log Q_{ij}}{\partial \log X_{ij}}\right)}_{\theta_{ij}^X \text{(Output Elasticity)}} \Big/ \underbrace{\left(\frac{W_{ij}^X \,.\, X_{ij}}{P_{ij} \,.\, Q_{ij}}\right)}_{\alpha_{ij}^X \text{(Expenditure Share)}} \tag{4}$$

where P denotes the price of output Q, W^X denotes the price of variable input X, MC is the marginal production cost, θ^X is the output elasticity with respect to the variable input, and α^X is the expenditure on that variable input as share of firm's revenue. The expression follows from the cost-minimization approach developed by Jan De Loecker in his various contributions (De Loecker and Warzynski 2012), technical details for which are provided in Appendix A.2. Because more than half of the firms in my data produce more than one product, I specifically follow De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) to estimate markups at firm-product level. I use material inputs as the variable input in production to compute the output elasticity.[9] Once the markups are estimated for all firm-products, I can obtain marginal costs using information on firm-product prices from the data and dividing it by corresponding markups. Finally, to avoid any potential effect of outliers for our results, I trim the sample at 5^{th} and 95^{th} percentiles of the markup distribution. Appendix Table A.2 shows the average and median markup by sector.

There are two main benefits of the cost-minimization approach to estimate markups. First, it allows to measure firms' markups without having to take a stand on many aspects of the theory such

[9]In principle, one could use labor as the variable input as in De Loecker and Warzynski (2012). However, labor markets in India are highly regulated and impose substantial firing costs on firms (Aghion, Burgess, Redding, and Zilibotti 2008). Using labor as flexible input, therefore, imposes an strong assumption on firms' adjustment costs. As an alternative, I have used electricity as the variable input in production because the ASI data also reports expenditure on electricity incurred by firms. The correlation between estimates obtained from electricity and material inputs as variable input measure is 80 percent. For the main analysis, I rely on the estimates obtained using material inputs because data on electricity expenses is missing for about 20 percent of the observations.

as imposing parametric assumptions on consumer demand, or the underlying nature of competition, or assumptions on the returns to scale. Equation 2 showed that markups depends on the market supply-conduct and underlying consumer demand, both of which are unobserved. The cost-minimization approach allows to estimate markups from production data without knowledge on any of these parameters. Second, this estimation procedure allows me to overcome two biases in markup estimates relative to existing work. First, due to data limitations, procedures that uses revenue-based measure of productivity estimation typically rely on industry-level price deflators. This leads to measurement error when firms produce differentiated products, can price differentiate or have market power. This is the 'output price bias' as described in (De Loecker, Goldberg, Khandelwal, and Pavcnik 2016). I use physical output instead of revenue which solves the output price bias. Second, unobserved differences in input quality across firms and over time could generate bias in productivity estimation by inducing an 'input price bias' (De Loecker, Goldberg, Khandelwal, and Pavcnik 2016; de Roux, Eslava, Franco, and Verhoogen 2020). I address the input price bias by adding as controls prices for input factors (wages and materials) and output to the production function. This controls for the unobserved variation in input quality by using information on output prices, with the intuition that input and output prices contain information on both output and input quality (Kugler and Verhoogen 2011).[10]

3 Data

1. Firm-level data. The primary data used in this analysis is Indian plant panel-data, the Annual Survey of Industries (ASI) maintained by the Ministry of Statistics. The basic unit of observation in the ASI is an establishment. I use the data from 1998 to 2009 that contain both consistent product level information and establishment location information during these years.[11] The sample frame for the survey is all manufacturing establishments in India that employ more than 10 workers. Establishments with more than 100 workers ("census" establishments) are surveyed every year, while smaller establishments are randomly sampled each year. The data contains establishment-level identifiers across years for both census and non-census establishments, allowing me to construct panel data for both types of establishments.[12] I match the establishment-level panel data to a separate ASI cross-sectional data previously maintained by the Ministry, allowing me to obtain the district in which the establishment is located.[13] The ASI allows owners who have more

[10] See De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) for a formal model and detailed discussion on the input price bias. Their methodology is also summarized in Appendix A.2.

[11] The ASI uses accounting year which runs from April 1 to March 31. I refer to each accounting year based on the start of the period; for example, the year I call "2000" runs from April 1, 2000 to March 31, 2001.

[12] See Martin, Nataraj, and Harrison (2017) for more details on the ASI data.

[13] A district is an administrative unit in India, with an average of 17 districts per state. A district is comparable to US county in size. On average, a district has approximately 2 million total residents.

than one establishment in the same state and industry to provide a joint return, but less than 5 percent of my sample do so, and the analysis is conducted at the level of the establishment. I treat each establishment as a separate firm but the results of the paper hold when I explicitly allow for only single-establishment firms.[14] I limit my analysis to domestic firms by excluding the firms that report non-zero share of their sales exported.

A key advantage of the ASI data is that it provides information on factory-gate wholesale prices for the reporting firms. ASI data enables me to track firm's product mix over time because Indian firms are required by the 1956 Companies Act to disclose product-level information on capacities, production, and sales in their annual reports. Product-level information is available for 80 percent manufacturing firms, which collectively account for more than 90 percent of labor force for the ASI manufacturing firms.

Firms report products in the ASI survey using ASI Commodity Classification (ASICC) codes which is the most refined level of product available in the data.[15] Table E.1 reports the basic summary statistics by two-digit NIC (industrial classification system for India) sector. Firms in ASI not only report total sales, but also report sales and quantity sold broken down by product. As the product definition is available at highly disaggregated level, unit values are interpreted as prices. I use this information to define per-unit price as (Total Sales Value)/(Total Quantity Sold).

2. Other data. Consumption data for households are from Indian National Sample Survey (NSS) conducted between years 1998 and 2009. The survey records total household expenditure and quantity bought by households across 256 product categories, which I use to construct per-unit prices at the household-level. The survey is a nationally representative repeated cross-sectional sample of about 500,000 households with sampling weights provided at the district-level. Weather data collected by the University of Delaware is used to construct a time series of rainfall received across Indian districts since the year 1960. These data are gridded by longitude and latitude lines. In order to match these to districts, I simply use the closest point on the grid to the center of the district and assign that level of rainfall to the district for each year. The agricultural data on district-level cropping patterns, crop prices and crop yields comes from the Ministry of Agriculture.

4 Stylized facts on markup variation and consumer demand

I document four facts consistent with *assortative matching* — that is, the tendency of wealthier consumers to source their consumption from goods produced by larger firms. I show that (1) larger

[14]Therefore, going forward, I use the term firms which will refer to the establishment.

[15]A product group is the most refined category to which a product belongs in the data. Few examples of product category include cotton shirts, wooden chair, black tea, sugar, cotton yarn. While unit of measurement could vary across groups, all products within the same group are measured in the same units.

firms incur higher marginal costs and charge higher markups for their products; (2) the positive relation between firm size, costs and markups is stronger in quality differentiated sectors; (3) richer households consume higher priced products; and (4) price elasticity of demand is decreasing in household income levels.

1. Firm-level facts. Panel (a) of Figure I shows the relation between log marginal costs and log number of employees, the closest proxy in the data for unobserved firm productivity.[16] The relationship controls for district-product-year fixed effects to account for any differences, both observed or unobserved, across regions that might contribute to differences in firm costs. The figure shows that within the same narrow product group and located in the same district, smaller firms incur lower marginal costs than larger firms. Specifically, firms with 10 percent larger labor force have 0.41 percent higher costs (Column 1, Panel (a) of Table I).

The positive relationship between marginal costs and firm-size might seem surprising at a first look. In standard production functions, marginal costs are inversely related to physical efficiency which would imply lower marginal costs for larger firms. However, the underlying assumption in those functional form for costs is of constant input prices across firms. This assumption is not valid when firms produce differentiated goods that will require variation in input quality, and therefore, will be reflected in differences in the input prices (Kugler and Verhoogen 2011). Under this production function, firm's productivity and the input quality are complementarity to each other and marginal costs will increase in firm-size.[17] The evidence in Columns 3-6 of Panel (a) in Table I (non-parametrically presented in Figure E.1) is consistent with such production function: larger firms use higher priced inputs, are more capital intensive, pay higher wages per-unit labor, and have higher physical productivity (TFPQ).

In fact, Panel (c) of Figure I shows that the positive correlation of marginal costs and markups with firm size is stronger in sectors with greater scope of quality variation, proxied using Rauch (1999) classification of product differentiation. Panel (b) of Table I reports these correlations. Column 1 shows that the positive relation between marginal costs and firm size, as well as input prices and firm size, is entirely driven by more differentiated sectors. More importantly, as Columns 3-6 of the table show, the underlying factors of marginal costs driving the different correlations

[16] Appendix Table E.2 shows that these results, as well as the results that follow, are robust if I use firms' total sales or fixed assets as alternate proxies for size. Labor force is my preferred proxy as unlike sales or physical productivity (which is estimated through the data), it does not induce a measurement error in the independent variable that could be correlated with estimated markups and marginal costs. The positive relationship between the size of firms' labor force and its productivity is documented in Bartelsman, Haltiwanger, and Scarpetta (2013).

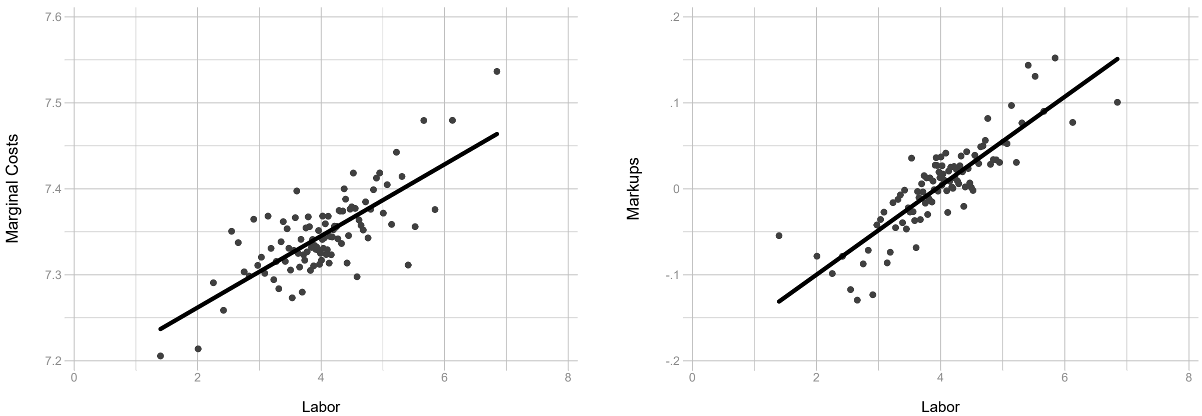
[17] The intuition for these differences in relationship of marginal costs with firm size across homogeneous and differentiated sectors is provided in Appendix A.1. The section shows that marginal costs are log-separable into physical productivity and input prices under the assumption that the underlying production function exhibits constant returns to scale (CRS) — a feature supported by the data, where the average sum of factor input shares is close to one (see Appendix Table A.1). It is important to note that the assumption of CRS is made only for the ease of interpreting how firm size relates to marginal costs. The assumption is not a part of the estimation exercise.

with size across the two sectors are input prices (which reflect input quality) and not the differences in the distribution of physical productivity (TFPQ) across the two sectors.

Figure I: Firms' markups, marginal costs and size

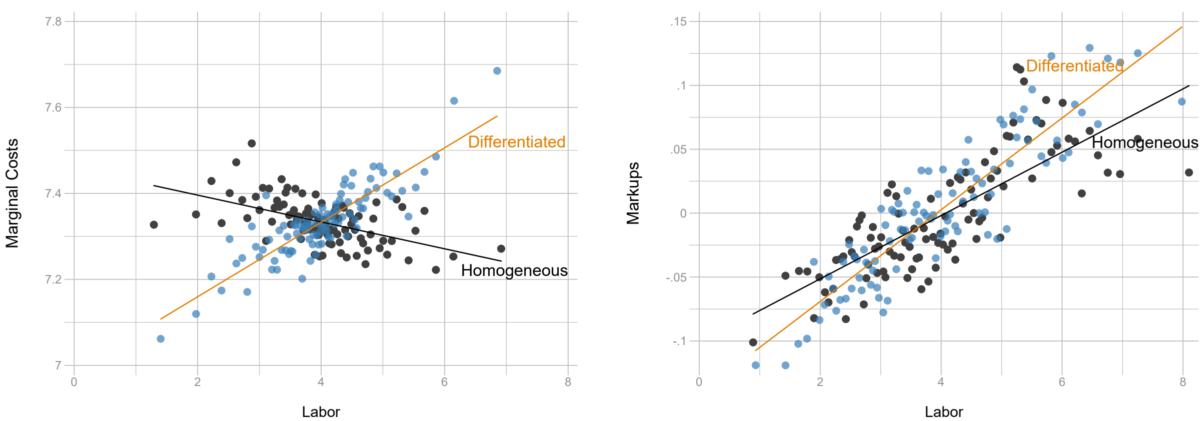
Average

(a) log marginal costs (b) log markups



By quality differentiation

(c) log marginal costs (d) log markups



All variables are measured in logs. The figure shows the relation between firm's per-unit markups, marginal costs and labor force. The top panels shows the average relation by firm-size, and the bottom panel shows the relation by quality-differentiation using the definition in Rauch (1999). The specification controls for district-by-product-by-year fixed effects. Each dot represents 1% of observations.

Next, panel (b) of Figure I documents the central findings of the paper: larger firms also charge higher markups for their products. As before, the values on both axes are after controlling for district-product-year fixed effects. This ensures that I am not comparing markups across regions which might differ along unobserved consumer characteristics or market structure (Anderson,

Table I: Baseline Correlations: Firm-size, markups and costs

	Dependent variable: log of ...					
	Marg.Costs	Markups	Input Price	K/L	Wages	TFPQ
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel (a). Average					
(log) labor	0.041***	0.056***	0.063***	0.098***	0.189***	0.150***
	[0.009]	[0.007]	[0.013]	[0.017]	[0.008]	[0.005]
	Panel (b). By quality differentiation					
(log) labor	-0.023*	0.077***	0.051***	0.073***	0.184***	0.155**
	[0.013]	[0.003]	[0.007]	[0.020]	[0.010]	[0.066]
(log) labor ×	0.117***	0.009**	0.019**	0.046**	0.008	-0.010
1(different. good)	[0.016]	[0.004]	[0.008]	[0.020]	[0.009]	[0.074]
Observations	167,221	167,221	443,022	167,221	167,221	167,221
R-squared	0.870	0.638	0.410	0.656	0.803	0.458
Industry f.e.	✓	✓	✓	✓	✓	✓
District-prod.-year f.e.	✓	✓	✓	✓	✓	✓

All variables are measured in logs. The estimates in Panel (a) are from the specification: $\log y_{ijt} = \alpha_k + \alpha_{djt} + \beta \ \log(\text{labor})_{it} + u_{ijt}$, where y_{ijt} is the variable of interest for product j produced by firm i belonging to industry k located in district d in year t. The estimates in Panel (b) are from the specification: $\log y_{ijt} = \alpha_k + \alpha_{djt} + \beta_1 \ \log(\text{labor})_{it} + \beta_2 \ [\log(labor)_{it} \times 1(\text{different. good})_j] + u_{ijt}$. 1(different. good) is a dummy equal to 1 if a product is classified as differentiated. Standard errors clustered at district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Rebelo, and Wong 2018). Panel (d) of Figure I shows that the positive correlation of markups with firm size is stronger in sectors with greater quality differentiation. Table I summarizes these correlations. Firms with 10 percent larger labor force charge 0.56 percent higher markups (Column 2, Panel (a)). Column 2 in the bottom panel of Table I shows that positive relationship between firm size and markups is stronger in more differentiated sectors.[18]

2. *Household-level facts.* Panel (a) of Figure II documents the relationship between per-unit price for a manufactured good consumed by households and their income. The estimates are after controlling for region-by-product fixed effects, where region is either a town or village and is finer geographical unit than a district, which allows to compare price differences within the same product group (e.g., clothes) for households located in a narrow geographical region. I also include controls

[18]In Appendix B, I conduct multiple tests that provide evidence inconsistent with measurement error as a potential driver of these correlations.

of households' primary occupation, size, religion and social group which absorbs any observable differences across households that might affect their consumption choices. The evidence shows that wealthier households consume higher-priced products within a narrow definition of a product group.

Next, I estimate the price elasticity of demand across income groups. I rely on the demand curve expression $\log Q_{hgi} = -\sigma_{hgi} \log P_i$ for good i by household h in income group g. Because NSS does not provide with a panel data on households, I estimate price elasticities at the income group level. This implies that all households within an income group, $\mathcal{H}_g$, have same price elasticity (that is, $\sigma_{hgi} = \sigma_{gi} \; \forall \; h \in \mathcal{H}_g$). Next, for simplicity and technical limitations, I make an assumption that the *relative* price elasticity across income groups is same over the product space. That is, $\sigma_{gi}/\sigma_{g'i} = \sigma_{gi'}/\sigma_{g'i'} \; \forall \; g, g' \in \mathcal{G}$ and $\forall \; i, i' \in \mathcal{I}$, where $\mathcal{G}$ is the set of all income groups and $\mathcal{I}$ is the set of all products that households consume from. The assumption implies while household income groups could differ in their price elasticities across products, the ratio of this difference is same across all income income groups.

Assumption 1. *The relative price elasticity of demand across income groups is same for all product groups.*

Comparing two product varieties i and k gives the following relation between their expenditure (E_{hi}, E_{hk}), and their prices (P_i, P_k):

$$\log \frac{E_{hgi}}{E_{hgk}} = (1 - \sigma_g) \left[\log \frac{P_i}{P_k} \right]$$

where $E_{hgi} = P_i Q_{hgi}$. Under Assumption 1, σ_g can be estimated from the above expression by only considering goods without any quality differentiation.[19] The above equation can be taken to the data using:

$$\log \left(\frac{E_{ihgrt}}{E_{khgrt}} \right) = \alpha_{gir} + \beta_g \log \left(\frac{P_{irt}}{P_{krt}} \right) + \nu_{ihgrt} \tag{5}$$

where r is the region (i.e. town or village) and t is the year of survey. E_{ihgrt} is the household expense on particular product sold at price P_{irt}. As $\beta_g = 1 - \sigma_g$, the above specification provides an estimate of elasticity by income group σ_g. I benchmark $k = 0$ with the most frequent commodity consumed in a region.

The OLS estimate of σ_g will be potentially biased because unobserved taste shocks in the error term could be correlated with price changes. I follow Faber and Fally (2020) and address

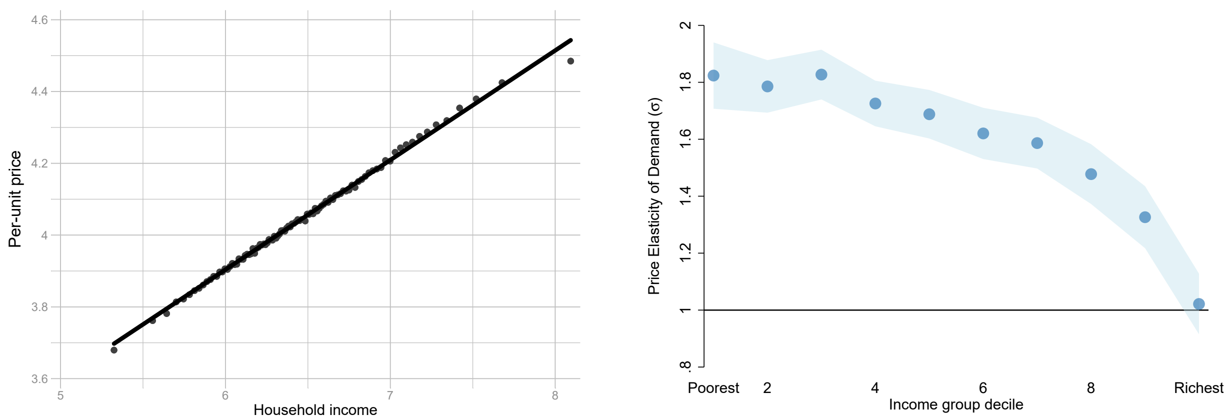
[19]This information is sourced from NSS Consumption Survey and includes products recorded under "grains" category. It includes quantities and prices for rice, wheat, jowar, bajra, maize, barley, and small millets. This is important because — as quality could be correlated with both demand and prices — estimating σ_g based on products with differences in quality could generate bias in the estimation for differentiated products.

this issue by instrumenting local prices $\Delta \log P_{irt}$ with state-level leave-out mean price changes $\frac{1}{N-1}\sum_{k\neq i}\Delta \log P_{krt}$. The instrument identifies the local average treatment effects where the complier group of the instrument will be local and regional sellers for the products. Panel (b) of Figure II shows the estimates of price elasticity of demand across 10 income groups. The price elasticity is similar and high for the lower three income groups and then gradually decreases with income levels before reaching unity for the richest income group. Table E.3 conducts the above exercise parametrically: the price elasticity of demand of the lowest income group quintile is 1.7 times higher than that of the richest quintile.

Figure II: Household income, product prices and price elasticity of demand

(a) per-unit product prices

(b) price elasticity of demand (σ)



Panel (a) shows the relation between log per-unit prices for manufactured goods paid by households and log household income (as proxied by total consumption). The specification controls for product-by-region-by-year fixed effects and household controls (industry of occupation, type of occupation, size religion and social group). Each dot represents 1% of observations. Panel (b) reports the estimates of price-elasticity of demand across income groups (σ_g) based on the estimating equation 5. The estimates are based on a IV-2SLS specification that instruments price of a good with state-level leave out mean price changes. 95% confidence intervals are represented by shaded blue area. Bold circles indicate results that are significant at the 10% level, and hollow circles statistically insignificant from 0 at the 10% level. Source: NSS.

In sum, the evidence presented in this section strongly rejects constant markups across firms. Markups are increasing in firm-size and this relationship is stronger in sectors that are differentiable in product quality. Results from the household consumption baskets show that wealthier households have lower price elasticity of demand. They also consume higher priced goods than poor households within the same product category which suggests they source a higher share of their consumption from larger firms. This evidence is consistent with a *demand-based* markup channel: producing better quality and selling to wealthier, less demand elastic households lead larger firms to incur higher costs and charge higher markups.

Discussion. Before proceeding forward, two important points relative to the existing work on variable markups are worth noting. First, it is the combination of being able to obtain both marginal costs and markups that provides support for the demand-based markup channel. While the observation that markups are increasing in firm-size is independently made — either through direct or indirect evidence — in recent work (Zhelobodko, Kokovin, Parenti, and Thisse 2012; Dhingra and Morrow 2019; Edmond, Midrigan, and Xu 2019), the relation between costs and firm-size is negative. The results above show that while these models fit well the correlations documented in the homogeneous sector, they are inconsistent with the relations in the differentiated sector. Similarly, in supply-side models of variable markups such as Atkeson and Burstein (2008); Edmond, Midrigan, and Xu (2015) where consumers have CES preferences and firms compete in imperfectly competitive environment, markups are higher for larger firms. The results on price elasticity of demand across consumers and positive relation between cost and firm-size both do not fit well these models. Appendix C discusses these (and other) alternative models of firm-heterogeneity, and Appendix Table C.1 compares the correlation between firm-size, markups and costs as made across these frameworks.

Second, and to the best of my knowledge, the work on variable markups does not incorporate preference heterogeneity on the demand-side. Most theoretical work on variable elasticity of demand that is able to generate positive association between markup and size assumes an underlying demand structure with a representative agent á la Kimball (1995) and Klenow and Willis (2016). By its nature, it abstracts away from consumer heterogeneity. This paper provides a micro-foundation for variable elasticity of demand that arises, in part, due to differences in consumer preferences. This suggests that ignoring such consumer heterogeneity as source of variable markups can have distributional consequences across the income distribution.

