Markup Accounting*

Juan Holguín UWO

Sergio Ocampo UWO

Sergio Salgado Wharton

March 26, 2026

Abstract

We document new facts of markup dispersion across the firm size distribution. Markup differences between firms of similar size account for over three-fourths of overall dispersion, calling for mechanisms that affect markups beyond market concentration. To study these mechanisms, we develop an analytical oligopolistic competition model of variable markups that accounts for the observed joint distribution of markups and firm size, including the large mass of small firms with high markups and the presence of large firms with small markups. The key ingredient is the introduction of firm-specific demand elasticity shifters that account for differences in markups between firms of similar size. We apply this model to Indian, Colombian, and U.S. firm data and find consistent results throughout. Without demand elasticity shifters, equilibrium markups collapse to those of standard static models of oligopolistic or monopolistic competition, which capture less than half of the observed variation in markups and counterfactually assigns most of the variation to differences *between* firms of different size. In aggregate, the model implies at least five times larger efficiency losses from markup dispersion relative to standard models of heterogeneous markups.

JEL: D2, D4, E2, L1, O4.
Keywords: Markups, productivity, demand, concentration, oligopoly.

*Holguin: jholguin@uwo.ca.
Ocampo: socampod@uwo.ca; Web: https://sites.google.com/site/sergiocampod.
Salgado: ssalgado@wharton.upenn.edu; Web: https://sergiosalgado.net.

We are especially thankful to Juan Herreño for his involvement in the early stages of this project and to Davesh Raval who kindly shared data on Colombian firms with us. We are grateful for comments from Lucia Casal, Mons Chan, Santiago Franco, Basile Grassi, Rafel Guntin, Mahmood Haddara, Kyle Herkenhoff, Joachim Hubmer, Guido Menzio, Simon Mongey, David Rivers, and Dominic Smith. We also thank Javier Martinez for his assistance in developing this project.

1. Introduction

There is substantial variation in markups across firms within industries. In this paper, we document new facts of the distribution of markups, showing that markup dispersion is concentrated among firms with similar sizes, accounting for over three-fourths of the standard deviation of markups. In particular, using data from India, Colombia, and the U.S. we show that there is a large dispersion in markups among small firms, including firms charging markups well above their market's average, as well as dispersion among large firms, including firms charging below-average markups. These facts call for forces beyond those directly linked to the scale of firms in accounting for markups, such as sales concentration in oligopolistic markets (see, among others, Atkeson and Burstein 2008; Edmond, Midrigan, and Xu 2023) or variable elasticity of demand in monopolistically competitive markets (see, among others, Kimball 1995; Klenow and Willis 2016).

In the second part of the paper, we disentangle the roles of market concentration, firm heterogeneity, and demand factors in explaining markup dispersion by developing an oligopolistic competition model of variable markups with non-CES demand that exactly accounts for the joint distribution of markups and firm sizes. We find a prominent role for demand factors that shift the elasticity of demand faced by firms in accounting for the dispersion of markups, particularly within firm size groups. Without these factors, the model generates a counterfactually low markup dispersion, driven primarily by differences between firms of *different sizes* and not by dispersion between firms of *similar sizes* as in the data.

We model oligopolistic product markets with finitely many firms that differ in their marginal costs, as in Atkeson and Burstein (2008). Firms face demand functions with variable elasticity derived from a "*homothetic direct implicit additivity*" aggregator (Matsuyama 2023) that generalizes the demand function in Kimball (1995). We exploit

the flexibility afforded by this aggregator to introduce idiosyncratic demand elasticity shifters. We solve analytically for equilibrium markups when firms compete strategically. Markups depend jointly on market structure—through firms' market shares and elasticities of demand—and firms' idiosyncratic factors—namely their marginal costs and demand elasticity shifters.

These mechanisms explain two types of firms that standard models struggle to account for: small (or otherwise less productive) firms with high markups, and large firms with low markups—both prominent features of the data. Thus, the two dimensions of firm heterogeneity we model break the standard link between markups, productivity, and firm size in static models of variable markups, and make it possible to account for the entire markup distribution, including differences between firms of the same size.[1]

We estimate markups for Indian, Colombian, and U.S. firms, with different methodologies and use our model to decompose markups into the roles of market concentration, firm scale, and idiosyncratic forces shaping demand elasticities. We find a robust pattern across countries, periods, and methodologies: Over half of the variation in markups between firms comes from variation between firms of similar size as opposed from firms of different sizes charging different markups.

In the first part of our analysis, we follow the bulk of the literature and estimate markups using revenue data at the industry or sector level, as in De Loecker, Eeckhout, and Unger (2020) and Raval (2023). This allows us to include a larger population of firms in our estimates. However, relying on revenue data introduces well-known biases in markup estimation, as shown in Bond, Hashemi, Kaplan, and Zoch (2021). Moreover, using industry classifications to define markets can result in too-broad a market definition, as a given industry contains firms competing in different product markets. This biases measures of market concentration, as discussed in Smith and Ocampo

[1] In theories of variable markups with a single dimension of firm heterogeneity, typically productivity, only high-productivity—low marginal cost—firms can charge low prices that lead to higher scale of production and market shares, both resulting in higher markups.