5 Isolating the role of consumer demand for markup variation

The equilibrium relationship between markups and firm size documented in Section 4 does not identify the causal effect of demand composition. Equation 2 suggests that larger firms could charge higher markups in equilibrium because they have larger market shares or because they face variable elasticity of demand. Moreover, variable elasticity of demand could arise from sources other than firms' demand composition.[20] To isolate the role of demand composition for markups, I propose a research design that uses quasi-exogenous changes to consumer demand across the income distribution. These changes in demand affect the demand composition of firms differently because consumers across income levels differ in the shares of their consumption basket sourced from large-, mid-, and small-sized firms. I then study how firms change their markups in response

[20] For example, see theoretical work by Edmond, Midrigan, and Xu (2019); Dhingra and Morrow (2019); Behrens, Mion, Murata, and Suedekum (2020).

to changes in their demand composition.[21]

5.1 Empirical Strategy

The objective is to understand how firms adjust their markups in response to changes their demand. The equilibrium relation between price P_{ijt} and quantity Q_{ijt} for firm i and product j is given by:

$$\log P_{ijt} = \alpha_0 + \alpha_1 \log Q_{ijt} + \nu_{ijt}$$

Using the identity $\log P_{ijt} = \log \mu_{ijt} + \log MC_{ijt}$, the above relation can be rewritten as:

$$\log \mu_{ijt} = \alpha_0 + \alpha_1 \log Q_{ijt} + (\nu_{ijt} - \log MC_{ijt}) \tag{6}$$

Estimating 6 using ordinary least squares (OLS) methods could lead to biased estimates of α_1. Any correlation between markup and quantities will not identify the causal effect of demand on markups because of (i) reverse causality: higher priced products (the ones with higher quality) could observe an increase in their demand, that is causality might run from markups to quantities; (ii) omitted variable bias: changes along the demand curve, i.e. changes to marginal costs of production, could change firms' markups and therefore the demand Q_{it}; and (iii) measurement error: estimates could be mechanically negative as prices are calculated as product revenue divided by its quantity sold. A solution to this is to obtain an exogenous demand shifter to Q_{ijt} that is unlikely to be correlated with the firms' marginal cost and the market structure. I propose one such instrument for changes in firm's demand: changes to consumer income due to local rainfall fluctuations.

Similar to many other developing countries, majority of the poor in India are employed in the agricultural sector. About 66 percent of males and 82 percent of females in rural India report agriculture as their principal economic activity (Mahajan and Gupta 2011).[22] More than two-third of farmed area in India is rain-fed; and thus agricultural production and rural income are considerably dependent on rainfall. Rainfall exhibits significant variation across districts and over years, generating income changes for poor households in those districts. More importantly, these weather-induced changes to income are *transitory* in nature and lack any persistence across years even within districts (Table E.6). These income changes over years affect the demand for firms that cater more to the poor households than firms that cater less to them. To see this, notice that quantity Q_{hijt} demanded by a income group h over time t is a function of the prices P_{ijt}, the price

[21] See Appendix D for a formal derivation of these predictions. The section presents a demand-based model of variable markups by linking differences in expenditure across the consumer-income distribution to firm-size distribution through product quality.

[22] The relationship between agricultural employment and income levels across districts is evident from Figure E.2. The figure shows that average income in the district is systematically decreasing in its share of population employed in the agricultural sector.

index P_{ht} faced by the group, income I_{ht} for that group, and other factors ν_{hijt}:

$$Q_{hijt} \equiv D(P_{ijt}, P_{ht}, I_{ht}, \nu_{hijt}) \tag{7}$$

The derivative with response to the third argument is $D_3 > 0$ which implies an exogenous income shifter for group h will increase the demand Q_{hijt} from that group.[23] The aggregate demand for firm i is the sum of its total sales to each consumer group, i.e. $Q_{ijt} = \sum_h Q_{hijt}$. Assortative matching dictates that firms across the size distribution — more so in the quality differentiated sector — differ in their share of sales made to different income groups, and therefore, differences in income changes across consumer groups affect the demand Q_{ijt} for firms with varied intensities.[24]

To estimate how rain shocks affect firm outcomes, I run the following specification:

$$\log y_{ijt} = \beta \text{ Shock}_{dt} + \alpha_{ij} + \alpha_{jt} + \gamma \tilde{X}_{ijt} + \epsilon_{ijt} \tag{8}$$

where y_{ijt} is the year t outcome of interest (demand, quantity sold, costs, and markups) for product j produced by firm i located in district d. Shock_{dt} are local rain shocks as defined below. As products produced by different firms could differ across various characteristics, I include firm-product fixed effects α_{ij} which absorbs any time-invariant firm-product unobservables (for example, any constant quality differences). The presence of product-year fixed effects α_{jt} controls for product-specific inflation and any macro-economic shock at the product level. $\tilde{X}_{ijt}$ are set of firm and market level controls described as they are used in Section 5.2. The reduced form coefficient β in the specification is straightforward to interpret as the elasticity of the response of firm-product level outcomes to rain shocks.

Following the non-linear relationship between local rainfall deviations in a year and agricultural yields in Figure E.3, I define a positive shock if the annual rainfall measure is above the 70^{th} percentile and negative shock as rainfall measure below the 30^{th} percentile within the district. The positive and negative shocks should not be taken in an absolute sense as I am not comparing districts that usually receive higher rainfall to those that usually receive lower rainfall. This measure simply captures high or low-rainfall years for each district during 1960-2009. For the analysis, I define rain shock as equal to +1 for positive shock, -1 for negative shock, and 0 otherwise. This definition is similar to the one employed in Jayachandran (2006). The mean value of the rain shock measure

[23] For the ease of exposition, I have abstracted away from presence of household savings. In reality, it is possible that households might smooth their consumption by saving more in response to transitory income shocks. However, as discussed later in the identification assumptions, the evidence on marginal propensity to consume in response to transitory income changes directly refutes this possibility.

[24] Using weather driven income changes has an additional advantage over other measures of local income changes that could be driven by changes in aggregate price levels (for e.g., industry level wage growth). To see this, we can decompose $\Delta \log I_{ht}$ into a function of aggregate prices $f(P_{ht})$ and a residual variation independent of prices ϵ^I_{ht}: $\Delta \log I_{ht} = f(P_{ht}) + \epsilon^I_{ht}$. Rain shocks have the advantage of affecting the residual variation ϵ^I_{ht}.

is -0.14 with standard deviation of 0.78. Columns 1-2 of Table E.4 show the effect of rain shocks on local agricultural outcomes: positive rain shocks increase crop yields in the district by 5 percent and revenue by 3.5 percent.[25]

Identification Assumptions. Consistent estimation of β in specification 8 requires two conditions to be satisfied: relevance of rain shocks, that is, Shock_{dt} and $\log Q_{ijt}$ should be correlated; and exclusion restriction, that is, Shock_{dt} is uncorrelated with ϵ_{ijt}. Relevance can be directly tested in the data — local rainfall deviation should be strongly correlated with the local income and the quantity demanded for poor households. Two results lend strong support to the hypothesis that rain shocks change the *relative* demand of the poor households, and affect the demand disproportionately across the firm-size distribution.

First, I show that rain shocks affect the wages of population employed in agriculture. Column 3 of Table E.4 shows the effect of rain shocks on incomes of the poor: daily wages in agricultural sector increase by 2.7 percent. Rain shocks do not affect wages for households employed outside agricultural sector or for non-rural labor force (Columns 4 and 5). Next, I document that poor households have higher marginal propensity to consume (MPC) out of temporary income changes. Figure E.5 reports the distribution of MPC across the income distribution.[26] For same increases in income (and conditional on prices), quantity demanded increases more for the poor population. Taken together, these results provide strong support that rain shocks generate significant variations in demand for the poor households.

Second, I check how rain shocks affect firms' idiosyncratic demand. Following the influential work by Foster, Haltiwanger, and Syverson (2008), I use the production data and obtain firm's idiosyncratic demand by isolating total quantity movement from quantity movement due to a change in supply-side change in prices. Specifically, I estimate firm-product level demand-shifters η_{ijt} using:

$$\log Q_{ijt} = \gamma \log P_{ijt} + \alpha_{jt} + \eta_{ijt} \tag{9}$$

α_{jt} absorbs yearly changes at the product-level (in both supply and demand), and η_{ijt} are firm-product demand shifters. Estimating equation 9 could lead to positive bias in estimates of price elasticity γ, because firms could respond to demand shifters η_{ijt} by increasing prices. I overcome this problem by using changes to marginal cost as instrumental variables (IV) for supply-side price

[25] I run the following specification: $y_{dct} = \beta \times \text{Shock}_{dt} + \alpha_{dc} + \alpha_{ct} + \epsilon_{dct}$, where the outcome variable is either average yield (output per hectare) or revenue for crop c across fifteen major crops in India and α_{dc} and α_{ct} are the district-crop and crop-year fixed effects.

[26] I follow Gruber (1997) and calculate the MPC using the observed drop in consumption upon unemployment. Using a monthly panel data on 100,000 households from CMIE household consumption data, I estimate the following regression for household h in town v at month t: $\Delta \log E_{hgvt} = \alpha_g \, \Delta \log I_{hgvt} + \beta_h + \beta_{vt} + \epsilon_{hvt}$, where β_h is the household fixed effect, β_{vt} is a town-year fixed effect that captures the total resources available in the town-month and aggregate shocks in month t, and g is the income group. As the regression is run on a panel data at household-month level, the coefficient α_g is identified of the variation in *within* household income across months.

shifters. Marginal costs incorporate firms' idiosyncratic cost-shifters through changes in their input prices and firm's technology (see Appendix A.1).[27] Thus, it has explanatory power over firms' prices which are unlikely correlated with short-run changes in demand.

The demand estimates γ from specification 9 are shown in Table E.5. The first column provides the OLS estimates, on average and across each of the industries, and the last two columns provides the results using firm-products' marginal costs as IV for their prices. I find that average estimated IV elasticities are negative, and range from -4.5 to -1.9 across the industries. The IV estimates of elasticities are also 2 to 4 times more elastic than OLS estimates, consistent with the upward bias due to simultaneity in the OLS estimates.

I use the residuals η_{ijt} from the demand function estimation to provide evidence on the relevance of rain shocks for firm-level demand. Table II reports the correlation of demand shocks η_{ijt} with rainfall shocks (using specification 8). Column 1 shows that firms' estimated idiosyncratic demand increases by 1.2 percent during years of positive rain shocks. Similar result is obtained if I instead use quantity sold by firms as a direct measure of firms' demand (Column 2). Next, I estimate the effects of rain shocks on firms' demand across quartiles of firm-size distribution using:

$$\log y_{ijt} = \sum_{r=1}^{4} \beta^r . \left(\text{Shock}_{dt} \times T_i^r\right) + \alpha_{ij} + \alpha_{jt} + \gamma \tilde{X}_{ijt} + \epsilon_{ijt} \tag{10}$$

where $r \in \{1, 4\}$ indexes each of the four quartiles of the size distribution and T_i^r are dummy variables taking the value of 1 when firm i belongs to quartile r.[28] Panel (a) and (b) of Figure III shows the effects of rain shocks on firms' demand across the size distribution. The effects are strongest for smallest firms and for firms in the middle of the size distribution, and gradually decrease to zero for the largest firms.

The second identification condition that rain shocks should satisfy is exclusion restriction. That is, rain shocks should affect markups only through changes to demand curve faced by firms. While this assumption cannot be directly tested, I believe the richness of the production data allows me to test whether rain shocks might affect firms' supply curve. As mentioned before, observing prices at firm-product level provides estimates of marginal costs along with markups, allowing to test whether (and how) rain shocks affect marginal costs across firms. Columns 3-6 of Table II report the correlations of rain shocks with firms' marginal costs and its underlying components. I

[27] Relative to Foster, Haltiwanger, and Syverson (2008), I use marginal costs instead of physical productivity (TFPQ) as an instrument for prices. Unlike TFPQ which is estimated at firm-level, marginal costs are estimated at firm-product level and provides greater variation. See Appendix A.1 for the discussion of log-separability of marginal costs into TFPQ and input prices.

[28] I define these quartile using firm size (using firm's first occurrence in the panel) based on its labor force relative to two-digit industry average. Using 2-digit industries instead of products increases the number of observations within each quartile and reduce the noise associated with misclassfication.

Table II: Effect of rain shocks on firms' idiosyncratic demand and marginal costs

	Dependent variable:					
	Demand Shifter	log quantity	log of ... marg. cost	TFPQ	wage	input price
	(1)	(2)	(3)	(4)	(5)	(6)
Shock$_{dt}$ (-1/0/+1)	0.012**	0.014***	-0.004	-0.011	0.001	0.001
	[0.005]	[0.005]	[0.007]	[0.008]	[0.002]	[0.008]
Observations	133,094	133,094	133,094	59,965	102,541	239,100
R-squared	0.898	0.975	0.952	0.887	0.922	0.931
Firm f.e.				✓	✓	
Firm-product f.e.	✓	✓	✓			✓
Product-year f.e.	✓	✓	✓	✓	✓	✓

All dependent variables but Column 1 are in logs. Shock$_{dt}$ is defined as +1(-1) if the rainfall in the monsoon months is above(below) the 70th(30th) percentile of the district's usual distribution for monsoon rainfall. It takes the value of 0 if the rainfall is between 30th-70th percentile of district's usual distribution. Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

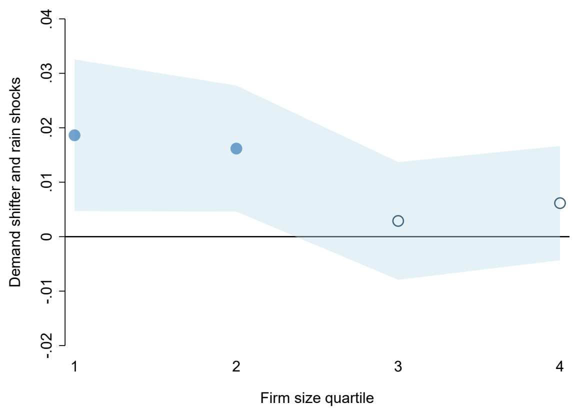
do not find any evidence that rainfall shocks affect marginal costs on average, or firms' physical productivity (TFPQ), wages and prices of material inputs. Figure III Panel (c) shows that rain shocks do not affect marginal costs across the firm-size distribution, and Figure E.4 shows that there is no affect of rain shocks on TFPQ, wages, input prices, and fixed capital across small, medium and large firms.[29]

While rain shocks do not affect marginal costs on average, they could still have a non-zero effect on costs for some firms. For example, an increase in demand could affect costs through changes in X-inefficiencies for few firms and not others. These changes in firms' costs could have an independent supply-side effect on markups and generate bias in estimated β. Therefore, I control for marginal costs in specification 8 in order to isolate markup responses due to changes in demand from rain shocks. This addresses any omitted variable bias by absorbing any component in the error term that might be correlated with both markup changes and quantity produced.

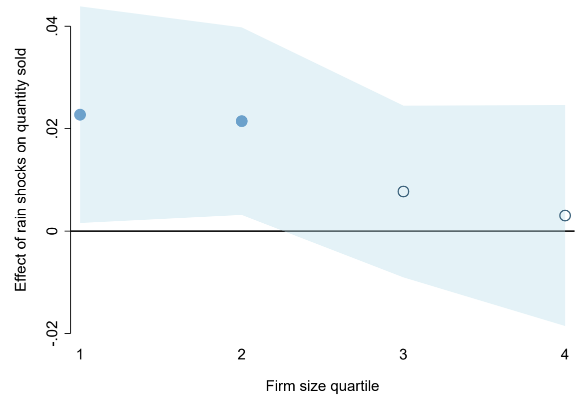
[29] In addition to the above identification assumptions, a separate assumption that needs to be satisfied is that an increase in income for poor households could should not decrease their long-run price elasticity. While this is not in violation of exclusion restriction, this could independently affect markups. For example, poor households can become less price-elastic if higher income in the current year due to better rainfall is predictive of higher income in the future years. As discussed later in section 5.2, a decrease in demand elasticity in years of positive rainfall shock should led to a increase in markups. However, I find that markups decrease in years of positive rain shocks and thus this mechanism should bias, if anything, the estimates towards zero. I also test for serial correlation of rainfall within districts because serially correlated rainfall shocks could induce permanent shifts in the price-elasticity of demand. Table E.6 shows an absence of any serial correlation in rain shocks.

Figure III: Effect of rain shocks on demand and costs (across firm-size distribution)

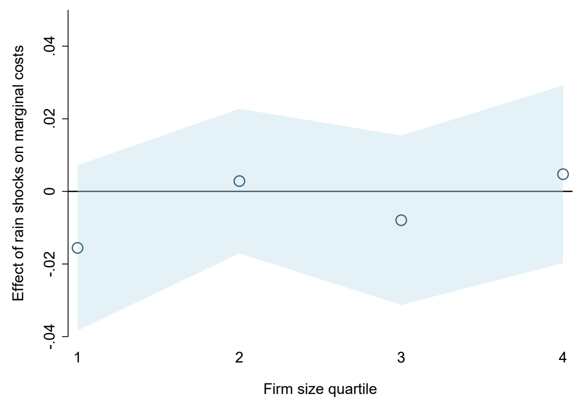
(a) effect on demand shifter η_{ijt}



(b) effect on quantities sold Q_{ijt}



(c) effect on marginal costs mc_{ijt}



The figure shows the estimates of the effect of rain shocks on firms' demand and marginal costs across the firm-size distribution based on specification 10. All specifications control for firm age and size quartile-year fixed effects. 95% confidence intervals are represented by shaded blue area. Bold circles indicate results that are significant at the 10% level, and hollow circles statistically insignificant from 0 at the 10% level.

5.2 Results

Table III presents the main results on how average markups respond to rain shocks. Column 1 shows that firms lower their markups by 0.5 percent in years of positive rain shocks. In Columns 2-8, I show that the results are robust to inclusion of various controls. Because multi-plant firms might be less responsive to local shocks, I restrict the analysis to only single plant firms in Column 2. In Column 3, I include controls for firms' age to allow for markup changes as the firm grows (Peters 2020). In Column 4, I include controls for firms' size quartile and its interaction with year fixed effects to allow for differences in aggregate shocks across size groups. Column 5 includes controls for past two-years of rain shocks in the district to allow for any effects from lagged changes in demand. In Column 6, I control for market access measure constructed from Allen and Atkin

(2016), which is a weighted average rainfall deviation for each district d' connected to district d , where the weights are proportional to the distance between the two districts. Column 7 controls separately for an in-state and an out-state market access measure to allow for separate impact based on whether other districts d' are in the same state as district d or outside the state. Finally, in Column 8 I allow for combined effect of controls from Columns 2-7. As can be seen, addition of these controls has no significant effect on the estimate of average effects of rain shocks on markups.[30]

Table III: Average effect of rain shocks on firms' markups

	Dependent variable: log markup							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shock$_{dt}$ (-1/0/+1)	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005**	-0.004*	-0.004*
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Observations	133,094	122,828	133,094	133,094	133,094	133,094	133,094	133,094
R-squared	0.989	0.990	0.989	0.989	0.989	0.989	0.989	0.989
Firm-product f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Product-year f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Controls	Baseline Specification	Single-plant firms	+ Age control	+ Size-year control	Past 2-year shocks controls	National Market access control	In + out-state market access	(3)-(7) controls

The table reports the average effects of rain shocks on markups, based on specification 8. Shock$_{dt}$ is as defined in the text. All columns include firm-product, product-year fixed effects and control for log marginal costs. Standard errors clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Mechanism. Next, I provide evidence supporting the role of consumer heterogeneity in driving markup variation. To see this, notice that the markup elasticity to firm i's idiosyncratic demand shocks η_{it} is given by:

$$\underbrace{\frac{d\log\mu_{it}}{d\log\eta_{it}}}_{\text{markup change}} = \frac{1}{\tilde{\sigma}_{it}-1} \times \underbrace{\left[-\frac{d\log\tilde{\sigma}_{it}}{d\log Q_{it}}\right]}_{\substack{\text{change due to } \Delta \text{ slope} \\ \text{(slope effect)}}} \times \underbrace{\left[\frac{d\log Q_{it}}{d\log\eta_{it}}\right]}_{\substack{\text{change due to } \Delta Q \\ \text{(size effect)}}} \tag{11}$$

Equation 11 shows that the necessary condition for firms to change its markups in response to a change in demand is that a shift in firms' demand (change due to Δ Q, that is, the *size effect*) is accompanied by shift in the slope of firms' residual demand curve (change due to Δ slope, that is,

[30] I have used alternative definitions of rain shocks, the results of which are documented in Table E.7. The significant relationship of markups and the null effects on marginal costs remain robust to various specifications of rain shocks. I also analyze the effects of rain-induced local demand shocks on exporters. Markups for exporters are a function of the demand that they face in export markets, rather than the local demand. Therefore, exporters should largely be unaffected by the demand changes from rain shocks. However, if rain shocks were indeed common supply shocks to firms, we would expect then to affect firm costs. Table E.8 shows that neither markups or marginal costs are affected by local rain shocks for exporters.

the *slope effect*). A change in slope for residual demand could occur due to changes in demand composition, or due to change in market structure (for e.g., from a change in competition).[31]

So how does one separate the sources behind markup changes? If markup responses to demand shifts are due to changes in demand composition, then an increase in demand from the poor households increases the demand elasticity *only* for firms that sell to both rich and poor households, forcing them to lower their markups. Under assortative matching, these firms are proxied in my data by firms in the middle of the size distribution. Therefore, changes to demand composition, and hence markups, should be strongest for firms in the middle of the size distribution. Smallest firms cater largely to poor — and therefore, homogeneous — consumer base, implying that rain shocks should not affect their demand composition (and markups). The consumers for largest firms are rich households, and therefore, rain shocks do not affect the demand for these firms. That is, in the context of equation 11, the slope effect for smallest firms and the size effect for largest firms are zero which implies that markups should not change for these firms. More importantly, this non-monotonic pattern should only be present in quality-differentiated sectors.

If instead the reduction in markups from higher demand is due to an increase in competition among firms, then the effects should be strongest for the smallest firms which observe the largest increase in demand (but no significant changes to their demand composition). Table C.1 provides a comparison of how markups will change across the firm-size distribution under different models of variable markups proposed in the literature. The table shows that it is only under the demand composition channel, that markup responses to demand shocks will be non-monotonic across the firm-size distribution.

To test this prediction, I estimate the effect of rain shock on markups across quartiles of firm-size distribution using specification 10, estimation results for which are plotted in panel (a) of Figure IV. As the figure shows, rain shocks only affect markups in the middle of the size distribution. The estimates are reported in Table E.9. The coefficient of -0.7 to -0.9 percent and -0.5 to -0.8 percent in the second and third quartile, respectively, of the size distribution is more than two to three times larger than the lowest quartile (which are insignificant across all specifications). Firms in the largest size quartile also do not change their markups. The estimates remain stable after inclusion of various controls from Table III (Columns 2-8).[32]

[31] The intuition for this is as follows. Under a model of monopolistic competition with differentiated products, the equilibrium markup for any firm is given by $\left(\frac{\sigma}{\sigma-1}\right)$, where σ is the *firm-level* demand elasticity. Demand for each firms' product is the sum of demands from rich and poor consumers. The firm-level elasticity is a weighted average of the poor and rich households, with weights equal to the group's share in total demand of the firm. For firms' pricing, the only difference among the different types of consumers is that poor households have a higher price elasticity than rich households (i.e., $\sigma_{poor} > \sigma_{rich}$). Assuming that σ_{poor} and σ_{rich} remains same in response to a *transitory* income shock, an increase in demand from poor household increases the demand elasticity for firms. As a result, firms will lower their markups.

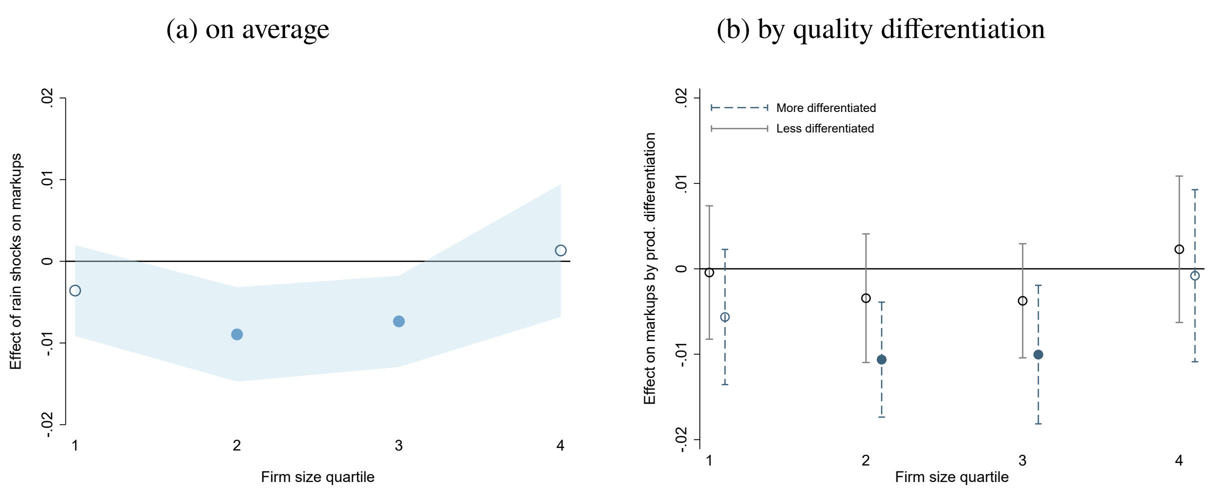
[32] In Panel (a) of Figure E.6, I conduct falsification tests using rain shocks realized in the next year rather than the current year. Markups are not responsive to these placebo shocks. I also do not find any evidence in support of past

Next, I test whether firms producing more differentiated goods change their markups more by estimating the following specification:

$$\log \mu_{ijt} = \sum_{p \in \{0,1\}} \sum_{r=1}^{4} \beta_p^r \, . \, (\text{Shock}_{dt} \times T_i^r \times (Z_{ij} = p)) + \alpha_{ij} + \alpha_{jt} + \Gamma' X_{ijt} + \epsilon_{ijt} \tag{12}$$

where Z_{ij} takes the value of 1 for firms in sectors with greater scope for quality differentiation. If differences in taste over quality are driving the assortative matching, we should observe that the non-monotonic pattern of markup responses should be more prominent in more differentiated sectors. Panel (b) of Figure IV shows that the results are consistent with this interpretation.

Figure IV: Effect of rain shocks on markups across firm-size distribution



Panel (a) shows the estimates of effects of rain shocks on markups across the firm-size distribution. Panel (b) shows the estimates of the effect of rain shocks on markups across the firm-size distribution by scope of quality-differentiation, using specification 12. All specification includes firm-product, product-year fixed effects and controls for firm age, size quartile-year fixed effects, and log marginal costs. 95% confidence intervals are represented by shaded area in panel (a), and by vertical lines in panel (b). Bold circles indicate estimates significant at the 10% level, and hollow circles statistically insignificant from 0 at the 10% level.

5.3 Alternate mechanisms

I consider alternative explanations for lower markups in periods of increased demand. First, under imperfect competition, incumbent firms could decrease markups due to entry of new firms, or introduce new products in response to higher demand. Second, firms might collude in setting markups and the incentives to deviate from such collusive agreements could increase when demand increases. Third, consumers might increase their shopping search intensity when their income

demand responses (observed from last year's rain shocks) having any persistent effects on future markups (Panel (b) of Figure E.6).

increases. Fourth, financially constrained firms could raise markups when facing negative demand shocks. A common distinction between these explanations and the demand-based markup channel is that the observed non-monotonic pattern of markup responses across the firm-size distribution documented above is unique only to the latter. I nevertheless examine each of these explanations separately and find empirical evidence inconsistent with any of them.

Firm entry and exit. Incumbent firms could lower their markup if new firms enter the market during high demand. This endogenous supply-side response to an increased demand increases the competition and exerts downward pressure on markups. I directly test for firm's entry and exit in the data. ASI data reports the year of establishment for firms as well as whether a firm is operational during the survey year. Table E.10 show that there is no evidence of excess entry or exit of firms in response to rain shocks.[33]

New product introduction. Firms might introduce new products in response to higher demand, putting downward pressure on markups for the existing products (Jaravel 2019). Two pieces of evidence suggest that this is unlikely the channel in my setting. First, it is the size of the market, and not the composition of the market, that matters for introduction of new products. Table E.11 rejects the hypothesis: effects of rain shocks on markups does not differ across districts with different *levels* of rural population. While the interaction with districts' *share* of rural population is significant, the estimate for interaction of rain shocks with district's *total* rural population is statistically insignificant. Second, the ASI data records product entry and exit, allowing me to test for this channel directly. Figure E.7 shows no effects on number of products across the firm-size distribution in response to rain shocks.

Collusion. In standard models of firm collusion, it is difficult to sustain collusion when demand changes frequently. This is because temptation to renege from a collusive agreement is higher during periods of temporary increase in demand because the gains from reneging are increasing in current demand but the loss from punishment increases in future (and uncertain) demand. If firms are indeed strategically adjusting their markups to build customer base then markups should decrease only in periods of higher demand. In periods with a drop in demand, however, markups should remain unchanged. The setting allows me to observe markup responses across both positive and negative demand shocks. Figure E.8 confirms that the non-monotonic effects of rain shocks on markups are present across both positive and negative rain shocks. Therefore, the prediction from models of firm collusion does not hold support in the data.

[33] Intuitively, firm entry or exit seems a remote candidate to drive the observed effects. Establishing a new firm requires substantial capital investment, labor hiring and it seems unlikely that firms would incur these large costs given the shifts in consumer demand induced by rain shocks are temporary in nature (Table E.6).

Consumer search. Consumers might increase their search intensity and shop more outlets during periods of high demand, appearing to be more price sensitive to firms. While both increased consumer search and changes in demand composition would affect markups, they emphasize different mechanisms due to which firms would lower markups when demand increases. Under the consumer search channel, time-varying demand elasticity faced by firms is a result of increased search activity. As a result, higher search intensity in periods of increased demand would predict a positive association between a firm's demand elasticity and changes to its demand. In the context of this paper, this implies that smallest firms should see the largest increase in their price elasticity and lower their markups. However, I find that markup responses are only present for firms in the middle of the size distribution.

Financial constraints. Firms facing costly external financing may raise their markups when faced with negative demand shocks (Gilchrist, Schoenle, Sim, and Zakrajsek 2017). In these models, consumers have persistent habit over firms' products. This allows financially distressed firms to increase their markups and increase cash holdings, allowing them to avoid liquidation in the short-run. Two results rule out the financial channel as a potential driver for the results. First, as reported in Table E.12, the estimates of interaction of rain shocks with firm-size are robust when I include as controls the differential effect of rain shocks depending on firm's financial strength, proxied by firm's cash ratio $\left(\frac{\text{Cash}}{\text{Cash + Fixed Assets}}\right)$, and it's financial leverage $\left(\frac{\text{Debt}}{\text{Fixed Assets}}\right)$. Second, as documented previously, markups for smallest firms do not change in response to negative rain shocks. Smaller firms are more likely to have binding financial constraints, and therefore, the results on no effect on markups for these firms is in contrast with a financial constraint channel.

Altogether, the evidence above shows that differences in demand composition across firms — arising from assortative matching on quality between firms and consumers — are necessary to rationalize the patterns of markup dispersion observed in the data. However, these results leave two related questions open. First, they do not indicate how large are the misallocation losses due to variable markups. Second, from these results, no conclusion can be drawn on the quantitative contribution of demand- and supply-side factors for misallocation losses. In the next section, I address both questions by providing an approach to estimate gains from reallocation under variable markups, and quantify the losses arising separately due to demand- and supply-side factors.

6 Aggregate Implications

Markups are a source of misallocation. Dispersion in markups could arise through differences in nature of competition faced by firms (supply-side factors), or due to differences in consumer preferences (demand-side factors), or a combination of both. Consumer preferences, however, are not susceptible to reallocation and the aggregate productivity gains that could be attained from a hypothetical reallocation exercise will, of course, be correspondingly smaller. In the end, any exercise computing reallocation gains is specific to the underlying model or the hypothetical exercise. For example, Hsieh and Klenow (2009) propose dispersion in revenue productivity as a measure of allocative inefficiency. The gains from reallocation in their framework is proportion to the variance of the dispersion in revenue productivity.

I take a different, yet complementary, approach and ask "how much aggregate productivity gains can be achieved if we remove underlying distortions through a tax-subsidy policy?".[34] Such an approach would need two objects to estimate gains from reallocation. First, it requires an expression that relates aggregate productivity gains to the underlying distortions. For this I rely on a first-order approximation for aggregate productivity growth from Petrin and Levinsohn (2012). An advantage of this approach is that it does not impose structure on underlying demand or market structure.[35] Second, it requires a tax-subsidy policy to counteract the underlying distortions. I consider a policy that serves a planner's objective to equalize marginal revenue products (MRP) across firms within industry under a fixed aggregate supply of resources, while (erroneously) assuming that any variation in MRP across firms arises only due to presence of exogenous distortions. The main result from the exercise is that when markup distortions are endogenous, firms could adjust their markups in response to tax-subsidy policies. This substantially lowers the gains from the intended reallocation policy.

6.1 Analytical framework

Aggregate productivity growth. In this subsection, I describe the aggregate productivity growth decomposition from Petrin and Levinsohn (2012). The change in aggregate productivity for a sector s is the difference between changes in output and input costs within that sector:

$$dAP_s = \sum_{i \in I_s} P_i dQ_i - \sum_{i \in I_s} W_i^X dX_i$$

where Q_i is the gross output of firm i, and P_i is firm's price, I_s is the number of firms in sector s. I define the total productivity growth in the economy as the weighted average of sector-level

[34] I consider the policy of optimal subsidy because it provides a set of simple, easy-to-act rules that can be targeted by policies, based solely on divergence between market prices and social costs. As Dixit (1985) notes: "a distortion is best countered by a tax instrument that acts directly on the relevant margin". Once the relevant margin has been traced, a tax-subsidy policy can be imposed to close the gap.

[35] This is especially relevant given that any model-specific exercise for computing reallocation gains will only be as credible as the underlying model of demand and supply, making such an exercise susceptible to model mis-specification (Haltiwanger, Kulick, and Syverson 2018; Asker, Collard-Wexler, and De Loecker 2014) or measurement error (Bils, Klenow, and Ruane 2018; Gollin and Udry 2021).

productivity growth:

$$dAP = \sum_s \gamma_s \, . \, dAP_s$$

where γ_s is the share of total output in the economy coming from sector s.[36] As the expression is similar across all sectors, for convenience I omit the notation s going forward. Setting aside firms' entry and exit, the aggregate productivity growth can be decomposed into a within-firm productivity improvement ("technical efficiency") term and an across-firm allocation ("reallocation") term.[37]

$$\text{APG} = \underbrace{\sum_i \lambda_i d\log\Omega_i}_{\text{APG (within)}} + \underbrace{\sum_i \lambda_i \left(\theta_i^X - \alpha_i^X\right) d\log X_i}_{\text{APG (reallocation)}} \tag{13}$$

where θ_i^X is the output elasticity with respect to the input, α_i^X is the input expenditure as share of firm's revenue, Ω_i is firm's technical efficiency, and $\lambda_i \equiv \frac{P_i Q_i}{\sum_i P_i Q_i}$ is firm's (Domar 1961) weight. The output elasticity θ_i^X is obtained by estimating the production function as described in Section 2. The revenue share of firms' input expenses α_i^X and Domar weights λ_i are obtained directly from the data.

Reallocation policy. Equation 13 shows that reallocation gains are directly related to $d\log X_i$. From Equation 3, the firms input demand X_i is a function of markups μ_i and exogenous distortions τ_i^X (that is, $X_i \equiv X(\mu_i, \tau_i^X)$). The implies that the change in input demand to a tax/subsidy S_i is:

$$d\log X_i \quad = \quad \left[\frac{\partial\log X_i}{\partial\log\mu_i}\frac{\partial\log\mu_i}{\partial\log S_i} + \frac{\partial\log X_i}{\partial\log\tau_i}\frac{\partial\log\tau_i}{\partial\log S_i}\right] d\log S_i \tag{14}$$

I define firms' pass-through rate Γ_i as the elasticity of firm's price to its costs as $\Gamma_i \equiv \frac{\partial\log P_i}{\partial\log S_i} = \left[1 + \frac{\partial\log\mu_i}{\partial\log S_i}\right]$.[38] Substituting for this expression in 14, and with some algebra, yields the following relationship between input demand and subsidy (see Appendix A.3 for a detailed derivation):

$$d\log X_i = -\left[\frac{\Gamma_i}{\theta_i^X - \alpha_i^X}\right] d\log S_i \tag{15}$$

There are two factors that affect resource reallocation across firms in response to a subsidy. First, firms with low-demand elasticities pass-through only a fraction of those subsidies into their prices.

[36] Following the literature, I use 4-digit NIC industry classification to define the sectors. There are 125 sectors in the data.

[37] As described in Section 3, while the ASI data used in this paper provides consistent firm identifiers across years, it only surveys about one-third of firms in consecutive years, making it difficult to identify the contribution of firms' entry and exit to aggregate productivity growth.

[38] Notice that the knowledge of pass-through rate is necessary and sufficient to assess the reallocation gains under variable markups. I do not need information on how market structure or demand faced by firms will change in response to targeted subsidies. Indeed, a combination of those responses is exactly what the pass-through rates will capture.

This fraction is dictated by the pass-through rates for firms. Second, conditional on limited pass-through firms facing low demand elasticities witness less changes in their quantities demanded, and therefore, change their input demand by less.

Equation 15 shows that the reallocation gains can be estimated for a proposed tax-subsidy policy S_i. I obtain such a policy by considering a social planner with the following objective and constraint: (i) the planner equalizes marginal revenue products for inputs (or, alternatively, markups) across firms within a sector; (ii) the planner faces a fixed supply of aggregate factors. These conditions are standard in the static misallocation literature. When dispersion in marginal products is assumed to be exogenous, the expression of tax/subsidy for firm i takes the following form (see Appendix A.3 for details):

$$d \log S_i = \left[\left(\sum_i \left(\frac{\tilde{X}_i}{\sum_i \tilde{X}_i} \right) \, . \, \log \text{MRPX}_i \right) - \log \text{MRPX}_i \right] \tag{16}$$

where $\tilde{X}_i = \left(\frac{X_i}{\theta_i^X - 1} \right)$. Define $w_i^X = \left(\frac{\tilde{X}_i}{\sum_i \tilde{X}_i} \right)$, and imputing the reallocation policy 16 back in expressions 15 and 13 provides with the following expression for reallocation gains under endogenous markups:

$$\text{APG-R}(\Gamma_i) = \sum_i \lambda_i \, \Gamma_i \, \left[\log \text{MRPX}_i - \left(\sum_i w_i^X \, . \, \log \text{MRPX}_i \right) \right] \tag{17}$$

It is clear from equation 17 that under variable markups, potential gains from reallocation will be affected by the pass-through rate. Under well-known case of monopolistic competition and CES demand, pass-through is complete ($\Gamma_i = 1$), and therefore, any reallocation targeted at exogenous wedges will increase aggregate productivity as intended. For example, in Hsieh and Klenow (2009), the expression 17 provides the exact quantification on productivity losses from misallocation when $\Gamma_i = 1$. Under variable markups, however, the pass-through rate is incomplete ($\Gamma_i < 1$). While the larger firms face more distortions and would need a large subsidy, they have lower pass-through rate and will change their prices less relative to their subsidies. This lowers gains from any targeted reallocation policy.

Pass-through and the demand curvature. The final task requires to decompose the gains in reallocation due to demand- and supply-factors. I do it in a parsimonious way and rely on the functional form of pass-through rate proposed in Weyl and Fabinger (2013) and Atkin and Donaldson (2015). I use the following general expression for Γ_i:

$$\Gamma_i = \left[1 + \frac{1 + \delta_i}{\Phi_i} \right]^{-1} \frac{1}{\mu(\Phi_i, \sigma_i)} \equiv \left[1 + \frac{1}{\Phi_i} \, . \, \frac{\chi_i}{\sigma_i} \right]^{-1} \frac{1}{\mu(\Phi_i, \sigma_i)} \tag{18}$$

where $\delta_i = \left[\frac{\partial \log\left(\frac{\partial P_i}{\partial Q_i}\right)}{\partial \log Q_i}\right]$ is the elasticity of the slope of inverse demand curve, and $\mu(\Phi_i, \sigma_i)$ is the markup from equation 2. With some rearranging, $\delta_i = \frac{\chi_i}{\sigma_i} - 1$, where $\chi_i = \left(1 + \frac{\partial \log\left(\frac{\partial Q_i}{\partial P_i}\right)}{\partial \log P_i}\right)$ is the elasticity of slope of demand ("super-elasticity") and σ_i is the demand elasticity. Therefore, the level of pass-through Γ_i depends on (i) elasticity of slope of demand χ_i, (ii) demand elasticity σ_i, and (iii) competitive structure of industry Φ_i. To assess how demand- and supply-factors affect the reallocation gains in 17 through their effects on pass-through rates, one would need to separately observe χ_i, σ_i and Φ_i. However, a primary challenge is that none of these parameters are observable to researchers — indeed, if they were observed one could have used that directly to compute pass-through rates. I provide a strategy to separate out demand factors from competitive factors from firms' estimated markups and pass-through rate. To do so, I make the following assumption:

Assumption 2. *The slope of demand faced by firms in homogeneous goods sector is constant, but can vary across firms in quality differentiated sector due to differences in demand composition faced by firms.*

While Assumption 2 might seem strong, it is rationalized from the results presented in Section 4 and 5: differences in demand (composition) arising from assortative matching between firms and consumers on quality generates provides firm with additional market power in quality-differentiated sector. The assumption implies that firms in homogeneous sector face same slope of demand curve. That is, $\sigma_i = \sigma^{\text{non-diff}}$ in homogeneous sector, and differences in markups across firms in homogeneous sector are thus driven *only* by differences in competitiveness Φ_i. For firms in differentiated sector, however, variation in markups is driven by differences in slope of demand as well as competitive index. Under this assumption, differences in markups in homogeneous and differentiated goods sector can be used to obtain estimates of σ_i and Φ_i for all the firms. I can then use estimated pass-through rates to compute the super-elasticity χ_i for all firms. This allows to estimate reallocation gains under CES versus variable demand, and under different underlying market structure faced by firms.

6.2 From theory to estimation.

Identification of Γ, Φ, σ, and χ. I now describe the methodology to estimate firm-level pass-through rate Γ_i, and its underlying components.

Step 1: Estimate firm-level pass-through rates Γ_i: Firm-level pass-through Γ_i can be estimated using the information on prices, marginal costs and the following relationship:

$$\log P_{ijt} = \Gamma_i \log MC_{ijt} + \alpha_{ij} + \alpha_{jt} + \xi_{ijt} \tag{19}$$

I compute the pass-through rates both by using marginal costs directly and by instrumenting it with estimated quantity productivity (TFPQ). A potential issue in using marginal costs is that it is calculated using prices and markups, and therefore measurement error could generate upward bias in the estimates of Γ_i. Instrumenting the marginal costs with TFPQ addresses this issue. I also show later that the OLS and IV estimates of pass-through are not significantly different from one another, suggesting that the bias is not a primary concern.

Step 2: Recover estimates of competitive index Φ_i for homogeneous sector: Let $\hat{\mu}_i$ denote the estimate of firm markups μ_i derived in Section 4. Assumption 2 implies that any variation in markups in homogeneous sector arises only due to differences in the competitive index for firms (embedded in Φ_i). I can then use equation 2 to obtain the following relationship between markups, the slope of demand and the conduct parameter:

$$-\log\left(\hat{\mu}_{ij}^{-1} - 1\right) = \log \sigma^{\text{non-diff}} + \log \Phi_{ij} \tag{20}$$

where, following Atkeson and Burstein (2008) and Edmond, Midrigan, and Xu (2015), there is one-to-one mapping between firms' competitiveness and relative size of the firm within its industry.[39] This relationship is captured using the following polynomial model:[40] $\log \Phi_{ij} = \boldsymbol{\zeta}_z f(\log z_i) + \alpha_j + \epsilon_{ij}^{\phi}$. Substituting in equation 20 gives the following equation that I take to the data:

$$-\log\left(\hat{\mu}_{ij}^{-1} - 1\right) = \log \sigma^{\text{non-diff}} + \boldsymbol{\zeta}_z f(\tilde{\boldsymbol{z}}_{\boldsymbol{i}}) + \alpha_j + \epsilon_{ij}$$

The first term is identified through the constant in the regression. The second term captures the supply-side pass-through variation. The third term ensures that we compare firms within the same industry. The error term ϵ_{ij} captures the variation in Φ_i that is orthogonal to firms' market shares.

Step 3: Recover estimates of slope of demand: I use the estimates for $\hat{\Phi}_i$ obtained in Step 2 to estimate σ_i for firms in quality-differentiated sector using:

$$\hat{\sigma}_i = \left[\left(\frac{1}{\hat{\mu}_i} - 1\right)\hat{\Phi}_i\right]^{-1} \tag{21}$$

I assume that the relationship between firm-size and competitiveness index in Step 2, i.e. $\hat{\Phi}_i = \hat{\zeta}_{\mathbf{z}} f(\log z_i)$, also follows in the differentiated sector. This allows me to estimate $\hat{\Phi}_i$ in the differentiated sector. Combining $\hat{\Phi}_i$ with estimated markups 21 provides the estimates for slope of demand

[39] The firms' demand elasticity in Atkeson and Burstein (2008) is a linear function of its market shares.

[40] I allow Φ_i to be a flexible function of firm's market share, which I proxy by employment: Specifically, I use the following relationship $\log \Phi_i^{-1} = \zeta_z^1 \log z_i + \zeta_z^2 (\log z_i)^2 + \nu_i$, where z_i is firm i's relative employment in its industry. The estimates are $\zeta_z^1 = -0.12$ (t-stat of -8.5 with errors clustered at firm-level), and $\zeta_z^2 = 0.028$ (t-stat of 4.19).

in differentiated sector. The slope of demand in homogeneous sector is $\sigma^{\text{non-diff}}$.

Step 4: Recover estimates of super-elasticity for all firms: Finally, I use the estimates for $(\hat{\Phi}_i, \hat{\sigma}_i)$ obtained in Step 2 and 3 to estimate χ_i for all firms. Let $\hat{\Gamma}_i$ denote the unbiased estimate of Γ_i obtained from from Step 1. I use the relationship 18 to obtain estimates of χ_i:

$$\hat{\chi}_i = \left(\frac{1}{\hat{\Gamma}_i \hat{\mu}_i} - 1 \right) \hat{\Phi}_i \, \hat{\sigma}_i \tag{22}$$

6.3 Results

I start by documenting the results on pass-through rates, and its underlying components using the methodology described in Section 6.2. I then use the estimated parameters to quantify the aggregate productivity gains from reallocation under variable markups observed in the data, and under counterfactual scenarios with different parameter values of conduct and demand.