(2025) in the context of the retail industry, where they compare industry- and product-based markets.

To address these concerns, we rely on Indian manufacturing data from the *Annual Survey of Industries* (ASI). The Indian data is particularly advantageous as it includes the prices and quantities of products being sold and inputs being used. This allows us to overcome biases in markup estimates generated by estimation based only on revenue (Bond, Hashemi, Kaplan, and Zoch 2021) and to define product markets in which firms compete, as opposed to grouping firms by industry or sector. We estimate markups building on the methodology of De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) and Ackerberg, Caves, and Frazer (2015) using data on single-product firms for our estimation of production functions, avoiding well-known input-allocation biases present in multi-product firms.

Our main results describing the role of markup dispersion between firms with similar sizes are preserved, and in fact strengthened, when using this methodology based on price data and markets defined at the product level. The results are also preserved when focusing on single-product firms for which there is no imputation at the time of calculating markups from the production function estimates, and in additional exercises addressing potential measurement error. The robustness of our results on the drivers of markup dispersion between revenue- and price-based measures of markup is consistent with the results of De Ridder, Grassi, and Morzenti (2026), who report that revenue-based estimates bias the level of markups but preserve their dispersion.

Markup dispersion between similarly-sized firms in the data results in a prominent role for idiosyncratic demand elasticity shifters when accounting for the distribution of markups. To show this, we exploit the analytical solution of equilibrium markups together with observed changes in markups and market shares to identify the super-elasticity of demand. The super-elasticity is the relevant measure for how much the elasticity of demand (and hence markups) vary with firms' scale and thus of the degree

of idiosyncratic variation in elasticities beyond the forces embedded in the market aggregator and the market structure through concentration.

To better understand the forces at play, we conduct a decomposition of the observed markup and firm size distribution into the roles of market structure, size, and elasticity shifters. We exploit the fact that the model includes as limiting cases the oligopolistic competition model with constant elasticity of demand of Atkeson and Burstein (2008), in which markups depend only on market concentration, the monopolistic competition model with variable elasticity of demand of Kimball (1995), in which markups depend only on firms' location along a common demand curve, and a hybrid model that combines strategic interactions and variable demand elasticity as in Amiti, Itskhoki, and Konings (2019, Appendix D). These cases differ from our full model in the absence of elasticity shifters.

We evaluate these models based on their ability to replicate the observed joint distribution of markups and market shares, as well as their success in matching the high degree of markup dispersion found within narrow size bins. We find models without elasticity shifters capture less than half of the observed dispersion in markups. These alternative specifications fail to capture the significant mass of small, high-markup firms and counterfactually assign the majority of markup variation to differences between firms of different sizes. Specifically, while the data shows that 77% of markup variance occurs between similarly sized firms, standard models attribute at least two-thirds of the variation to differences across the size distribution. Moreover, the markups implied by these models have a correlation of less than 0.25 with measured markups even when matching the same market-level markups.

The performance of these models in capturing the distribution of markups has direct implications for the aggregate impact of misallocation. To see this, we compare the distance-to-the-efficiency frontier in our model versus alternative specifications following Baqaee and Farhi (2020). We find output losses from markup dispersion in the

product-markets we study of 28 percent when we account for the entire joint distribution of markups and firm size. These costs are five to ten times higher than those implied by standard models. This discrepancy arises because our approach identifies a significant pool of highly productive small firms whose products are harder to substitute—a margin of misallocation that remains largely invisible in models where size and productivity are more strictly linked.

Finally, we also contrast the resulting distributions of markups, productivity, and firm size. Markups are on average higher for more productive firms, in line with standard models of variable markups, but this relationship is not strong, as there is a lot of variation in marginal costs for a given level of markups. Nevertheless, productive firms do not command a higher market share on average.[2] We find that market shares are on average higher for low productivity firms, despite large variation in marginal costs among small firms. These findings are reconciled by the role demand elasticity shifters play in our model in separating markups from market shares and firm scale.