Pass-through rates. Table IV shows the results on pass-through estimates from equation 19. The average pass-through rate is 55 percent (OLS estimates, Column 1), and 70 percent (IV estimates, Column 4). Columns 2 and 5 show that larger firms pass-through less of changes in costs into their prices. Column 3 and 6 show that the negative relation between pass-through and firm size is stronger in quality-differentiated sectors.

Figure E.9 shows the estimates of pass-through rate and its underlying components across the firm-size distribution separately for homogeneous and differentiated sectors. I winsorize all estimates at 5%. Panel (a) shows that pass-through rates are decreasing in firm size, and this relationship is stronger in quality differentiated sectors. Panel (b) shows the negative relationship between pass-through rate and firm size also reflects lower competition faced by larger firms. The similarity in the slope of two lines is just mechanical — by construction (Step 3 above) the relation between firm-size and competitiveness index is same across two sectors. Panel (c) shows that larger firms in quality-differentiated sector face higher slope of inverse demand (i.e., less elastic demand curve) relative to smaller firms. As demand composition is not a feature of homogeneous sector, the slope of inverse demand with firm-size is zero for this sector. Finally, Panel (d) shows that the elasticity of slope of demand is increasing in firm-size in the differentiated sector.

Counterfactuals. Next, I use these parameters to estimate aggregate productivity under various scenarios, and assess the role of demand factors in reducing aggregate productivity gains from reallocation. Column 2 of Table V presents the results under planners' objective to equate MRP of material inputs across firms within a sector. I start by calculating gains from reallocation under the natural benchmark scenario of exogenous markups (Scenario (1)). When markups are assumed

Table IV: Estimates of pass-through rates

	Dependent variable: log price$_{ijt}$					
	OLS estimation			IV estimation		
	(1)	(2)	(3)	(4)	(5)	(6)
log mc$_{ijt}$	0.553***	0.678***	0.664***	0.700***	0.739***	0.855***
	[0.006]	[0.015]	[0.025]	[0.008]	[0.023]	[0.036]
log mc$_{ijt}$ $\times$ log labor$_{it}$		-0.026***	-0.019***		-0.009*	-0.014*
		[0.003]	[0.005]		[0.005]	[0.008]
log mc$_{ijt}$ $\times$ 1(diff)$_i$			0.030			-0.166***
			[0.031]			[0.046]
log mc$_{ijt}$ $\times$ log labor$_{it}$ $\times$ 1(diff)$_i$			-0.014**			0.004
			[0.007]			[0.010]
Observations	131,557	131,557	131,557	131,557	131,557	131,557
R-squared	0.408	0.411	0.411			
Kleibergen-Paap F-stat				8392.57	847.44	132.57
Firm-product f.e.	✓	✓	✓	✓	✓	✓
NIC4 - year f.e.	✓	✓	✓	✓	✓	✓

The table reports estimates of pass-through rates from specification 19. Standard errors clustered by firm-level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

to be exogenous wedge, they do not react to underlying environment or policy changes. I plug in $\Gamma_i = 1$ along with the estimates of tax-subsidy policy, weighted-average MRPs and Domar weights in equation 17. The first row shows that assuming markups to be exogenous would give us an estimate of about 47 percent for the aggregate productivity gains from reallocation.

In the second row of Table V, I allow for firms to adjust their markups in response to the tax-subsidy policy (Scenario (2)). As described in Section 6.1, this markup adjustment is captured by firm-level pass-through rates Γ_i. I plug in the estimates of Γ_i from the data in equation 17. As reported in the second row, while the estimated productivity gains from reallocation are still significant and positive (15.3 percent), they are an order of magnitude lower than the benchmark case of complete pass-through. This is because high markup firms are also the firms that have the lowest pass-through of subsidies into their prices.

Next, in Scenario (3), I analyze the role of demand-based markup channel by applying the restriction that all firms face maximum competitiveness, while holding fixed their estimated slope of demand and super-elasticity. Specifically, I allow all firms to face the maximum competition within 4-digit industries every year observed in the data ($\Phi_i = \Phi^{\max}$). I use the estimated $\Phi^{\max}$ and plug it into:

$$\text{APG-R}\left(\chi_i, \sigma_i, \Phi_i^{\max}\right) = \sum_i \lambda_i \, \Gamma_i\left(\chi_i, \sigma_i, \Phi_i^{\max}\right)\left[\log \text{MRPX}_i - \sum w_i^X . \log \text{MRPX}_i\right]$$

Table V: % change in aggregate productivity from reallocation

	Pass-through Γ_i considered:	Reallocation gains from equating	
		MRP inputs	markups
	(1)	(2)	(3)
(1) Complete pass-through	1	46.9%	36.6%
(2) Incomplete pass-through	actual (from data)	15.3%	8.8%
(3) Maximum competitiveness in data + estimated demand	$\Gamma(\chi_i, \sigma_i, \Phi^{\max})$	33.2%	57.0%
(4) Minimum competitiveness in data + estimated demand	$\Gamma(\chi_i, \sigma_i, \Phi^{\min})$	16.9%	16.8%

The table reports gains from a reallocation policy that equates marginal revenue products for materials across firms within 4-digit industries (Column 2), and markups (Column 3). The reallocation gains are calculated by averaging annual gains across the sample period.

The estimate in third row of Table V shows that if all firms faced maximum competitiveness within the sector, aggregate productivity gains are 33.2 percent. While these gains are 13.7 percent point lower than the gains when pass-through is assumed to be completed, they are still order of magnitude higher than the observed pass-through in the data. This implies that the demand-driven markup dispersion lowers the aggregate productivity gains from reallocation by about 30 percent.[41]

Finally, I also created counterfactual Scenario (4) where I allow competitiveness index Φ_i to be the least competitive environment (that is, the minimum competition within 4-digit industries in the data $\Phi_i = \Phi^{\min}$)), while keeping fixed the estimated demand parameters. The results are reported in fourth row in Table V. The aggregate reallocation gains are 16.9 percent which are closer to Scenario (2) that uses observed pass-through rates, suggesting that limited competition faced by firms also generates large misallocation losses. In Column 3 of the table, I conduct the exercise with the objective function of equalizing markups (instead of MRP of inputs) and find similar results.

6.4 Caveats

The exercise above is a partial equilibrium analysis, and as such, comes with few caveats. First, it does not take into account entry and exits of firms. In this sense, the exercise is static in nature. This allows me to concentrate attention on *static misallocation* — and its implication for aggregate

[41] This number is calculated by considering the amount of reduction in reallocation gains (46.9 percent to 33.2 percent) that can be explained by going from Scenario (1) of complete pass-through to Scenario (3) which uses estimated demand but holds competition to its maximum level estimated in the data.

productivity — in a spirit that is closer to much of the existing work. An additional margin through which variable markups reduce aggregate productivity is the selection of firms, as suggested by Dhingra and Morrow (2019); Behrens, Mion, Murata, and Suedekum (2020).

Second, the first-order decomposition above might not be a good approximation to quantify productivity gains if the distortions are large and a subsidy policy can have higher-order effects on productivity growth. However, when trying to understand the impact of enacted policies under endogenous markups, one could sum over first-order approximations of policy effects each year to obtain the non-linear approximation of the effects of policy over a longer time horizon (Baqaee and Farhi 2019a).

Finally, the analysis does not consider effects on consumer welfare. It has instead focused on the aggregate productivity in the manufacturing sector, given the available data. One could study the changes to consumer price indices across the income distribution (derived in Appendix D, equation D.1) to assess consumer welfare consequences. However, such an exercise will require complete information on quality-adjusted product prices in the consumption baskets, which is not available in the NSS data. The methodology recently proposed by Atkin, Faber, Fally, and Gonzalez-Navarro (2020) provides a promising direction to estimate consumer welfare across the income distribution in absence of detailed price information. I leave for future work the application of such techniques to estimate consumer welfare effects from reallocation policies under endogenous markups.

7 Conclusion

There is now an increasing evidence documenting higher markups for larger firms. The empirical evidence on the sources driving this correlation is, however, rather scarce. In this paper, I provide evidence on how demand-side characteristics affect the equilibrium distribution of markups across firms. I also assess the implications of this demand-driven markup dispersion for understanding misallocation losses. My results provide strong support for models that feature heterogeneous demand elasticities across firms, which are able to generate variable markups. However, I go a step further in documenting the interaction of consumer and firm heterogeneity in driving these variable markups. I show that heterogeneity in consumer preferences — that is, differences in their demand elasticities and preferences over quality — across income distribution translates into heterogeneity in markups charged by firms: lower demand elasticity of wealthier households allows larger firms to charge higher markups. While this demand-driven variable markups generate misallocation across firms, the losses from such misallocation are limited.

I use a sufficient statistic, firms' pass-through rate, to correct bias in aggregate reallocation gains under endogenous markup adjustments by firms. I find that pass-through rates are decreasing in firms size, with the relationship stronger in quality-differentiated sector. These differences in pass-through rates are driven by both differences in the slope of demand curve and market structure faced by firms. I propose a methodology — supported by the empirical evidence presented — that uses differences in markups and pass-through rates across homogeneous and differentiated sector to identify how differences in the demand characteristics across firms affect their pass-through rates. The main finding is that gains from reallocation are lower by about 30 percent under demand-driven variable markups than when markups are assumed to be exogenous.

Like much of data available in developing countries, I do not directly observe the characteristics of consumers that buy from firms. Yet this paper shows that inferences on how consumer demand affects firms' prices — and its underlying components — can still be made by combining available production data for firms with natural experiments. With separate data on prices and quantities (rather than revenues), — and despite imposing minimal assumptions on demand or market structure faced by firms — differences in markups and marginal costs across firms and sectors, and how firms change their prices in response to changes in their costs can inform us to a great extent about sources behind firms' market power and how that affects aggregate productivity.

References

Aghion, P., R. Burgess, S. J. Redding, and F. Zilibotti (2008): "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India," *American Economic Review*, 98(4), 1397–1412.

Allen, T., and D. Atkin (2016): "Volatility and the Gains from Trade," *Working Paper*.

Anderson, E., S. Rebelo, and A. Wong (2018): "Markups Across Space and Time," *Working Paper*.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013–1063.

Atkeson, A., and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998–2031.

Atkin, D., A. Chaudhry, S. Chaudhry, A. K. Khandelwal, and E. Verhoogen (2015): "Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan," *American Economic Review*, 105(5), 537–544.

Atkin, D., and D. Donaldson (2015): "Who's Getting Globalized? The Size and Implications of Intranational Trade Costs," *Working Paper*.

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2020): "Measuring Welfare and Inequality with Incomplete Price Information," *Working Paper*.

Baqaee, D. R., and E. Farhi (2019a): "A Short Note on Aggregating Productivity," *Working Paper*.

——— (2019b): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135(1), 105–163.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 103(1), 305–34.

Bau, N., and A. Matray (2020): "Misallocation and Capital Market Integration: Evidence From India," *Working Paper*.

Behrens, K., G. Mion, Y. Murata, and J. Suedekum (2020): "Quantifying the Gap Between Equilibrium and Optimum under Monopolistic Competition," *The Quarterly Journal of Economics*, 135(4), 2299–2360.

Bergquist, L. F., and M. Dinerstein (2020): "Competition and Entry in Agricultural Markets: Experimental Evidence from Kenya," *American Economic Review*, 110(12), 3705–47.

Bils, M., P. J. Klenow, and C. Ruane (2018): "Misallocation or Mismeasurement?," *Working Paper*.

David, J. M., and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109(7), 2531–67.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016): "Prices, Markups, and Trade Reform," *Econometrica*, 84(2), 445–510.

De Loecker, J., and F. Warzynski (2012): "Markups and Firm-Level Export Status," *American Economic Review*, 102(6), 2437–71.

de Roux, N., M. Eslava, S. Franco, and E. Verhoogen (2020): "Estimating Production Functions in Differentiated-Product Industries with Quantity Information and External Instruments," *Working Paper*.

Dhingra, S., and J. Morrow (2019): "Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity," *Journal of Political Economy*, 127(1), 196–232.

Dixit, A. (1985): "Chapter 6 Tax policy in open economies," vol. 1 of *Handbook of Public Economics*, pp. 313–374. Elsevier.

Domar, E. D. (1961): "On the Measurement of Technological Change," *The Economic Journal*, 71(284), 709–729.

Edmond, C., V. Midrigan, and D. Y. Xu (2015): "Competition, Markups, and the Gains from International Trade," *American Economic Review*, 105(10), 3183–3221.

——— (2019): "How Costly Are Markups?," *Working Paper*.

Faber, B., and T. Fally (2020): "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *Working Paper*.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Ganapati, S., J. S. Shapiro, and R. Walker (2020): "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes," *American Economic Journal: Applied Economics*, 12(2), 303–42.

Gilchrist, S., R. Schoenle, J. Sim, and E. Zakrajsek (2017): "Inflation Dynamics during the Financial Crisis," *American Economic Review*, 107(3), 785–823.

Goldberg, P. K., and R. Hellerstein (2012): "A Structural Approach to Identifying the Sources of Local Currency Price Stability," *The Review of Economic Studies*, 80(1), 175–210.

Gollin, D., and C. Udry (2021): "Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture," *Journal of Political Economy*, 129(1), 1–80.

Gruber, J. (1997): "The Consumption Smoothing Benefits of Unemployment Insurance," *The American Economic Review*, 87(1), 192–205.

Haltiwanger, J., R. Kulick, and C. Syverson (2018): "Misallocation Measures: The Distortion That Ate the Residual," *Working Paper*.

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

——— (2014): " The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 129(3), 1035–1084.

Jaravel, X. (2019): "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector," *The Quarterly Journal of Economics*.

Jayachandran, S. (2006): "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries," *Journal of Political Economy*, 114(3), 538–575.

Kimball, M. S. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit, and Banking*, 27(2), 1241–1277.

Klenow, P. J., and J. L. Willis (2016): "Real Rigidities and Nominal Price Changes," *Economica*, 83(331), 443–472.

Kugler, M., and E. Verhoogen (2011): "Prices, Plant Size, and Product Quality," *The Review of Economic Studies*, 79(1), 307–339.

Linder, S. B. (1961): *An Essay on Trade and Transformation*. New York: Wiley & Sons.

Mahajan, V., and R. K. Gupta (2011): *Non Farm Opportunities for Smallholder Agriculture*.

Martin, L. A., S. Nataraj, and A. E. Harrison (2017): "In with the Big, Out with the Small: Removing Small-Scale Reservations in India," *American Economic Review*, 107(2), 354–86.

Nakamura, E., and D. Zerom (2010): "Accounting for Incomplete Pass-Through," *The Review of Economic Studies*, 77(3), 1192–1230.

PETERS, M. (2020): "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 88(5), 2037–2073.

PETRIN, A., AND J. LEVINSOHN (2012): "Measuring aggregate productivity growth using plant-level data," *The RAND Journal of Economics*, 43(4), 705–725.

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

RESTUCCIA, D., AND R. ROGERSON (2008): "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 11(4), 707 – 720.

——— (2017): "The Causes and Costs of Misallocation," *Journal of Economic Perspectives*, 31(3), 151–74.

RODRIK, D. (1987): "Policy Targeting with Endogenous Distortions: Theory of Optimum Subsidy Revisited," *The Quarterly Journal of Economics*, 102(4), 903–911.

ROTEMBERG, M., AND T. K. WHITE (2020): "Measuring Cross-Country Differences in Misallocation," *Working Paper*.

SRAER, D., AND D. THESMAR (2020): "How to Use Natural Experiments to Estimate Misallocation," *Working Paper*.

WEYL, E. G., AND M. FABINGER (2013): "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, 121(3), 528–583.

ZHELOBODKO, E., S. KOKOVIN, M. PARENTI, AND J.-F. THISSE (2012): "Monopolistic Competition: Beyond the Constant Elasticity of Substitution," *Econometrica*, 80(6), 2765–2784.

Online Appendix for "Demand For Quality, Variable Markups and Misallocation: Evidence from India"

A Technical Appendix (For Online Publication)

A.1 Log-separability of marginal costs in physical productivity and input prices

This section shows that marginal costs (in logs) is additively separable into physical productivity (Ω) referred to as TFPQ, and a function of input prices $\phi(\boldsymbol{W})$ under the following assumptions: (i) the production function is Hicks-neutral and it is homogeneous of degree one, that is, it exhibits constant return to scale (CRS), (ii) firms take input prices as given. Assumption (i) of CRS technology can be verified in the data and the results in the next section A.2 (Table A.1) show that firms across most sectors indeed exhibit returns to scale of one. Under these assumptions, the production function takes the following general form: $Q = \Omega \, . \, F(\boldsymbol{X})$, where Q is firms' output (measured in physical quantities) and $\boldsymbol{X}$ is a vector of inputs. The first-order condition for a cost-minimizing firm with respect to inputs X^m and X^n with prices W^m and W^n is:

$$\frac{F_{X^m}(\boldsymbol{X})}{F_{X^n}(\boldsymbol{X})} = \psi\left(\frac{X^n}{X^m}\right) = \frac{W^m}{W^n} \tag{A.1}$$

where $\psi(.)$ is an increasing function. The first equality in the above expression following from Assumption (i) of homogeneity of the production function. Because all homogeneous functions are homothetic, $\frac{F_{X^m}(.)}{F_{X^n}(.)}$ is an increasing function of $\left(\frac{X^n}{X^m}\right)$. Defining $\Psi = \psi^{-1}(.)$, one can invert the second inequality to obtain the input demand function:

$$X^n = \Psi\left(\frac{W^m}{W^n}\right) X^m \tag{A.2}$$

Substituting the input demand in $Q = \Omega \, . \, F(\boldsymbol{X})$, and using the assumption of degree one homogeneity of production function provides:

$$\frac{Q}{\Omega} = X^m F(\boldsymbol{\gamma_i W}) \Rightarrow X^m(\boldsymbol{W}) = \frac{Q}{\Omega} g_m(\boldsymbol{W}) \tag{A.3}$$

where $g_m(.)$ is a function of input prices. Replacing the input demand in cost function $C(\boldsymbol{X}, \boldsymbol{W}) = \sum_m X^n W^m$ yields:

$$C(Q, \boldsymbol{W}) = \frac{Q}{\Omega} \sum_m g_m(\boldsymbol{W}) W^m \tag{A.4}$$

Define $\phi(\boldsymbol{W}) = \sum_m g_m(\boldsymbol{W})W^n$. Finally, this provides us with the following functional form of marginal costs:

$$MC(\boldsymbol{W}) = \frac{C(Q, \boldsymbol{W})}{Q} = \frac{1}{\Omega}\phi(\boldsymbol{W}) \tag{A.5}$$

Taking logs of equation A.5, and defining $\rho(X, Y)$ as the covariance between variables X and Y, and $var(X)$ as variance of X to arrive at the following relationship:

$$\frac{\rho(\log mc(\boldsymbol{W}), \log \Omega)}{var(\log \Omega)} = -1 + \frac{\rho(\log \phi(\boldsymbol{W}), \log \Omega)}{var(\log \Omega)} \tag{A.6}$$

Because firms are price-takers in input market, $\rho(\log \Omega, \log \phi(\boldsymbol{W})) = 0$ in homogeneous sector. This implies that marginal costs are decreasing in firm-productivity in this sector. For firms in differentiated sector, the relation between marginal costs and productivity depends on whether there are complementarities between the physical productivity and input quality (which is reflected positively in the input prices). When these complementarities are large enough — for example in Kugler and Verhoogen (2011); Bastos, Silva, and Verhoogen (2018) — marginal costs can increase in firm productivity.[42]

A.2 Estimating markups and marginal costs

This section provides broad overview of the procedure for estimating markups in the Indian manufacturing data. The framework primarily builds on the methodology in De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) for estimating markups for multi-product firms.

Framework. Consider a production function for firm i and product j in year t: $Q_{ijt} = \Omega_{it} \,.\, F_{jt}(\boldsymbol{X}_{ijt}, \boldsymbol{K}_{ijt})$, where $\{\boldsymbol{X}_{ijt}, \boldsymbol{K}_{ijt}\}$ is a vector of variable and static inputs X and capital stock K which is assumed to be dynamic. The presence of a variable and static input X implies that it is chosen in the same time period it is used and only affects current profits. The latter rules out inventories or adjustment costs.

The cost-minimization problem for firms yields following expression for markups at firm-

[42] For general expression for correlation of marginal costs with firm size S, expression A.6 changes to $\rho(\log mc(\boldsymbol{W}), \log S)/var(\log S) = -\rho(\log \Omega, \log S)/var(\log S) + \rho(\log \phi(\boldsymbol{W}), \log S)/var(\log S)$. Therefore, the same argument presented above applies whenever more productive firms are also larger in their size $\rho(\log \Omega, \log S) > 0$. In the main analysis, I have used firms' labor force — instead of physical productivity Ω — as a proxy for firm size (S). The choice of labor force as a proxy for firm-size is intentional because unlike sales or productivity productivity (which is estimated through data), the use of labor force is not susceptible to measurement error in the independent variable that could be correlated with estimated marginal costs. Appendix Table E.2 shows that these results are robust if I use firms' total sales or fixed assets as alternate proxies for size.

product level:

$$\mu_{ijt} = \theta^X_{ijt} \left(\alpha^X_{ijt}\right)^{-1}, \quad \text{where} \quad \underbrace{\theta^X_{ijt} = \frac{\partial \log Q_{ijt}(.)}{\partial \log X_{ijt}}}_{\substack{\text{output elasticity} \\ \text{(obtained by estimating } F_{jt}(.))}}, \qquad \underbrace{\alpha^X_{ijt} = \frac{W^X_{ijt} X_{ijt}}{P_{ijt} Q_{ijt}}}_{\substack{\text{share of input expenditure} \\ \text{(only observe for single product firms)}}}$$

In addition to the fact that θ^X_{ijt} needs to be estimated, the main challenge when working with multi-product firms is that α^X_{ijt} is not observed in the data at the firm-product-level. I eventually estimate markup using:

$$\hat{\mu}_{ijt} = \hat{\theta}^X_{ijt} \cdot \frac{P_{ijt} Q_{ijt}}{\exp(\hat{\rho}_{ijt}) \tilde{W}^X_{it}} \tag{A.7}$$

where ρ_{ijt} is the share of input expenditure attributable to product j. The data provides information on $(P_{ijt}, Q_{ijt}, \tilde{W}^X_{it})$, where $\tilde{W}^X_{it}$ is the total expenditure on variable input X. The estimation of $\hat{\mu}_{ijt}$ involves estimation of production function parameter θ^X_{ijt} and the input allocation $\hat{\rho}_{ijt}$.

Estimation. Taking the logarithm of the production function gives:

$$q_{ijt} = f_{jt}(x_{ijt}, k_{ijt}; \boldsymbol{\theta}) + \omega_{it} \equiv f_{jt}(\boldsymbol{z}_{ijt}; \boldsymbol{\theta}) + \omega_{it} \tag{A.8}$$

where notation in small caps denote logarithms of corresponding large cap variables. Thus, any changes to output over time occurs due to either (i) changes in input quantities or (ii) unanticipated shocks to productivity. Here $\boldsymbol{z}_{ijt} = \{x_{ijt}, k_{ijt}\}$ is a vector of (log) physical inputs and $\omega_{it} = \log(\Omega_{it})$. The production coefficients $\boldsymbol{\theta}$ need to be identified.

Three biases arise in the estimation of the production function. First, output price-bias could arise when output is constructed by deflating firm revenues by an industry-level price index. A difference from existing work that relied on data on revenue is that here q_{ijt} is in physical units of output. This solves the output-price bias. Second, because I only observe input expenditure, and not input quantities, I need to modify the above expression with the use of input expenditure: $z_{ijt} = \rho_{ijt} + \tilde{z}_{it} - w^z_{ijt}$, using $W^Z_{ijt} Z_{ijt} = \tilde{\rho}_{ijt} \left[\sum_p W^Z_{ijt} Z_{ijt}\right] = \tilde{\rho}_{ijt} \tilde{Z}_{it}$. Here $\tilde{z}_{it}$ is the firm-level expenditure on input Z and w^Z_{ijt} is the deviation of the unobserved (log) firm-product-specific price from the (log) industry-wide input price index. This yields the following decomposition of expression A.8:

$$q_{ijt} = f_j(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}; \boldsymbol{\theta}) + \underbrace{a(\rho_{ijt}, \tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \boldsymbol{\beta})}_{\text{Input Allocation Bias}} + \underbrace{b(\boldsymbol{w}_{ijt}, \rho_{ijt}, \tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \boldsymbol{\theta})}_{\text{Input Prices Bias}} + \omega_{it}$$

The objective is to address the two sources of biases: "Input Allocation Bias" and "Input Prices Bias". I now discuss the steps involved in addressing these biases, and the estimation of production function and input allocation.

Addressing input allocation bias. I address input allocation bias by focusing on single product firms. For these firms, $\rho_{ijt} = 1$ and hence $a(.) = 0$. That is for single-product firms, the true production function will not suffer from input allocation bias. I also drop sub-script j due to its redundancy for single-product firms:

$$q_{it} = f(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}; \theta) + B(\boldsymbol{w}_{\boldsymbol{it}}, \tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \boldsymbol{\theta}) + \omega_{it}$$

where $b(\boldsymbol{w}_{\boldsymbol{it}}, \tilde{\boldsymbol{x}}_{\boldsymbol{it}}, \boldsymbol{\theta})$ is the input-prices bias. I use three inputs in the (deflated) input expenditure vector $\tilde{z}_{it}$: labor ($\tilde{l}$), materials ($\tilde{m}$) and capital ($\tilde{k}$). Thus, $\tilde{z}_{it} = \{\tilde{l}_{it}, \tilde{m}_{it}, \tilde{k}_{it}\}$. I also address the selectivity of firms, that is the entry and exit of firms in and out of single-product firms using strategy in De Loecker, Goldberg, Khandelwal, and Pavcnik (2016).

Addressing input price bias. I now address input-price bias. As I only see expenditure on inputs and not their quantities, the procedure requires treatment of unobserved input prices. I address this using the following procedure: I assume that input price function depends on firms location $\boldsymbol{G}_{\boldsymbol{i}}$ and input quality v_{it}. Information on input quality can be obtained from output price p_{it}, market share $\boldsymbol{ms}_{\boldsymbol{it}}$, product dummies $\boldsymbol{D}_{\boldsymbol{j}}$, and location $\boldsymbol{G}_{\boldsymbol{i}}$. The idea is that in absence of direct measures of input quality, information on output prices and their market share within a product category and location are informative of input quality. I also include rain shocks in my estimation to allow for the possibility that rain shocks may, but need not, change input prices. I assume an input price control function: $w_{it}^{x} = w_t(p_{it}, \boldsymbol{ms}_{\boldsymbol{it}}, \boldsymbol{D}_{\boldsymbol{j}}, \boldsymbol{G}_{\boldsymbol{i}}, r_{it})$. The input price bias function takes the form:

$$b(w_{it}, \tilde{z}_{it}, \boldsymbol{\theta}) = b((p_{it}, \boldsymbol{ms}_{\boldsymbol{it}}, \boldsymbol{D}_{\boldsymbol{j}}, \boldsymbol{G}_{\boldsymbol{i}}, r_{it}) \times \tilde{\boldsymbol{z}}^{\boldsymbol{c}}_{\boldsymbol{it}}; \boldsymbol{\theta}, \boldsymbol{\delta})$$

where $\tilde{\boldsymbol{z}}^{\boldsymbol{c}}_{\boldsymbol{it}} = \{1, \tilde{\boldsymbol{z}}_{\boldsymbol{it}}\}$. This allows for input expenditure vector to affect input prices by itself and separately through the interaction with the input price control function.

Productivity process, moment conditions and identification. Next, I next the identification of production function. I follow literature on production function estimation and control for unobserved productivity ω_{it} using static input demand equation for materials: $\tilde{m}_{it} = m_t(\Omega_{it}, \tilde{k}_{it}, \tilde{l}_{it}, \boldsymbol{\kappa}_{it})$, where $\boldsymbol{\kappa}_{it} = \{\boldsymbol{G}_{\boldsymbol{i}}, r_{it}, p_{it}, \boldsymbol{D}_{\boldsymbol{j}}, \boldsymbol{ms}_{\boldsymbol{it}}\}$. Inverting this provides a control function for productivity: $\omega_{it} = h_t(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \boldsymbol{k}_{\boldsymbol{it}})$. To estimate the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$ I form moments based on innovation in productivity shock ξ_{it}:

$$\omega_{it} = g(\omega_{it-1}, r_{it-1}, SP_{it}) + \xi_{it}$$

where SP is the probability of remaining single-product. Again, I include local rain shocks r_{it} in the last year to allow for the possibility that it may affect productivity. Next, I estimate the production function parameters using the following steps standard in the literature (see De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) for detailed a discussion on the estimation techniques).

1. Use the original production equation with input price bias and run the first stage:

$$q_{it} = \phi_t(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \boldsymbol{\kappa_{it}}) + \epsilon_{it} \quad \text{where} \quad \phi_t(.) = f(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}; \boldsymbol{\theta}) + B(\boldsymbol{w_{it}}, \tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \boldsymbol{\theta}) + \omega_{it}$$

2. This allows to get productivity as a function of $(\boldsymbol{\theta}, \boldsymbol{\delta})$:[43]

$$\omega_{it}(\boldsymbol{\theta}, \boldsymbol{\delta}) = \hat{\phi}_{it} - f(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}; \boldsymbol{\theta}) - b((p_{it}, \boldsymbol{ms_{it}}, \boldsymbol{D_j}, \boldsymbol{G_i}, r_{it}) \times \tilde{\boldsymbol{z}}^{\mathbf{c}}_{\boldsymbol{it}}; \boldsymbol{\theta}, \boldsymbol{\delta})$$

3. Obtain the innovation in productivity as function of $(\boldsymbol{\theta}, \boldsymbol{\delta})$ using the law of motion for ω_{it}:

$$\xi_{it}(\boldsymbol{\theta}, \boldsymbol{\delta}) = \omega_{it}(\boldsymbol{\theta}, \boldsymbol{\delta}) - E(\omega_{it}(\boldsymbol{\theta}, \boldsymbol{\delta})|\omega_{it-1}, r_{it-1}, SP_{it})$$

4. Finally, build moment conditions that identify the parameters using $E(\xi(\boldsymbol{\theta}, \boldsymbol{\delta})\boldsymbol{Y}'_{\boldsymbol{it}}) = 0$ where $\boldsymbol{Y_{it}} = \{m_{it-1}, l_{it}, k_{it}, \boldsymbol{\kappa_{it-1}}\}$ along with the higher order terms and interactions

Recovering input allocation for multi-product firms. The final step requires estimation of input allocation parameter for multi-product firms, $\rho_{ijt} = \ln \frac{W^X_{ijt} X_{ijt}}{\tilde{X}_{it}} \, \forall X \in \{V\}$. To do so, I first eliminate unanticipated shocks and measurement error using $\hat{q}_{ijt} \equiv E[q_{ijt}|\phi_t(\tilde{z}_{it}, \kappa_{it})]$. The production function can then be written as: $\hat{q}_{ijt} = f(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \hat{\boldsymbol{\theta}}, \hat{w}_{ijt}, \rho_{ijt}) + \omega_{it}$, where $\hat{q}_{ijt}$ is obtained through first-stage estimation. I use translog for the production function functional form. I can use the estimation of $\hat{\theta}$ to decompose this translog into a component separately dependent on ρ_{ijt}:

$$\hat{\omega}_{ijt} \equiv \hat{q}_{ijt} - f_1(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \hat{\boldsymbol{\theta}}, \hat{w}_{ijt}) = f_2(\tilde{\boldsymbol{z}}_{\boldsymbol{it}}, \hat{w}_{ijt}, \rho_{ijt}) + \omega_{it}$$

where I have $\hat{w}_{ijt}$ from the input price estimation. Using the translog functional form for the production function yields: $\hat{\omega}_{ijt} = \omega_{it} + \hat{a}_{ijt}\rho_{ijt} + \hat{b}_{ijt}\rho^2_{ijt} + \hat{c}_{ijt}\rho^3_{ijt}$. This expression provides with $J+1$ equations in $J+1$ unknowns $(\omega_{it}, \rho_{i1t}, ..., \rho_{iJt})$ for each multi-product firm-year. Recall that all the parameters $(\hat{a}_{ijt}, \hat{b}_{ijt}, \hat{c}_{ijt})$ are functions of $(\hat{\theta}, \hat{w}_{ijt})$. With the estimates of ρ_{ijt}, the markups can be obtained from the expression A.7.

[43] I assume a translog functional for $f(.)$. Unlike Cobb-Douglas, the use of translog function has the advantage that the output elasticities with respect to inputs depend on the level of input factors. As input factors are observed in the data for each year, the use of translog functional form also allows for time-varying output elasticities with respect to each input. The qualitative results remain robust if Cobb-Douglas production function is used instead.

Results. In this subsection, I first present summary statistics on output elasticities and markups, across sectors. I then cross-validate the measures of estimated markups and marginal costs by analyzing correlations in the data and comparing it to correlations documented in other settings.

Summary statistics. Table A.1 reports the output elasticities and returns to scale across industries and on average. The estimated coefficients for most sectors are close to constant returns to scale, with modest within-industry variation. The average returns to scale in the Indian manufacturing sector is 1.06.

Table A.1: Output Elasticities and Returns to Scale

	Labor	Capital	Material	RTS		Labor	Capital	Material	RTS
Sector	(1)	(2)	(3)	(4)	Sector	(1)	(2)	(3)	(4)
Food and beverages	0.06	0.05	0.82	0.94	Non-metal minerals	0.37	0.07	0.51	0.95
	[0.05]	[0.04]	[0.10]	[0.09]		[0.19]	[0.06]	[0.22]	[0.17]
Tobacco products	0.42	0.06	0.80	1.27	Basic Metals	0.06	0.04	0.92	1.02
	[0.19]	[0.05]	[0.14]	[0.21]		[0.05]	[0.03]	[0.06]	[0.06]
Textiles	0.19	0.03	0.96	1.17	Fabricated metal	0.26	0.16	0.74	1.17
	[0.12]	[0.03]	[0.05]	[0.15]		[0.19]	[0.09]	[0.19]	[0.15]
Wearing Apparel	0.35	0.04	0.52	0.91	Machinery	0.13	0.09	0.80	1.02
	[0.11]	[0.04]	[0.25]	[0.16]		[0.13]	[0.09]	[0.14]	[0.09]
Leather products	0.14	0.09	0.89	1.13	Electric	0.13	0.16	0.87	1.15
	[0.09]	[0.04]	[0.09]	[0.08]		[0.08]	[0.10]	[0.07]	[0.16]
Paper products	0.50	0.25	0.63	1.39	Motor vehicles	0.06	0.06	0.83	0.95
	[0.28]	[0.19]	[0.19]	[0.40]		[0.03]	[0.04]	[0.05]	[0.04]
Printing	0.15	0.14	0.71	1.00	Other transport	0.22	0.29	0.66	1.17
	[0.12]	[0.12]	[0.07]	[0.18]		[0.11]	[0.14]	[0.18]	[0.15]
Chemicals	0.06	0.02	0.89	0.97	Furniture	0.38	0.15	0.47	1.00
	[0.03]	[0.02]	[0.04]	[0.04]		[0.23]	[0.13]	[0.24]	[0.22]
Rubber and Plastic	0.17	0.29	0.40	0.87	**Average**	**0.21**	**0.12**	**0.73**	**1.06**
	[0.16]	[0.21]	[0.21]	[0.26]					

Notes: The table reports the estimated output elasticities for the production function, estimated within 2-digit industries. Columns 1-3 report the estimated output elasticity for each factor of production. Standard deviations of the output elasticities are reported in brackets. Column 4 reports the returns to scale.

Table A.2 reports the mean and median markups across each two digit industry in the manufacturing sector. The mean markup is 2.84 and the median markup is 1.50 with a standard deviation of 5.65, suggesting wide variation in markups across firms. These averages, however, mask considerable heterogeneity across industries. For example, the median markup about 1 in non-metallic minerals, whereas it is 2.42 for firms in industries that produce computing equipment.

Cross-validation. Next, I perform three exercises to validate the estimates of markups and marginal costs. First, I examine how markups vary with firms' exporting behavior. There is extensive evidence that markups are systematically higher for exporting firms than domestic firms, and markups increase upon export entry (De Loecker and Warzynski 2012; Atkin, Khandelwal, and Osman 2017; Garcia-Marin and Voigtländer 2019). Although my sample size for exporters is small,

Table A.2: Markups, by industry

Industry	Markups Mean	Markups Median	Industry	Markups Mean	Markups Median
Food and beverages	1.46	1.10	Non-metallic minerals	1.30	0.97
Tobacco products	2.53	2.37	Basic metals	2.52	1.80
Textiles	2.54	1.75	Fabricated metal	3.75	1.82
Clothing	3.14	1.08	Machinery	6.23	2.16
Leather products	4.15	1.93	Electrical mach. & comm.	3.87	1.76
Wood products	3.67	1.94	Medical equipments	5.83	2.42
Paper products	1.28	1.17	Automobiles	5.50	1.60
Printing and publishing	3.19	1.42	Other transportation	3.35	1.29
Chemicals	3.38	1.77	Furniture	2.66	1.50
Rubber and plastic	3.72	1.34	**Total**	**2.84**	**1.50**

Notes: The table displays the mean and median markups across 2-digit industries between 1998-2009. The tables trims observations that are below and above 5th and 95th percentile in each industry.

I do find that markups are higher for exporters (Columns 1 -3 of Table A.3), and are increasing in share of sales exported by firms (Columns 4 - 6 of Table A.3).

Table A.3: Markups and export status

	Dependent variable: log (markup)					
	(1)	(2)	(3)	(4)	(5)	(6)
1(exporter)	0.076***	0.067***	0.060*	-	-	-
	[0.019]	[0.020]	[0.036]	-	-	-
% of sales exported	-	-	-	0.093***	0.091***	0.168**
	-	-	-	[0.033]	[0.034]	[0.073]
Firm-product f.e.			✓			✓
Product-year f.e.	✓	✓	✓	✓	✓	✓
District-year f.e.		✓	✓		✓	✓

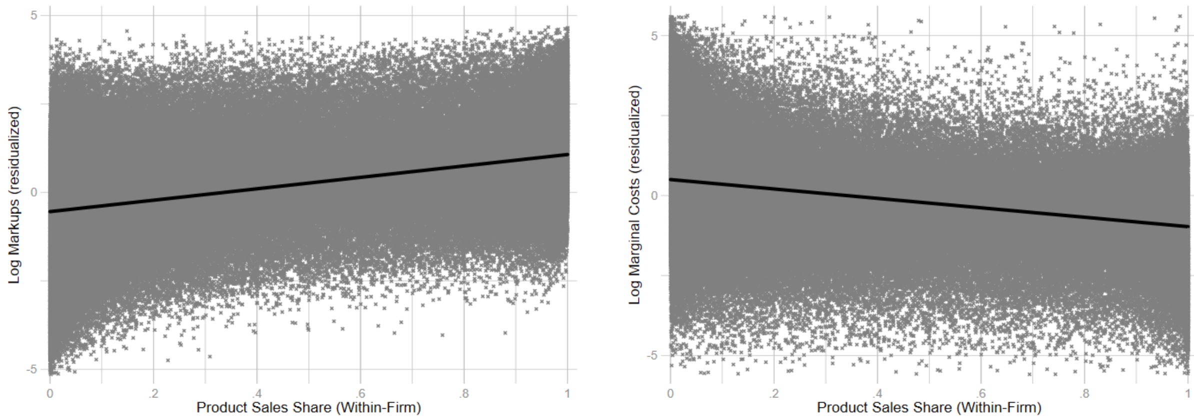
Notes: Standard errors are clustered at firm-level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Second, I analyze how markups and marginal costs vary across products within a firm as function of their share of sales. Theoretical work by Mayer, Melitz, and Ottaviano (2014) suggests that multi-product firms feature a core competency wherein their core product has the lowest marginal cost. Figure A.1 provides evidence consistent with this hypothesis. It plots the markups and marginal costs against the share of sales made for that product within each firm. Markups rise as the firm move towards its core competency, and costs decrease. These correlations are obtained without imposing any assumptions on the demand system or market structure. Despite this, the patterns are remarkably consistent with the multi-product firm literature.

Third, Figure A.2 reports the correlations between marginal and average cost of production

Figure A.1: Markups and marginal costs as share of sales within-firm

(a) Markups (b) Marginal Costs



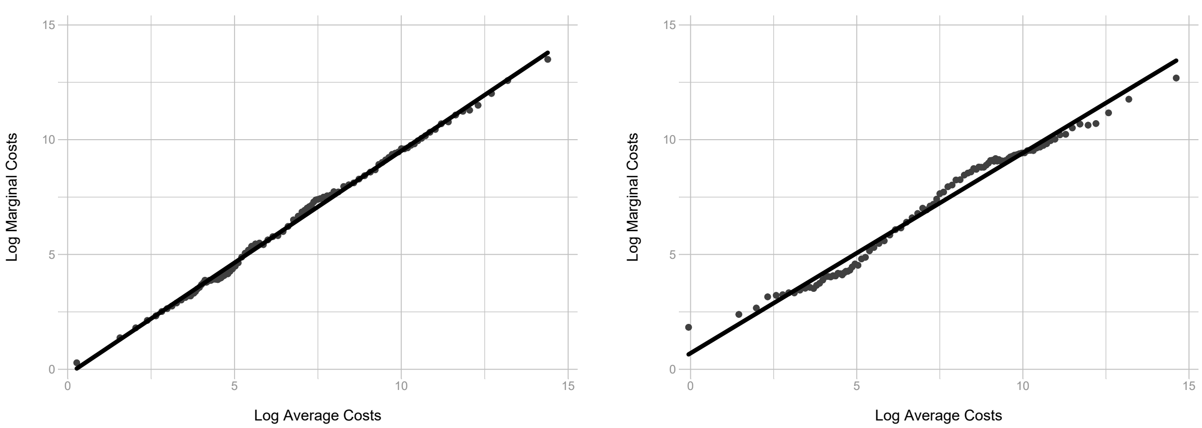
Notes: Markups and marginal costs are demeaned using product-year, firm-year and district-year fixed effects and outliers are trimmed at above and below 95th percent and 5th percent.

from ASI data, at the firm-product level. Panel (a) reports the correlation for single-product firms. In Panel (b), I also include multi-product firms in the sample. To calculate the average costs at firm-product level for multi-product firms, I multiply the firm-level cost reported in the ASI data with the share of sales across products within the firm. The figures show that marginal costs are tightly related to the average costs for the firms.

Together, these correlations provide credibility on the estimates of markups and marginal costs obtained from the cost-minimization approach.

Figure A.2: Relation between marginal costs and average costs

(a) Single product firms (b) All firms



Notes: The figure plots (log) marginal costs and (log) average costs for single-product and for all firms.

A.3 Appendix: Implications of Variable Markups for Reallocation Gains

This section complements Section 6 of main text. I first derive the relationship between the input demand X_i and any general firm subsidy S_i when markups are variable. I then derive a tax-subsidy policy S_i that guides the objective of a planner to reallocate resources while facing a fixed supply of aggregate resource.

1. Relation between input demand and tax-subsidy policy.

The change in input demand with respect to a change in subsidy is given by the total derivative of $\log X_i$:

$$d\log X_i \quad = \quad \left[\frac{\partial \log X_i}{\partial \log \mu_i}\frac{\partial \log \mu_i}{\partial \log S_i} + \frac{\partial \log X_i}{\partial \log \tau_i}\frac{\partial \log \tau_i}{\partial \log S_i}\right] d\log S_i \tag{A.9}$$

I use the fact that the subsidy (tax) reduces (increases) the marginal costs for a firm by exactly the amount of the tax-subsidy provided. That implies that $\frac{\partial \log \tau_i}{\partial \log S_i} = 1$. Next, I define firms' pass-through rate Γ_i as the elasticity of firm's price to its costs:

$$\Gamma_i \equiv \frac{\partial \log P_i}{\partial \log S_i} = \left[1 + \frac{\partial \log \mu_i}{\partial \log S_i}\right] \tag{A.10}$$

Substituting for these expression in A.9 yields:

$$d\log X_i = \left[\frac{\partial \log X_i}{\partial \log \mu_i}(\Gamma_i - 1) + \frac{\partial \log X_i}{\partial \log \tau_i}\right] d\log S_i \tag{A.11}$$

Because a input level can only be changed by varying the output level, the above expression can be rewritten as:

$$d\log X_i = \left[\underbrace{\frac{\partial \log X_i}{\partial \log Q_i}}_{=1/\theta_i^X}\ \underbrace{\frac{\partial \log Q_i}{\partial \log P_i}}_{=-\sigma_i}\ \left(\underbrace{\frac{\partial \log P_i}{\partial \log \mu_i}}_{=1}(\Gamma_i - 1) + \underbrace{\frac{\partial \log P_i}{\partial \log \tau_i}}_{=1}\right)\right] d\log S_i \tag{A.12}$$

where $\sigma_i = \mu_i/(\mu_i - 1)$ is the firms' demand elasticity. Using markup relationship from cost-minimization (equation 4) provides:

$$d\log X_i = -\left[\frac{\Gamma_i}{\theta_i^X - \alpha_i^X}\right] d\log S_i \tag{A.13}$$

2. Deriving a tax-subsidy policy.

I consider the marginal revenue product as the relevant margin of distortion. There are few advantages to do so. First, a large literature has now documented large dispersion in marginal product of inputs across manufacturing firms in developing countries (including India).[44] Viewed through a standard model of demand and supply, presence of distortions in marginal products is evidence of distortion. Second, it can be readily computed in the data using data on firms' revenue, input quantities, and estimates of output elasticity with respect to the inputs. The latter can be estimated using firm-level production data (as described in the A.2). Because distortions are not directly observed, marginal revenue product is considered a relevant summary statistic. Assuming the distortions are exogenous, a tax-subsidy policy can be devised using information on firms' marginal products. The policy should be such that the marginal product of input X is equalized across firms within an industry. Formally, let S_i represent the firm-level tax-subsidy (where $S_i > 1$ implies a tax and $S_i < 1$ implies as subsidy). Then S_i is defined such that:

$$S_i = \frac{\overline{\text{MRPX}}}{\text{MRPX}_i} \Rightarrow d\log S_i = d\log \text{MRPX}_i \tag{A.14}$$

where $d\log \text{MRPX}_i \equiv \log \overline{\text{MRPX}} - \log \text{MRPX}_i$. To obtain subsidy policy from the data requires knowledge of $\overline{\text{MRPX}}$. For this, I impose the constraint that the aggregate supply of resources X in the economy is fixed. This is the usual constraint imposed in the literature analyzing *static* misallocation. The constraint of fixed aggregate supply of resources implies that $\sum_i dX_i = 0$.

With the above objective and constraint, I can proceed with the calculation of tax-subsidy policy S_i. Using equation 3, the relation between (changes in) marginal revenue product and input demand is:[45]

$$d\log \text{MRPX}_i = (\theta_i^X - 1)d\log X_i$$

Define $\tilde{X}_i = \frac{X_i}{\theta_i^X - 1}$. This yields

$$dX_i = \tilde{X}_i \left(\log \overline{\text{MRPX}} - \log \text{MRPX}_i\right)$$

Summing over dX_i and using the aggregate supply of resources constraint $\sum_i dX_i = 0$ provides

[44] See Hsieh and Klenow (2009); Asker, Collard-Wexler, and De Loecker (2014); David and Venkateswaran (2019) and references therein.

[45] To see this, recall that $\text{MRPX} = P\frac{dQ}{dX} = P\theta^X \frac{Q}{X}$. Taking logs and first differences provides: $d\log \text{MRPX}_i = \left(\frac{d\log Q_i}{d\log X_i} - 1\right) d\log X_i$.

with the expression for the equalized $\overline{\text{MRPX}}$:

$$\log \overline{\text{MRPX}} = \left(\frac{\sum_i \tilde{X}_i \ . \ \log \text{MRPX}_i}{\sum_i \tilde{X}_i} \right) \tag{A.15}$$

Substituting for the expression of $\overline{\text{MRPX}}$ from A.15 back in equation A.14 gives us the tax-subsidy policy:

$$d \log S_i = \left[\left(\sum_i \left(\frac{\tilde{X}_i}{\sum_i \tilde{X}_i} \right) \ . \ \log \text{MRPX}_i \right) - \log \text{MRPX}_i \right]$$

B A role for measurement error? (For Online Publication)

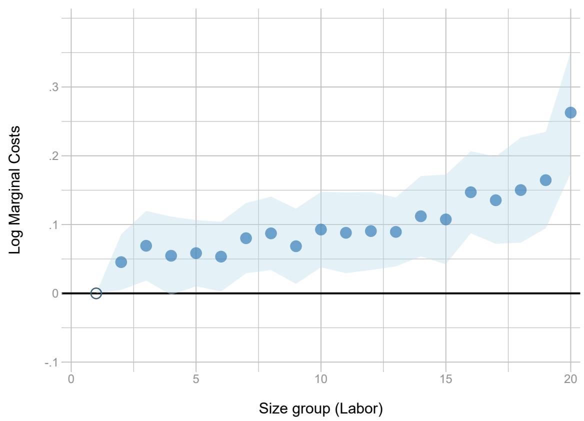
A potential concern with the analysis is that firm revenues or quantities might be measured with measurement error (ME). Notice that this is only an issue with the correlation results documented in section 4 and not for the estimates from identification strategy. In fact, an advantage of the identification strategy is that classical or non-classical ME will not affect the estimates because rain shocks are orthogonal to the error terms in estimated markup. I next address ME in correlations.

First, firms' employment might be positively associated with its prices if there are common reporting errors across the two variables across some years, generating positive bias in correlation estimates. To address this issue, I re-estimate the relation between firm's size and its costs and markups after instrumenting each firms' employment for every year using its initial employment (based on the first occurrence of each firm in the data) and its average employment across all years. The point estimates are virtually unaffected in these additional estimations. Second, I show that correlations documented in Figure I are robust to using the ranking of firms instead of using the levels. The use of ranks instead of levels relies less directly on the reported values and is less susceptible to correlations driven by outliers. Figure B.1 shows that the positive relations between costs and markups with firm size hold when using these ranking measures. Third, as documented in section 4, the correlation of marginal costs with firm size is of opposite signs across homogeneous and quality differentiated sector. Therefore, the correlation of ME across the two sectors will need to be of opposite sign to be able to explain positive relationship in differentiated sector and negative relationship of costs and size in homogeneous sector. This suggests that the correlations between firm size, markups and costs are not driven by ME bias, because such bias would have to also uniquely vary with quality differentiation.

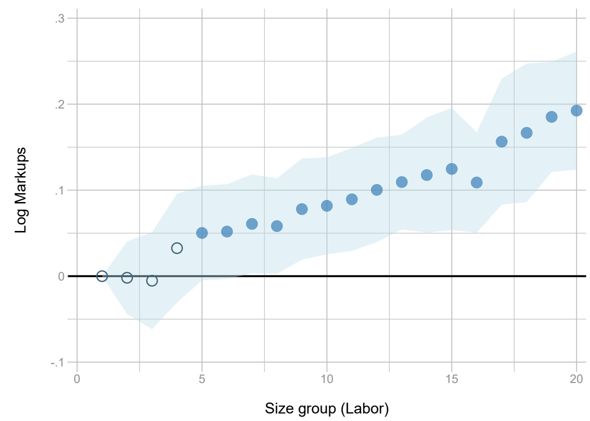
Figure B.1: Relation between firm's markups, marginal costs and size (by rankings)

Relationship by size groups

(a) Marginal costs

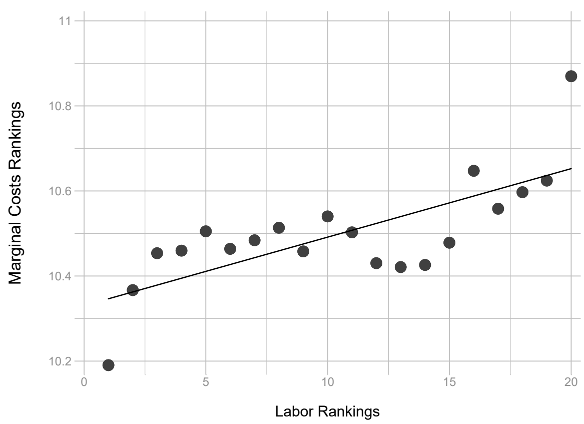


(b) Markups

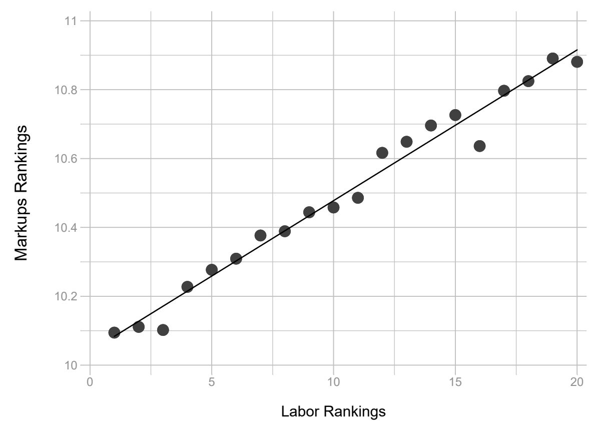


Relationship over discrete rankings

(c) Marginal costs



(d) Markups



The figure shows the relation between firm's per-unit marginal costs, markups and size by rankings across groups. Panel (a) and (b) show the relationship across firm size rankings, where a firms' ranking belongs to one of 20 size groups (based on the size of its labor force) within a district-product-year. Panel (c) and (d) also rank firms by their marginal costs and markups across 20 groups, and show how the rankings across marginal costs and markups relate to rankings across firm size.

C Comparison with existing models on firm heterogeneity (For Online Publication)

In this section, I compare the cross-sectional and time-series predictions from existing models from literature that feature firm heterogeneity and variable markups. For cross-sectional predictions, I compare the relationship between firm-size, with their markups and marginal costs. For time-series predictions, I compare the predictions on how firms across the size distribution would change their markups in response to demand shocks to the poor households. These predictions are summarized in Table C.1.

C.1 Models with monopolistic competition

Melitz (2003) is the benchmark efficiency sorting model with CES demand and this framework more efficient firms have lower marginal costs. A number of studies have incorporated Melitz (2003) framework with quality differentiation. Under these models, more productive and larger firms charge higher prices for their products. This higher price is a premium for quality and is driven by higher marginal costs: production of better quality entails expensive and better quality inputs (Kugler and Verhoogen 2011).

Under CES demand, however, all firms optimally charge a constant markup over marginal costs. Therefore, markups do not vary with firm size, and neither do they vary across time in response to demand shocks in either efficiency sorting or quality sorting frameworks.

Melitz and Ottaviano (2008) present an efficiency sorting model where firms face linear demand. Unlike CES preferences, the price elasticity of demand faced by firms is not constant in these models but rather depends on degree of competition among firms in these markets. Firms facing lower competition charge higher markups. Efficiency sorting implies that larger firms have lower marginal costs and offer lower prices, even though they have higher markups.

Zhelobodko, Kokovin, Parenti, and Thisse (2012) propose a variable markups framework with endogenous consumer demand elasticity. In their model, consumers have higher preference over larger varieties and consume more variety as their income increases. An increase in variety increases their demand elasticity and lowers firms' markups. Under this framework, we should expect the markups to decrease the most for smallest firms in response to an increase in demand from the poor.

Kneller and Yu (2016) embed quality differentiation in Melitz and Ottaviano (2008) framework. Firms with higher costs produce better quality and charge higher markup as they are able to command larger market share. However, in response to an increase in demand — irrespective of the income group from which demand increases — the markups increase.

C.2 Models with imperfect competition

In framework of Atkeson and Burstein (2008) and Edmond, Midrigan, and Xu (2015), firms face CES demand and compete in a olipolistic competitive market. Larger firms are more efficient and have lower costs. Larger firms also charge higher markups as they command higher market shares. An increase in demand from poor households, however, increases the market share for smallest and mid-size firms and they should increase their markups. Market share for largest firms shrink and they should lower their markups. Bastos, Silva, and Verhoogen (2018) incorporate quality sorting in Atkeson and Burstein (2008) framework. Due to quality sorting, costs increase with firm size in Bastos, Silva, and Verhoogen (2018). Relation between markups and firm size is similar to that in Atkeson and Burstein (2008).

Table C.1: Existing models of firm heterogeneity and variable markups

		Correlation between firm size and		Effect of Δ_t Demand from the poor on Δ_t Markup for ..		
Nature of firm heterogeneity, competition, and demand	Relevant Papers	Marginal Cost (1)	Markups (2)	Smallest firms (3)	Mid-size firms (4)	Largest firms (5)
Efficiency sorting, Monopolistic Comp., and CES	Melitz (2003)	-	0	0	0	0
Quality sorting, Monopolistic Comp., and CES	Verhoogen (2008) Kugler and Verhoogen (2011)	+	0	0	0	0
Efficiency sorting, Monopolistic Comp., and non-CES	Melitz and Ottaviano (2008) Zhelobodko et. al. (2012) Edmond, Midrigan, and Xu (2019)	-	+	- - -	- -	0
Quality sorting, Monopolistic Comp., and Linear	Kneller and Yu (2016)	+	+	+++	++	0
Efficiency sorting, Oligopolistic Comp., CES	Atkeson and Burstein (2008) Edmond, Midrigan, and Xu (2015)	-	+	+++	+/-	- - -
Quality sorting, Oligopolistic Comp., CES	Bastos, Silva, and Verhoogen (2018)	+	+	+++	++	-
Quality sorting, heterogeneous demand elasticities	This paper and data	+	+	-	- - -	-

The number of signs reflect the relative intensity of effects in Column 3-5. For example, +++ implies that the positive effect is higher when compared to ++, which is higher than +.

Taken together, this exercise shows that existing models cannot explain the three findings of the paper in combination: (a) larger firms have higher marginal costs (b) larger firms have higher markups (c) markups are decreasing for mid-sized firms in response to an increase in demand from the poor.

D A Simple Theoretical Framework (For Online Publication)

This section develops a simple model that features quality choice in a setting with heterogeneous households in consumption and heterogeneous firms in production. The model serves two purposes. First, it generates predictions on relation between firm-size, markups and costs that are consistent with the empirical correlations documented in Section 4. Specifically, the model predicts that markup dispersion in quality differentiated sector is generated due to assortative matching between firms and households. Second, and more importantly, the model generates testable prediction for how firms across the size distribution should change their markups in response to demand shocks across the income distribution, that I test in Section 5.

D.1 Model

The demand side features consumers that have non-homothetic preferences: consumers with different income levels vary in their quality valuations and demand elasticities. Specifically, when faced with identical prices, rich and poor households allocate their consumption expenditure differently across the quality ladder. The production side is a reduced-form version of the quality-choice model of Kugler and Verhoogen (2011) that features endogenous output quality choice across heterogeneous firms.

Demand. Consumers are indexed by h. As in Handbury (2019), consumers spend their income across two sectors: manufactured goods M and an outside good I. Their preferences follow a two-tier utility where the upper-tier utility depends on the consumption of outside good I: $U_h = (U_M(I_h), I_h)$. Following Handbury (2019), I assume that the outside good I is normal (and I_h is therefore analogous to the income level).[46] By making consumption on manufactured goods to be a function of outside good consumption, I allow introduction of non-homotheticity in a reduced-form manner, similar to Faber and Fally (2020). Each household derives utility from a product variety produced by firm i. Each firm produces a unique variety of product, and therefore i indexes both firms and products. Utility of household h from manufactured goods is defined by:

$$U_M(I_h) = \left[\sum_{i=1}^{I} (\zeta_i^{\nu_h} Q_{hi})^{\frac{\sigma_h - 1}{\sigma_h}} \right]^{\frac{\sigma_h}{\sigma_h - 1}} \quad \text{s.t.} \quad \sum_i P_i Q_{hi} \leq I_h$$

where each variety has a quality ζ_i, $\sigma_h > 1$ is households' demand elasticity, $\nu_h > 0$ is households' taste for quality. I assume that household utility from consuming better quality increases with their income level such that $\nu_1 < \nu_2$ if $I_1 < I_2$.[47] These preferences are common across households but non-homothetic as the utility from manufactured goods depends on income level I_h as well as households' taste for quality ν_h and demand elasticity σ_h. There are two advantages of working with this structure. First, I keep the price elasticity of demand to be constant within income groups but allow them to vary across income groups. Second, I impose no restriction on how price elasticity of demand depends on income and rather estimate it from data.

[46] I choose electricity to be the normal good, given the homogeneity of the good and its availability in all households' consumption baskets. I find that household expenditure on electricity follows a significant log-linear relation with respect to household income (slope of 0.248 with standard error of 0.005 when errors are clustered at town-level).

[47] In recent work, Comin, Lashkari, and Mestieri (2021) develop a framework that can rationalize increased willingness to pay for product quality with income levels.

Production. Each firm produces a single variety of product subject to a fixed cost F. The profit function for the firms is given by:

$$\pi_i = P_i Q_i - C'(Q_i)Q_i - F = \left(1 - \frac{1}{\mu_i}\right) Y_i - F$$

where P_i is the price of the product i, Q_i is quantity sold by firm, $C(Q_i)$ is the total cost and $Y_i = P_i.Q_i$ are the total sales made by the firm. I assume that marginal costs are increasing in firm's product quality. Specifically, following Kugler and Verhoogen (2011), the functional form for marginal costs is such that the total cost of firm is increasing in its quality and decreasing in productivity Ω_i and is given by $C(Q_i;\zeta_i,\Omega_i) = \frac{\zeta_i^\alpha Q_i}{\Omega_i} + k\zeta_i$, where $k > 0$ and $\alpha > 0$. Therefore, marginal cost for the firm is $C'(Q_i) = \frac{\zeta_i^\alpha}{\Omega_i}$, and is increasing in the quality of the product. Define the price index P_h faced by consumer group h as:

$$P_h = \left(\sum_{i=1}^{I} \left(\frac{P_i}{\zeta_i^{\nu_h}}\right)^{1-\sigma_h}\right)^{\frac{1}{1-\sigma_h}} \tag{D.1}$$

Let μ_i be the markups over marginal costs defined by $P_i = \mu_i C'(Q_i)$ gives the following expression for consumer demand curve:

$$Y_{hi} = \zeta_i^{(\sigma_h-1)(\nu_h-\alpha)}\, \mu_i^{1-\sigma_h}\, P_h^{-1}\, \Omega_i^{(\sigma_h-1)}\, I_h \tag{D.2}$$

Total sales made by firm i is given by $Y_i = \sum_i Y_{hi}$, where $Y_{hi} = P_i\, Q_{hi}$.

Firms' Optimization. In equilibrium consumers maximize utility. Firms take the consumers demand curve D.2 as given and choose their markup, and quality to maximize their profits, subject to free entry (zero profits). As all firms face same problem, I suppress subscript i for convenience:

$$\max_{\mu,\zeta}\ \pi(\mu) = \left(1 - \frac{1}{\mu}\right) \sum_h Y_h(\mu,\zeta,\Omega,I) - F$$

Proposition 1. *Average quality of household consumption basket increases in quality valuation ν_h.*

Proof: Define $s_{hi} = \frac{P_i Q_{hi}}{\sum_i P_i Q_{hi}}$ as share of household expenditure on product i. Because $\frac{ds_{hi}}{d\nu_h} = (\sigma_h - 1)(\log\zeta_i - \sum_i \log s_{hi} \log\zeta_i)$, this implies that household's expenditure shares within product group increase in ν_h for products with above average quality, and decrease in ν_h for below average quality products. Therefore, households with lower quality evaluations ν_h allocate higher share of their consumption expenditure to products with lower quality.

Proposition 2. *Product quality of a firm is increasing in its sales.*

Proof: Optimal quality produced by firm is given by:

$$\zeta_i = \frac{1}{k}\left[\left(\frac{\tilde{\sigma}_i - 1}{\tilde{\sigma}_i}\right)\ Y_i\ \left(\hat{\nu}_i - \alpha\right)\right]$$

where $\hat{\nu}_i = \left[\frac{\sum_h (\sigma_h - 1)\ \nu_h\ Y_{hi}}{\sum_h (\sigma_h - 1)\ Y_{hi}}\right]$, and $\tilde{\sigma}_i$ is defined below. Therefore, product quality of the firm is increasing in its sales. Intuitively, this is because for two firms with the same consumer base, the larger firm would be more profitable for a given quality upgrade.

Proposition 1 and 2 imply a sorting on product quality among consumer income distribution and firm size distribution — wealthier households have larger share of their consumption expenditure from larger firms. I refer to this pattern as *assortative matching* on product quality. As marginal costs are increasing in the underlying product quality, this implies that larger firms have higher marginal costs and wealthier households pay more for the products they consume. This is consistent with the correlations documented in Section 4. Next, I use equation D.2 to arrive at an expression for *firm-level* markup:

$$\mu = \frac{\sum_h \sigma_h Y_h(\mu, \zeta, \Omega, I_h)}{\sum_h (\sigma_h - 1) Y_h(\mu, \zeta, \Omega, I_h)} = \frac{\tilde{\sigma}}{\tilde{\sigma} - 1} \tag{D.3}$$

where $\tilde{\sigma}$ is the average demand elasticity faced by firm and given by:

$$\tilde{\sigma} = \frac{\sum_h \sigma_h Y_h(\mu, \zeta, \Omega, I_h)}{\sum_h Y_h(\mu, \zeta, \Omega, I_h)} = \sum_h \sigma_h \psi_h(\mu, \zeta, \Omega, I_h)$$

Here $\psi_h(\mu, \zeta, \Omega, I_h) = \frac{Y_h(\mu,\zeta,\Omega,I_h)}{\sum_h Y_h(\mu,\zeta,\Omega,I_h)}$ is the share of firm's sales made to the consumer group h. Equation D.3 allows for a new source of markup variation across firms: firms face heterogeneous market demand curves depending on composition of income groups demanding their products. These differences in demand composition faced by firms are dictated by assortative matching on product quality, leading to larger firms facing lower demand elasticities and charging higher markups.

D.2 Qualitative predictions

Prediction 1 (**Cross-section prediction**). *Under assortative matching, decreasing demand elasticities with income levels imply that producer markups are increasing in their size.*

Proof: Following Proposition 1, ψ_h is decreasing in quality for poor households. Thus, ψ_h for poor households is higher for smaller firms than larger firms. Based on D.3, this implies that firms' demand elasticity firm is decreasing in its size. As markup charged by firm is inversely related to its demand elasticity, markups of smaller firms are lower than larger firms. The prediction is consistent with the correlations documented in Section 4 (Figure I, panel (b) and (d)).

Markups responses to demand shocks. I now reintroduce subscripts for firm and time. Let I_{pt} be income for the poorest consumer group. The elasticity of markups with respect to I_{pt} is:

$$\frac{d\log\mu_{it}}{d\log I_{pt}} = \frac{-1}{\tilde{\sigma}_{it}(\tilde{\sigma}_{it}-1)}\frac{d\tilde{\sigma}_{it}}{d\log I_{pt}} = \frac{-1}{\tilde{\sigma}_{it}(\tilde{\sigma}_{it}-1)}\sum_k \sigma_{kt}\psi_{kt}\frac{d\log\psi_{kt}}{d\log I_{pt}}$$

Solving and replacing for the last term in summation gives us:

$$\frac{d\log\mu_{it}}{d\log I_{pt}} = \frac{-\psi_{hi}\times(\sigma_p-\tilde{\sigma}_{it})}{\tilde{\sigma}_{it}(\tilde{\sigma}_{it}-1)} = \frac{-\sum_{k\neq p}(\sigma_p-\sigma_k)\psi_{ki}\psi_{pi}}{\tilde{\sigma}_{it}(\tilde{\sigma}_{it}-1)} \tag{D.4}$$

It is clear that markup responses to income shocks to the poor depends on (i) share of sales made by firm across income groups ψ_{ki}, and (ii) difference between demand elasticity of the poorest income group relative to other income groups $(\sigma_p - \sigma_k)$.

Prediction 2 (Time-series prediction). *Firms lower their markups in response to an increase in demand from poor households. Moreover, this markup response to an increase in demand from poor households is non-monotonic across the firm-size distribution.*

Proof: Define $\chi_{pit} \equiv \frac{d\log\mu_{it}}{d\log I_{pt}}$ as the elasticity of firm i's markup to income shocks to the poor in year t. If poorest households have highest price elasticity of demand (i.e. $\sigma_p > \sigma_k > 1 \;\; \forall\, k$), then combined with the fact that $\psi_{ki} \geq 0 \;\; \forall\, k$ and $\tilde{\sigma}_i > 1$, equation D.4 implies $\chi_{pit} <= 0$. Thus, markups either decrease or stay same in response to positive income shocks to the poor.

Second, I am interested in how χ_{pit} varies with share of firm i's sale made to the poor ψ_{pit}. The proof for non-monotonic markup responses to an increase in demand from poor households follows in two steps, details for which are provided in Appendix D.3. First, there exists a unique $\psi_{pit} \in [0,1]$ for which the function $\frac{d\chi_{pit}}{d\psi_{pit}}$ takes the value of 0. Second, the function $\frac{d^2\chi_{pit}}{d\psi_{pit}^2}$ is strictly positive. This implies that the elasticity of markups to income shocks to the poor has a non-monotonic relation with respect to firm size.[48]

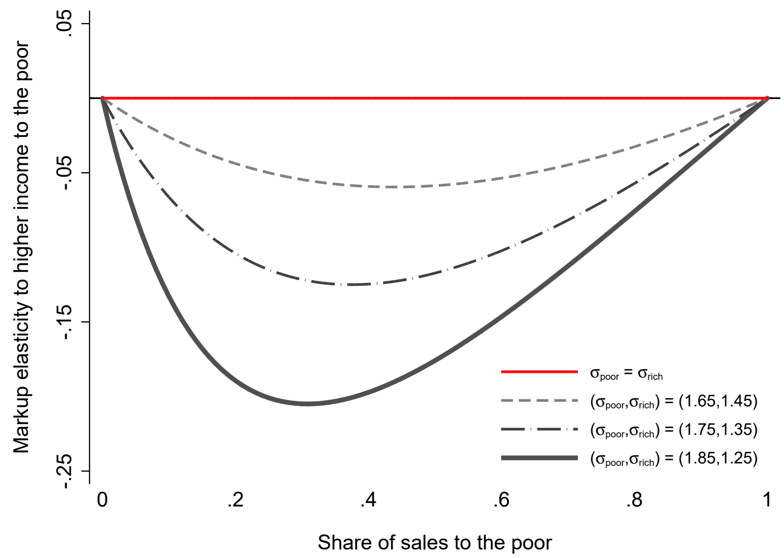
The following example illustrates this channel: Consider only two consumer groups in the economy - the poor and the rich consumers. As before, let I_{pt} be the income for the poor consumer group. Following equation D.4, the markup elasticities to the income shock I_{pt} is:

$$\frac{d\log\mu_{it}}{d\log I_{pt}} = -\frac{(\sigma_{poor}-\sigma_{rich})}{\tilde{\sigma}_{it}(\tilde{\sigma}_{it}-1)}\times\psi_{poor,i,t}\times(1-\psi_{poor,i,t})$$

Figure D.1 plots $\frac{d\log\mu_{it}}{d\log I_{pt}}$ from this specification as a function of share of sales made by firm to poorest income group, across various combinations of $(\sigma_{poor}, \sigma_{rich})$.

[48] As per Proposition 1, ψ_{pit} is monotonically decreasing in size of the firm and therefore the relation of χ_{pit} over firm size distribution follows the same relation between χ_{pit} and $(1-\psi_{pit})$.

Figure D.1: Elasticity of markups to positive income shocks to poor (as function of share of sales made to the poor by firm)



The figure shows simulated relationship of elasticity of markups to positive income shocks to households that are more price-sensitive, as a function of share of sales made by the firm to those households.

Two findings emerge. First, the elasticity of markups is zero in absence of any heterogeneity in demand elasticities (i.e. under CES demand), and in absence of assortative matching on quality (i.e in homogeneous goods sector). Second, under heterogeneous demand elasticities, markup elasticity is strictly convex with respect to share of sales made to the poor $\psi_{poor,i,t}$. The elasticity is highest for firms catering to both rich and poor households, while it approaches zero for firms making most of their sales to the poor households ($\psi_{poor,i,t} \rightarrow 1$), and for firms making most of their sales to the rich households ($\psi_{poor,i,t} \rightarrow 0$). The curvature of the function is also increasing in the demand elasticities gap between the two income groups. Intuitively, positive demand shocks to poor have a stronger positive effect on sales of firms that cater to a heterogeneous consumer base. This makes these firms pay more attention to the demand elasticity of its more price elastic consumer base, lowering their markups.[49]

[49]The model above has imposed few assumptions including a specific non-homothetic demand system. This particular demand system has the advantage of being simple, while providing tractable solutions and comparative statics. However, these functional forms are not crucial and the predictions hold under alternate demand system based on explicitly additive consumer preferences that endogenizes decreasing demand elasticities with income levels (Appendix Section D.4).

D.3 Mathematical Appendix

Proof for Prediction 2 I remove subscript i for convenience. Let $\chi_h(\psi_h) = f(\psi_h)\ .\ g(\psi_h)$, where

$$f(\psi_h) = \frac{-1}{\tilde{\sigma}(\tilde{\sigma}-1)} < 0 \quad \text{and} \quad g(\psi_h) = \sum_{k \neq h} (\sigma_h - \sigma_k)\psi_k \psi_h > 0$$

To see that the function $\chi_h(\psi_h)$ has a unique minimum, we first solve for $\chi_h'(\psi_h)$

$$\chi_h'(\psi_h) \;\; = \;\; f(\psi_h)\ .\ \left[(f(\psi_h)(2\tilde{\sigma}-1)\ .\ \psi_h + 1) \left(\sum_{k \neq h} (\sigma_h - \sigma_k)\psi_k \right) - \psi_h\ .\ \sum_{k \neq h} (\sigma_h - \sigma_k) \right]$$

Solving for $\chi_h'(\psi_h) = f'(\psi_h)\ .\ g'(\psi_h)$ = 0 gives

$$(f(\psi_h)(2\tilde{\sigma}-1)\ .\ \psi_h + 1) \left(\sum_{k \neq h} (\sigma_h - \sigma_k)\psi_k \right) = \psi_h\ .\ \sum_{k \neq h} (\sigma_h - \sigma_k)$$

The left hand side is decreasing in ψ_h and the right hand side is increasing in ψ_h. Therefore, there exists a unique $\psi_h^* \in [0,1]$ for which $\chi_h'(\psi_h^*) = 0$. Next, we solve for $\chi_h''(\psi_h)$:

$$\chi_h''(\psi_h) = f''(\psi_h)\ .\ g(\psi_h) + 2\ .\ f'(\psi_h)\ .\ g'(\psi_h) + f(\psi_h)\ .\ g''(\psi_h) \tag{D.5}$$

$$\begin{aligned} \text{where:} \quad f''g \;\; &= \;\; 2\frac{(f')^2}{f} g + 2f^2 \left(\frac{d\tilde{\sigma}}{d\psi_h} \right)^2 g \quad \text{and} \quad f'g' = -\frac{(f')^2}{f} g \\ \text{and} \quad g''f \;\; &= \;\; -2\ .\ f\ .\ \left[\sum_{k \neq h} (\sigma_h - \sigma_k) \right]\ .\ > 0 \end{aligned}$$

Substituting these expressions in D.5

$$\chi_h''(\psi_h) \;\; = \;\; 2f^2 - 2\ .\ f\ .\ \left[\sum_{k \neq h} (\sigma_h - \sigma_k) \right]\ .\ > 0$$

Therefore, χ_{ht} is a convex function with a unique minimum.

D.4 Alternative Demand System 1: Explicitly Additive Preferences

In this section, I consider an alternate demand system with explicitly additive preferences. Consumers have directly explicitly additive preferences (Generalized Stone-Geary preferences) and have heterogeneous quality valuations. The production side is the same as Section D.

Demand. Consumer h has Stone-Geary preferences over the consumption goods Q_{hi}

$$U_h = \sum_i \left[\zeta_i^{\nu_h} \left(Q_{hi} - \underline{Q}_{hi} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

where, as before, ζ_i is the product quality; ν_h captures the consumer's valuation of quality, which I assume is strictly increasing in the exogenous income level I_h. The price elasticity of demand for consumer h for product i is given by:

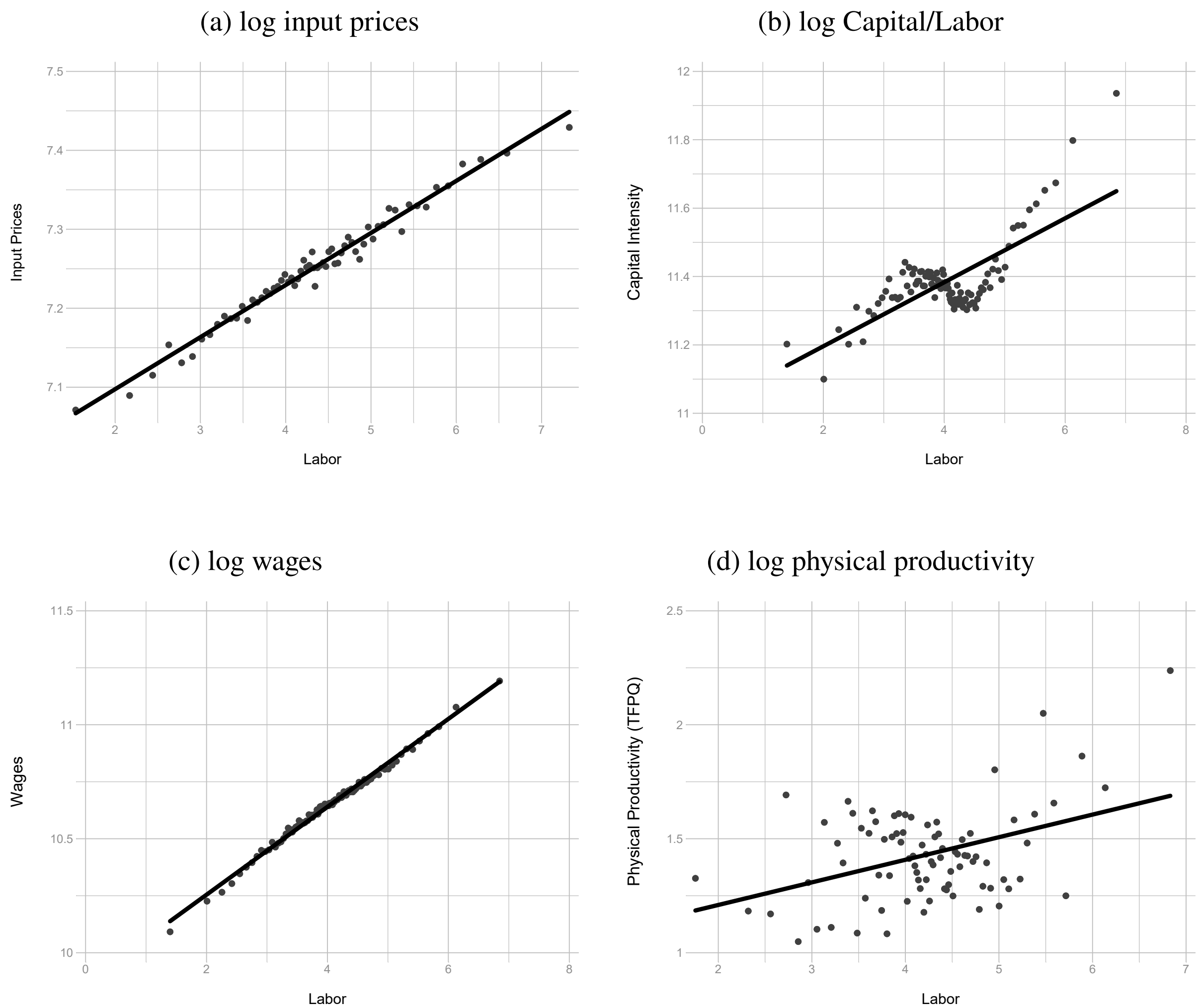
$$\sigma_{hi} \equiv -\frac{P_i}{Q_{hi}} \frac{dQ_{hi}}{dP_i} = \sigma \left(1 - \frac{\underline{Q}_{hi}}{Q_{hi}} \right) \left(1 + \frac{P_i \underline{Q}_{hi}}{I_h - \sum_h P_i \underline{Q}_{hi}} \right)$$

The price elasticity of consumer is decreasing with the amount of residual income. Therefore, wealthier households have lower price elasticity of demand.

For firm i, the demand elasticity is the *sales-weighted* average of price elasticity of demand of its consumer base: $\tilde{\sigma}_i = \sum_h \sigma_{hi} \psi_{hi}$. The greater the firm's share of sales made to a particular income group ψ_{hi}, the higher is the weight the firm places on that group's price elasticity of demand σ_{hi}. Since larger firms make larger share of their sales to wealthier households — and because demand elasticity σ_{hi} is lower for wealthier households —, larger firms charge higher markups.

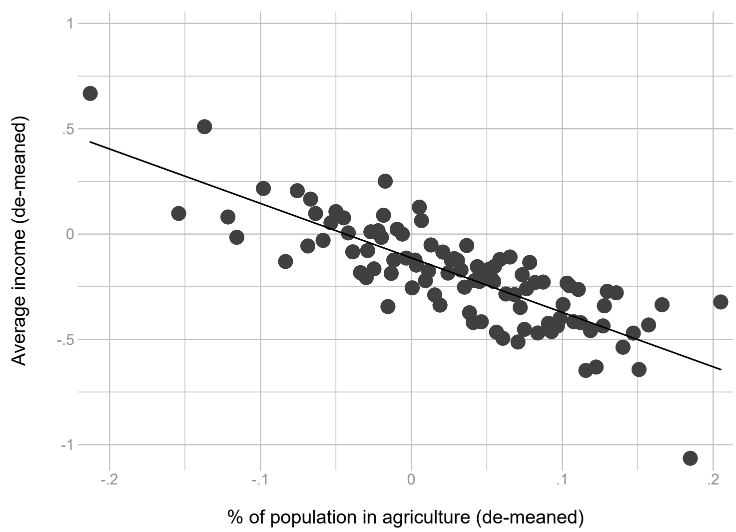
E Appendix Figures and Tables (For Online Publication)

Figure E.1: Relation between firm size and input factor costs



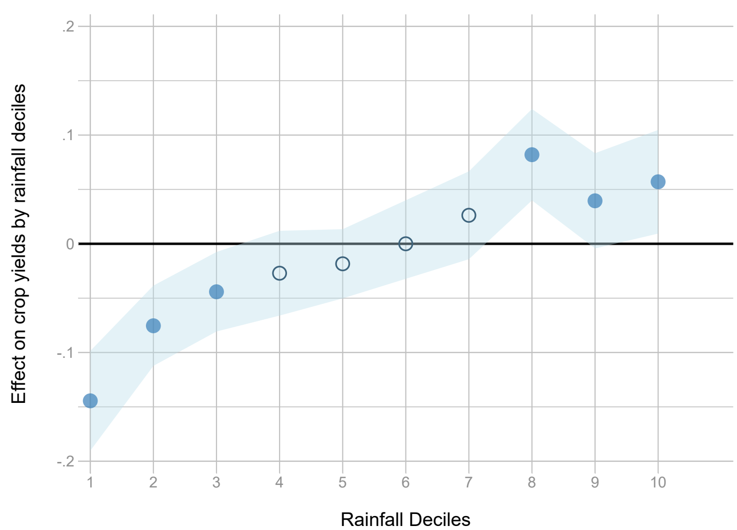
The figure shows the relation between firm size (as measured by its labor force) and input prices (Panel (a)), capital intensity (Panel (b)), wages per unit labor (Panel (c)), physical productivity TFPQ (Panel (d)). All variables are measured in logs. All specifications control for district-by-product-by-year fixed effects. Each dot represents 1% of observations. Source: ASI

Figure E.2: Share of agricultural population and average income in district



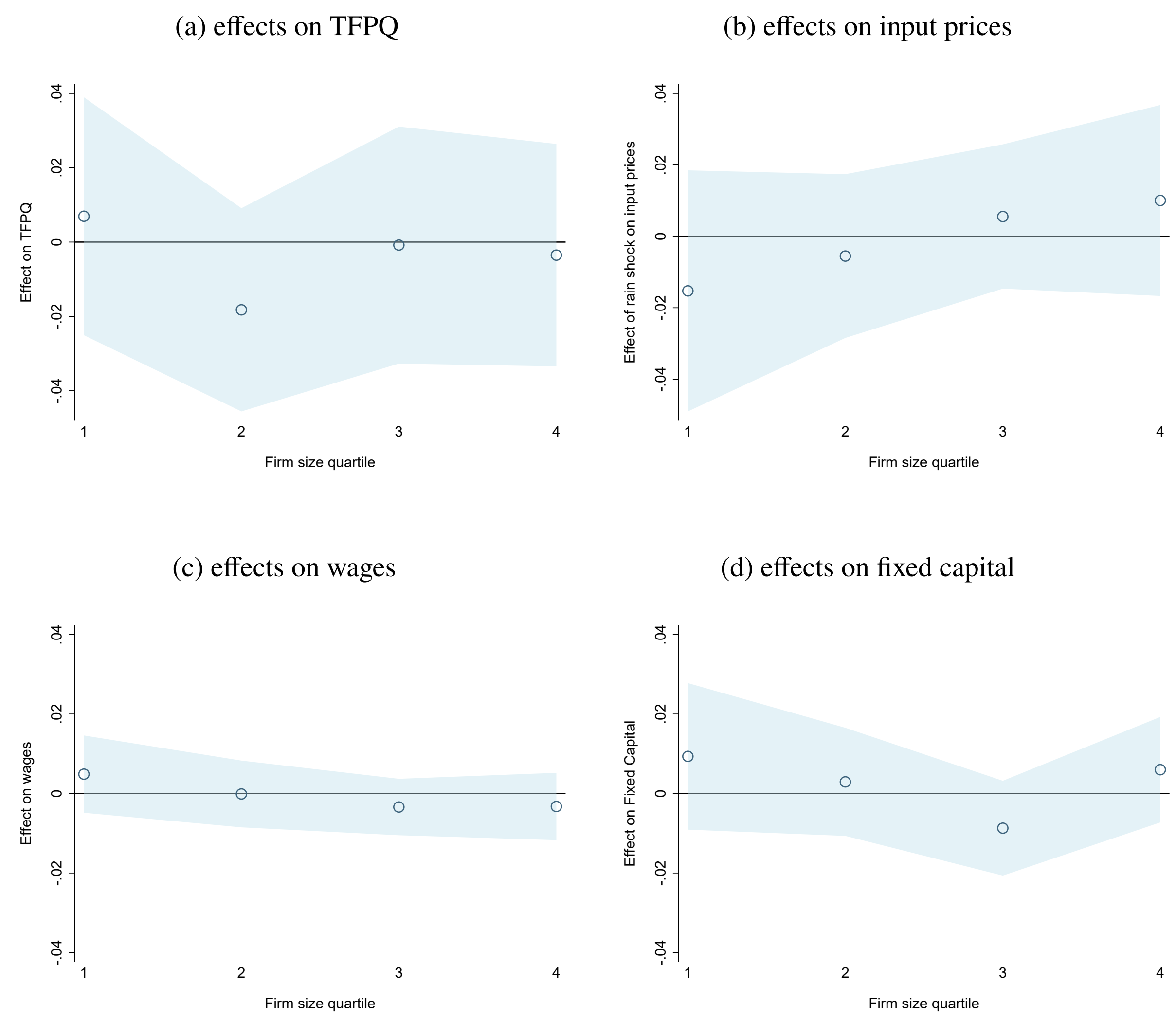
The figure plots the relation between share of population involved in agricultural activities and average income in the district. Both axes plot the residualized values after removing state fixed effects. The correlation is -4.52 and is significant at 1% levels (t = -5.39) when standard errors clustered at district level. Source: NSS

Figure E.3: Effect of rain shocks on agricultural yields



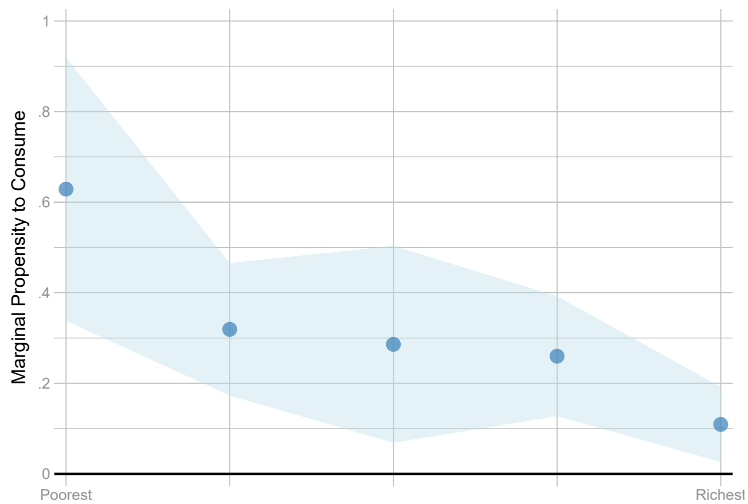
The figure plots coefficients and 95% confidence intervals from a regression of log crop yields on dummies for each decile of the rainfall distribution within the district. Log crop yields is the log of a weighted average of yields of the 15 crops for which data is available in the VDSA database. The yield for each crop has first been normalized by the mean yield of that crop in the district. Weights are the mean percentage of land area planted with a given crop in a district. Each decile dummy equals 1 if monsoon rainfall in the current year fell within the given decile of the district's usual rainfall distribution for that year and equals 0 otherwise. The omitted category against is the 6^{th} decile. Regression specification includes district and year fixed effects. Standard errors are clustered at district level.

Figure E.4: Effect of rain shocks on TFPQ, input prices, wages and capital



The figure shows the estimates of the effect of rain shocks across the firm-size distribution on TFPQ (Panel (a)), input prices (Panel (b)), wages per unit labor (Panel (c)), fixed capital (Panel (d)). All dependent variables are measured in logs. All specifications control for firm age and size quartile-year fixed effects. 95% confidence intervals are represented by shaded blue area. Bold circles indicate results that are significant at the 10 level, and hollow circles statistically insignificant from 0 at the 10% level.

Figure E.5: Marginal Propensity to Consume (MPC) across income groups

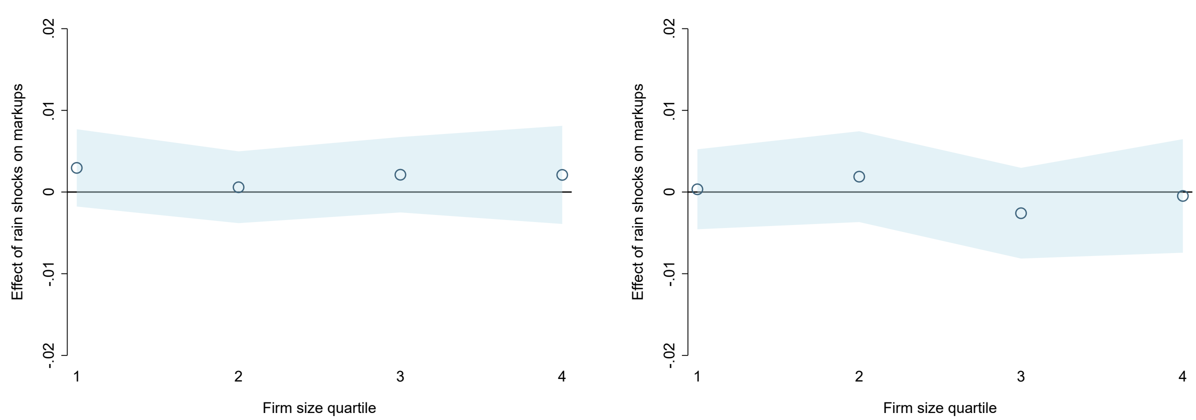


The figure reports the estimate of marginal propensity to consume (MPC) across income groups. It plots the estimates $\alpha(z)$ across five income groups based on the following specification: $\Delta \log x_{ivt}(z) = \alpha(z) \Delta \log y_{ivt}(z) + \beta_i + \gamma_{vt} + \epsilon_{ivt}$ where β_i is the household fixed effect and γ_{vt} is a town-year fixed effects. Changes in employment status are used as an instrument for changes in income. Source: CMIE

Figure E.6: Effects of placebo and past rain shocks on markups

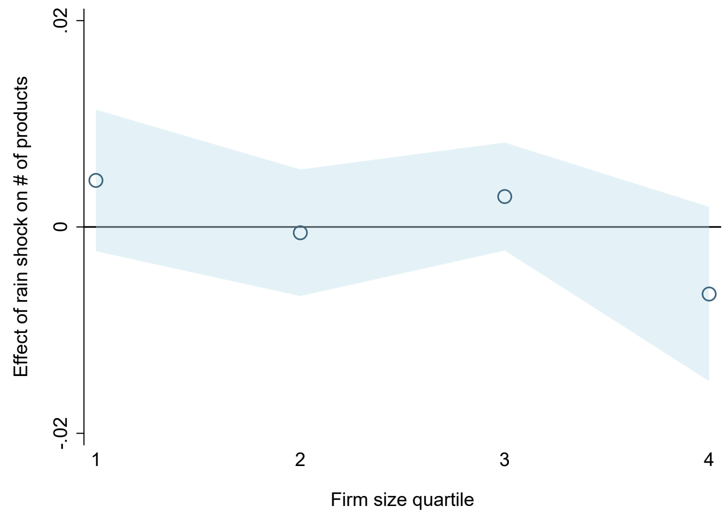
(a) Placebo (next year's) rain shocks

(b) Lagged (past year's) rain shocks



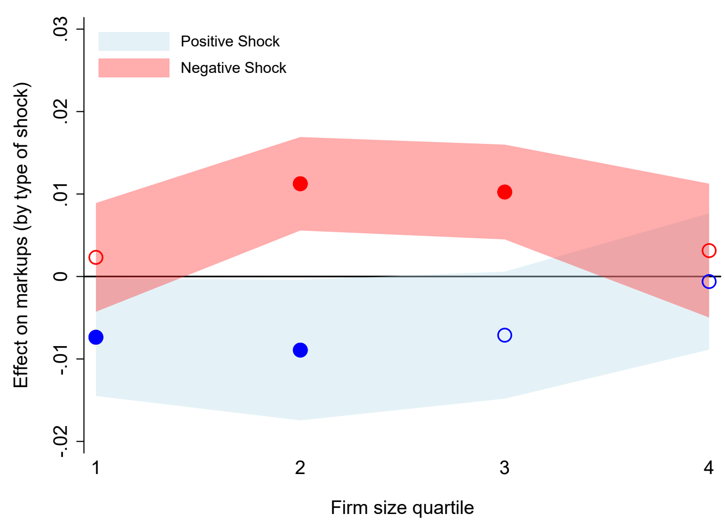
95% confidence intervals are represented by shaded blue area. Bold circles indicate results that are significant at the 10 level, and hollow circles statistically insignificant from 0 at the 10% level.

Figure E.7: Effect of rain shocks on number of products



The figure reports the heterogeneous effects of rain shocks on number of products based on specification: $\log y_{it} = \sum_{r=1}^{4} \beta^r.\ (\text{Shock}_{dt} \times Q_i^r) + \alpha_i + \alpha_{kt} + \epsilon_{it}$, where y_{it} are the number of products for firm i in year t. Shock_{dt} is as defined in the main text.

Figure E.8: Effect on markups by positive and negative shocks

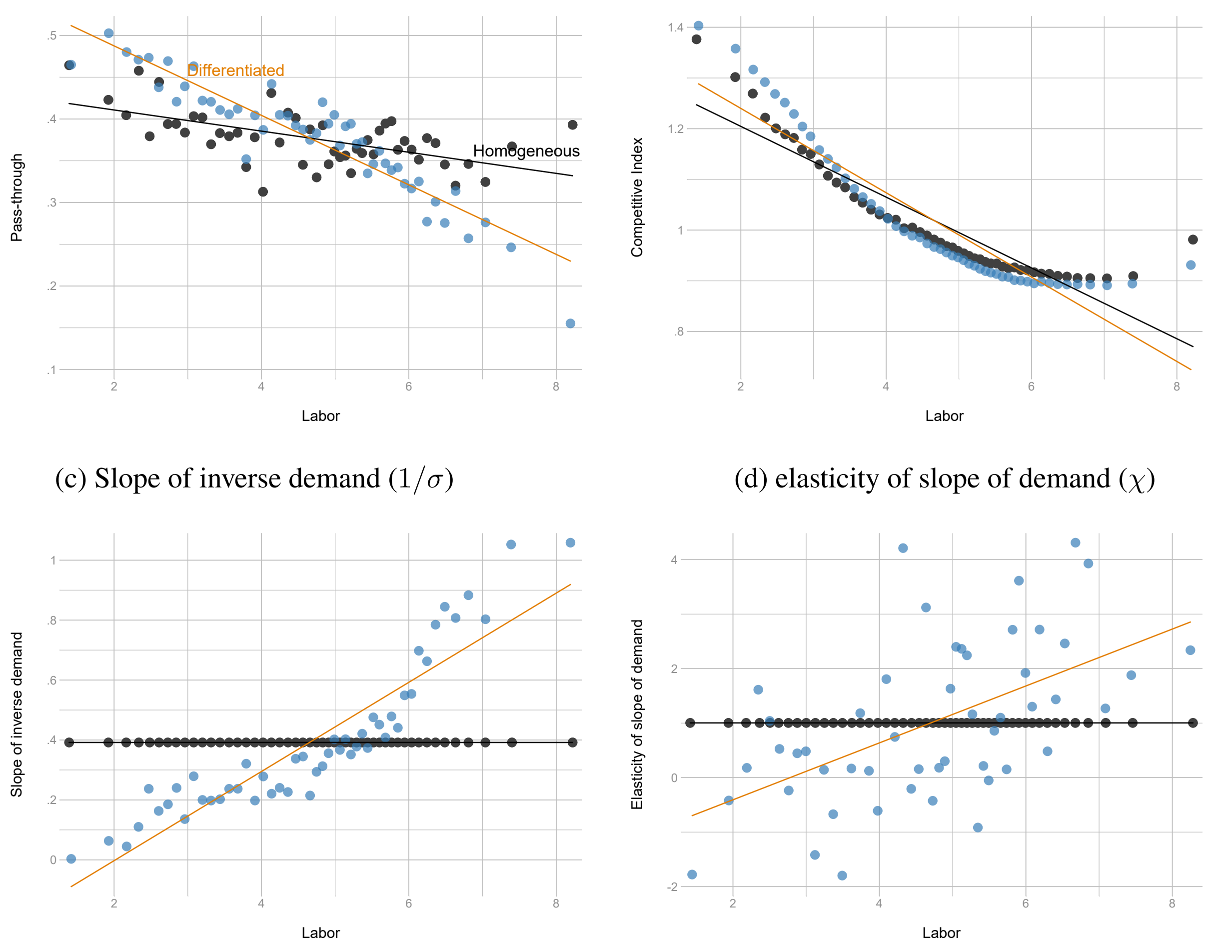


The figure reports the heterogeneous effects of positive and negative rain shocks on markups from the specification: $\log \mu_{ijt} = \sum_{r=1}^{4} \beta^{r,+}.\ (\text{Shock}_{dt}^{+} \times Q_i^r) + \sum_{r=1}^{4} \beta^{r,-}.\ (\text{Shock}_{dt}^{-} \times Q_i^r) + \alpha_{ij} + \alpha_{jt} + \Gamma' X_{ijt} + \epsilon_{ijt}$, where Shock_{dt}^{+} and Shock_{dt}^{-} takes the value of 1 if $\text{Shock}_{dt} = +1$ and $\text{Shock}_{dt} = -1$, and zero otherwise. Specification includes firm-product, product-year fixed effects and controls for firm age, size quartile-year fixed effects, and log marginal costs. 95% confidence intervals are represented by shaded area. Bold circles indicate estimates significant at the 10% level, and hollow circles statistically insignificant from 0 at the 10% level.

Figure E.9: Pass-through rate, competitiveness, demand slope and curvature across firm size

(a) Pass-through rate (Γ) (b) Competitive index (Φ)

(c) Slope of inverse demand ($1/\sigma$) (d) elasticity of slope of demand (χ)



The figure shows the estimates of pass-through Γ_i (Panel (a)), competitiveness index Φ_i (Panel (b)) , slope of inverse demand $1/\sigma_i$ (Panel (c)), elasticity of slope of demand (super-elasticity) χ_i (Panel (d)) as a function of firm-size.

Table E.1: Summary Statistics: Distribution of sales across industries

	Share of Output		Share of Output
15 Food and beverages	23%	26 Non-metal minerals	8%
16 Tobacco products	2%	27 Basic Metals	15%
17 Textiles	8%	28 Fabricated metal	1%
18 Wearing apparel	1%	29 Machinery	5%
19 Leather products	1%	31 Electric	2%
20 Wood products	0%	32 Communications prod.	1%
21 Paper products	1%	33 Medical equipment	0%
22 Printing	0%	34 Motor vehicles	6%
23 Coke products	7%	35 Other transport	4%
24 Chemicals	11%	36 Furniture	1%
25 Rubber and Plastic	3%		

The table reports the share of total output by 2-digit industries (averaged across years) in the Annual Survey of Industries (ASI) data.

Table E.2: Baseline correlations using alternative measures of firm size

	Dependent variable: log of ...				
Panel A:	Marg. Cost	Markup	Material Inputs	K/L	Wages
	(1)	(2)	(3)	(4)	(5)
(log) sales	0.020**	0.055***	0.048***	0.246***	0.164***
	[0.008]	[0.009]	[0.011]	[0.009]	[0.006]
R-squared	0.870	0.639	0.410	0.682	0.830
Panel B.	(1)	(2)	(3)	(4)	(5)
(log) assets	-0.011	0.067***	0.049***	0.614***	0.119***
	[0.007]	[0.004]	[0.008]	[0.015]	[0.005]
R-squared	0.870	0.641	0.410	0.879	0.809
Observations	167,221	167,221	443,022	167,221	167,221
Industry f.e.	✓	✓	✓	✓	✓
District-prod.-year f.e.	✓	✓	✓	✓	✓

The table reports the correlation from Table I using alternate definition of firm size based on total sales (Panel A) and total fixed assets (Panel B). Standard errors clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.3: Estimates of price-elasticity of demand (σ)

	Dependent variable: (1- σ)		
	OLS	IV	IV
	(1)	(2)	(3)
All households	-0.408***	-0.577***	-
	[0.022]	[0.031]	-
Poorest Quintile (Relative to Richest)	-	-	-0.726***
	-	-	[0.044]
2nd poorest Quintile (Relative to Richest)	-	-	-0.696***
	-	-	[0.040]
Median Quintile (Relative to Richest)	-	-	-0.602***
	-	-	[0.039]
2nd richest Quintile (Relative to Richest)	-	-	-0.504***
	-	-	[0.032]
Observations	103,767	103,767	103,767
R-squared	0.484		
F-stat		82.607	12.555
Region-product f.e.	✓	✓	✓
Quintile f.e.	✓	✓	✓

The table reports the estimate of price-elasticity of demand based on the estimating equation 5. Column 2-3 estimates are based on the IV specification that instruments change in price of a good with state-level leave out mean price changes (described in Section 5.2). Standard errors clustered at district level are reported in brackets. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table E.4: Rainfall induced income shocks for poor population

	Dependent variable: log of ...				
	Agricultural output		Daily wages		
	Crop yield (1)	Revenue per unit area (2)	Rural agri. labor (3)	Rural non-. -agri labor (4)	Non-rural labor (5)
Shock$_{dt}$ (-1/0/+1)	0.045***	0.035***	0.027***	-0.009	0.017
	(0.005)	(0.005)	[0.008]	[0.009]	[0.011]
Observations	38,280	38,280	115,852	102,910	154,939
R-squared	0.887	0.853	0.516	0.271	0.124
District-crop f.e.	✓	✓			
Crop-year f.e.	✓	✓			
District f.e.			✓	✓	✓
Year f.e.			✓	✓	✓

The table reports the effect of rain shocks on agricultural productivity and labor market. Standard errors clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.5: Estimates of price elasticities across industries

Sector	OLS (1)	IV (2)	Sector	OLS (1)	IV (2)
Pooled (Average)	-0.668***	-2.364***	Rubber and plastic	-0.634***	-2.916***
	[0.011]	[0.044]		[0.025]	[0.124]
Food and beverages	-0.472***	-4.543***	Non-metal minerals	-0.580***	-2.305***
	[0.022]	[0.168]		[0.024]	[0.080]
Tobacco products	-0.281***	-2.617***	Basic Metals	-0.641***	-2.060***
	[0.104]	[0.317]		[0.015]	[0.051]
Textiles	-0.530***	-4.462***	Fabricated metal	-0.675***	-2.014***
	[0.028]	[0.174]		[0.015]	[0.049]
Wearing apparel	-0.465***	-3.900***	Machinery	-0.783***	-2.011***
	[0.056]	[0.194]		[0.013]	[0.045]
Leather products	-0.510***	-3.438***	Electric	-0.747***	-1.919***
	[0.063]	[0.257]		[0.016]	[0.045]
Wood products	-0.726***	-2.348***	Communications	-0.733***	-2.033***
	[0.035]	[0.088]		[0.026]	[0.062]
Paper products	-0.718***	-2.559***	Medical equipment	-0.820***	-2.104***
	[0.032]	[0.136]		[0.026]	[0.061]
Printing	-0.709***	-2.534***	Motor vehicles	-0.626***	-1.986***
	[0.035]	[0.119]		[0.019]	[0.051]
Coke products	-0.460***	-3.241***	Other transport	-0.643***	-1.993***
	[0.039]	[0.136]		[0.018]	[0.050]
Chemical	-0.517***	-3.520***	Furniture	-0.702***	-2.531***
	[0.019]	[0.120]		[0.026]	[0.077]

The table reports the estimated of price elasticities from specification 9. Columns 2 report IV estimates where price is instrumented with marginal costs. All specifications include product-year fixed effects. Standard errors are clustered by firm-product level (N = 133,094). Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.6: Testing for serial correlation in rainfall

	Dependent variable: $\text{RainDeviation}_{d,t}$			
	1998-2009 (Sample Years)		1990-2014	
	(1)	(2)	(3)	(4)
$\text{RainDeviation}_{d,t-1}$	-0.007	-0.007	-0.013	-0.014
	(0.029)	(0.028)	(0.015)	(0.015)
$\text{RainDeviation}_{d,t-2}$	-	0.007	-	0.016
	-	(0.030)	-	(0.012)
Observations	3,116	3,116	7,850	7,850
R-squared	0.231	0.231	0.261	0.261
District f.e.	✓	✓	✓	✓
Year f.e.	✓	✓	✓	✓

This table tests for serial correlation in rainfall. The estimates are based on the following specification: $\text{RainDeviation}_{dt} = \alpha_d + \alpha_t + \beta_1 \text{ RainDeviation}_{d,t-1} + \beta_2 \text{ RainDeviation}_{d,t-2} + \epsilon_{dt}$, where $\text{RainDeviation}_{dt}$ is the rainfall deviation in district d and year t form the median rainfall of the district since 1960. Standard errors are clustered at the district level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.7: Robustness to definition of rain shocks

	Percentile cut-off for Positive/Negative Shocks				Deviations from
	80/20	80/30	85/15	90/10	the median
	(1)	(2)	(3)	(4)	(5)
	Panel A. Dependent variable: log (markup)				
Shock_{dt}	-0.005**	-0.005***	-0.005**	-0.005**	-0.002***
	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]
R-squared	0.989	0.989	0.989	0.989	0.989
	Panel B. Dependent variable: log (marginal costs)				
Shock_{dt}	0.007	0.002	0.008	0.007	0.003
	[0.007]	[0.007]	[0.008]	[0.010]	[0.003]
R-squared	0.952	0.952	0.952	0.952	0.952

The table shows the estimates from specification 8 based on alternate definitions of rain shocks. In Column 1-4 I use different cut-offs of rain shocks in equation 8: positive and negative shocks are defined as rain shocks above/below 80/20, 80/30, 85/15 and 90/10 percentiles. In Column 5, I use continuous measure of rain shock defined as rainfall deviation relative to the historical rainfall received in the district. All specifications include firm-product and product-year fixed effects. Standard errors clustered at district level are reported in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1. (N = 133,094)

Table E.8: Effect of rain shocks on exporters

	Dependent variable: log of ...					
	quantity	markup	marg. cost	quantity	markup	marg. cost
	(1)	(2)	(3)	(4)	(5)	(6)
Shock$_{dt}$ (-1/0/+1)	0.015	0.019	-0.010	-	-	-
	[0.023]	[0.027]	[0.030]	-	-	-
Rain deviations	-	-	-	0.012	0.006	-0.002
from median$_{dt}$	-	-	-	[0.010]	[0.010]	[0.012]
Firm-product f.e.	✓	✓	✓	✓	✓	✓
Product-year f.e.	✓	✓	✓	✓	✓	✓

Notes: The table analyzes the effect of rain shocks on exporters. Standard errors clustered at district level are reported in parenthesis. Significance: *** p<0.01, ** p<0.05, * p<0.1. (N=10,114)

Table E.9: Effects of rain shocks on markups across firm-size distribution

	Dependent variable: log markup							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shock$_{dt}$ (-1/0/+1)								
× First size quartile	-0.003	-0.003	-0.003	-0.004	-0.004	-0.002	-0.001	-0.002
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.003]
× Second size quartile	-0.009***	-0.010***	-0.009***	-0.009***	-0.010***	-0.008**	-0.007**	-0.008**
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]
× Third size quartile	-0.007**	-0.005*	-0.007**	-0.007**	-0.008***	-0.006*	-0.005	-0.006*
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]
× Fourth size quartile	0.001	0.000	0.001	0.001	0.001	0.003	0.003	0.003
	[0.003]	[0.004]	[0.003]	[0.004]	[0.003]	[0.004]	[0.004]	[0.004]
Observations	133,094	122,828	133,094	133,094	133,094	133,094	133,094	133,094
R-squared	0.989	0.990	0.989	0.989	0.989	0.989	0.989	0.989
Firm-product f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Product-year f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Controls	Baseline Specification	Single-plant firms	+ Age control	+ Size-year control	Past 2-year shocks controls	National Market access control	In + out-state market access	(3)-(7) controls

The table reports effects of rain shocks on markups across the firm-size distribution (β^r from specification 10). Shock$_{dt}$ is as defined in the text. All columns include firm-product, product-year fixed effects and control for log marginal costs. Standard errors clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.10: Firm's entry/exit in response to rain shocks

	1(entry)	1(exit)	1(entry)	1(exit)
	(1)	(2)	(3)	(4)
Shock$_{dt}$(-1/0/+1)	0.001	0.000	-	-
	[0.001]	[0.001]	-	-
Shock$^{+}_{dt}$	-	-	-0.002	-0.001
	-	-	[0.002]	[0.001]
Shock$^{-}_{dt}$	-	-	-0.003	-0.002
	-	-	[0.002]	[0.001]
Observations	226,275	226,275	226,275	226,275
R-squared	0.358	0.312	0.358	0.312
Firm f.e.	✓	✓	✓	✓
Year f.e.	✓	✓	✓	✓

The table reports the estimates of new firm entry or incumbent exit based on specification: $1(entry/exit)_{it} = \alpha_i + \alpha_t + \beta$. Shock$_{dt} + \epsilon_{idt}$, where 1(entry) takes the value of 1 in the first year of firm's operation and 1(exit) takes the value of 1 when a firm is reported to be Closed in the survey. Shock$^{+}_{dt}$ and Shock$^{-}_{dt}$ takes the value of 1 if Shock$_{dt} = +1$ and Shock$_{dt} = -1$, and zero otherwise. Standard errors are clustered at the district level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table E.11: Composition effect versus size effect

	Dependent variable: log (markup)		
	(1)	(2)	(3)
Shock$_{dt}$ (-1/0/+1)	-0.002	-0.010*	-0.010*
	[0.002]	[0.005]	[0.005]
Shock$_{dt}$ $\times$ **1**(High Share of rural pop.)$_d$	-0.006**	-	-0.007**
	[0.003]	-	[0.003]
Shock$_{dt}$ $\times$ log(Total rural population)$_d$	-	0.000	0.001
	-	[0.000]	[0.000]
Observations	133,094	133,094	133,094
R-squared	0.989	0.989	0.989
Firm-product f.e.	✓	✓	✓
Product-year f.e.	✓	✓	✓

The table reports the effects of rain shocks on markups, by share of rural population and total rural population. 1(High share of agricultural population)$_d$ takes the value of 1 if more than two-third of districts' population is rural as reported in the 2001 Census of India. Total rural population for the district is sourced from the 2001 Census of India. All columns include firm-product, product-year fixed effects and control for log marginal costs. Standard errors clustered at district level are reported in brackets. Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table E.12: Robustness to financial frictions

	Dependent variable: log (markup)		
	(1)	(2)	(3)
Shock_{dt} (-1/0/+1)			
× First size quartile	-0.003	-0.004	-0.003
	[0.003]	[0.003]	[0.003]
× Second size quartile	-0.009***	-0.009***	-0.009***
	[0.003]	[0.003]	[0.003]
× Third size quartile	-0.007**	-0.007**	-0.007**
	[0.003]	[0.003]	[0.003]
× Fourth size quartile	0.001	0.001	0.001
	[0.004]	[0.004]	[0.004]
Shock_{dt} (-1/0/+1)			
× Cash Ratio	-0.003		-0.003
	[0.015]		[0.015]
× Leverage		-0.000	-0.000
		[0.000]	[0.000]
Observations	132,746	132,746	132,746
R-squared	0.989	0.989	0.989
Firm-product f.e.	✓	✓	✓
Product-year f.e.	✓	✓	✓

The table tests for robustness of estimates after controlling for differential effects of rain shocks on firms' financial strength. All specifications include firm-product, product-year fixed effects and control for firm age, size quartile-year fixed effects and for log marginal costs. Coefficients on levels of financial strength are not reported for brevity. Standard errors clustered at district level are reported in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Appendix References (For Online Publication)

Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013–1063.

Atkeson, A., and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *American Economic Review*, 98(5), 1998–2031.

Atkin, D., A. K. Khandelwal, and A. Osman (2017): "Exporting and Firm Performance: Evidence from a Randomized Experiment," *The Quarterly Journal of Economics*, 132(2), 551–615.

Bastos, P., J. Silva, and E. Verhoogen (2018): "Export Destinations and Input Prices," *American Economic Review*, 108(2), 353–92.

Comin, D., D. Lashkari, and M. Mestieri (2021): "Structural Change With Long-Run Income and Price Effects," *Econometrica*, 89(1), 311–374.

David, J. M., and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109(7), 2531–67.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016): "Prices, Markups, and Trade Reform," *Econometrica*, 84(2), 445–510.

De Loecker, J., and F. Warzynski (2012): "Markups and Firm-Level Export Status," *American Economic Review*, 102(6), 2437–71.

Edmond, C., V. Midrigan, and D. Y. Xu (2015): "Competition, Markups, and the Gains from International Trade," *American Economic Review*, 105(10), 3183–3221.

——— (2019): "How Costly Are Markups?," *Working Paper*.

Faber, B., and T. Fally (2020): "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *Working Paper*.

Garcia-Marin, A., and N. Voigtländer (2019): "Exporting and Plant-Level Efficiency Gains: It's in the Measure," *Journal of Political Economy*, 127(4), 1777–1825.

Handbury, J. (2019): " Are poor cities cheap for everyone? Non-homotheticity and the cost of living across US cities," *Working Paper*.

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

Kneller, R., and Z. Yu (2016): "Quality Selection, Sectoral Heterogeneity and Chinese Exports," *Review of International Economics*, 24(4), 857–874.

Kugler, M., and E. Verhoogen (2011): "Prices, Plant Size, and Product Quality," *The Review of Economic Studies*, 79(1), 307–339.

Mayer, T., M. J. Melitz, and G. I. P. Ottaviano (2014): "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 104(2), 495–536.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Melitz, M. J., and G. I. P. Ottaviano (2008): "Market Size, Trade, and Productivity," *The Review of Economic Studies*, 75(1), 295–316.

Verhoogen, E. A. (2008): "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *The Quarterly Journal of Economics*, 123(2), 489–530.

Zhelobodko, E., S. Kokovin, M. Parenti, and J.-F. Thisse (2012): "Monopolistic Competition: Beyond the Constant Elasticity of Substitution," *Econometrica*, 80(6), 2765–2784.