Related Literature. Accounting for the distribution of markups thus requires the presence of large, unproductive firms and small, high-markup firms in equilibrium. We show how this can be achieved by introducing heterogeneity in the demand firms face into an otherwise standard model of oligopolistic competition. In this way, our results complement standard theories of firm dynamics and imperfect competition with a single-dimension of heterogeneity, where market power corresponds to market concentration. The importance of the role demand factors play is in line with results for Chilean exporters as in Blum, Claro, Horstmann, and Rivers (2024), Colombian manufacturing firms as in Eslava, Haltiwanger, and Urdaneta (2023), firm dynamics models based on Melitz (2003) that allow for heterogeneity in demand as Baqaee, Farhi,

[2]See Hubmer, Chan, Ozkan, Salgado, and Hong (2025) and Haddara (2026) for evidence of low productivity among large firms in Canadian administrative data.

and Sangani (2023), and theories of firm dynamics emphasizing competition in the acquisition and lock-in of customers, as in Afrouzi, Drenik, and Kim (2025), Hubmer and Nord (2025), Casal (2026), and Haddara (2026), where differences in productivity are accompanied by differences in the customer base of firms.

Our results also highlight the role of misallocation resulting from the dispersion of markups within firm sizes as opposed to the dispersion of markups between firms of different sizes (already emphasized by Edmond, Midrigan, and Xu 2023, among others). Large firms are not necessarily the most productive, and they coexist with small but productive firms. This complements a long-standing literature on the optimality of markup dispersion (see, for example, Dixit and Stiglitz 1977; Dhingra and Morrow 2019) by highlighting the role of demand, such as consumer behavior leading to markup heterogeneity as in Albrecht, Phelan, and Pretnar (2023) or non-CES preferences as in Zhelobodko, Kokovin, Parenti, and Thisse (2012).

Our paper also contributes methodologically by providing a unified analytical framework that integrates two of the leading workhorse models to study heterogeneous markups. These are models based on oligopolistic competition between firms facing demands with a common constant elasticity (see, among others, Atkeson and Burstein 2008; Edmond, Midrigan, and Xu 2015, 2023; Burstein, Carvalho, and Grassi 2025) and on monopolistic competition between firms facing a common demand curve with variable elasticity (see, among others, Kimball 1995; Dotsey and King 2005; Barde 2008; Behrens, Mion, Murata, and Suedekum 2020; Baqaee, Farhi, and Sangani 2023; Boar and Midrigan 2024; Afrouzi, Drenik, and Kim 2025; Herreño, Pinardon-Touati, and Thie 2025). Our results expand on those in Amiti, Itskhoki, and Konings (2019, Appendix D) by providing analytical formulas for equilibrium elasticities, markups, and aggregate market variables that include the workhorse models as special cases.

Our focus on accounting for the dispersion in markups between firms of similar size—which makes up most of the dispersion in the data—also complements the literature on

markup dispersion arising from search frictions that give market power to firms offer identical products (see, among others, Butters 1977; Varian 1980; Burdett and Judd 1983; Menzio 2024; Chernoff, Head, and Lapham 2024). The sources of markup dispersion driving variation among similarly-sized firms are different from those generated by search frictions that instead lead to markup dispersion linked to firm size, as firms charging lower prices sell with a higher probability.

2. Estimating Product-Level Markups

Our main analysis uses price and quantity data from Indian manufacturing firms to measure firm markups in narrowly defined product-markets. The data comes from the *Annual Survey of Industries* (ASI). Crucially for our purposes, the data not only contains information on firm sales and input expenditure, but also on output and input prices, distinguishing between sales of different products for multi-product firms. We use price data to estimate quantity-based production functions that overcome many of the biases introduced when estimating markups only on revenue data (e.g., Bond, Hashemi, Kaplan, and Zoch 2021).

The ASI follows manufacturing establishments, with large establishments surveyed annually and small establishments included on a sampling basis.[3] The complete dataset is for the years 2001 to 2008 and has a total of 340,129 establishment-product-year observations with an average of 24,245 unique firms per year. We further divide our data into two samples, one from 2001 to 2004 and one from 2004 to 2008. For space considerations, we focus on results for the second sample in Section 4, but we take into account both samples when accounting for the distribution of markups at the end of this section and in Section 5.

[3]Size is determined based on the number of workers. Between 2001 and 2003, all establishments with more than 200 workers were included; afterward, establishments with more than 100 workers were included. Small establishments are those not below these thresholds.

We also use data from the Colombian *Encuesta Anual Manufacturera* (EAM) and from US public firms for additional results. The Colombian markup estimates come from Raval (2023) and cover manufacturing firms from 1980 to 1989. The final dataset includes 2,551 unique establishments and 18,035 total establishment-year observations. For Colombia, we define markets at the 3-digit ISIC (Rev. 2) level. The data on US public firms come from De Loecker, Eeckhout, and Unger (2020). We use the subsample from 1980 to 1989, which includes 790 unique firms and 5,192 total observations. We define markets at the 3-digit NAICS industry level.

Data selection at the product level. In our main analysis, which uses data from India. We focus on *product markets* where manufacturing firms are most likely to compete and where firms are most likely to share a common production technology. This choice reduces the measurement error of market shares, as it provides a more precise definition of competition than alternative definitions at the sector or industry level. It also reduces measurement error in the estimation of production functions and markups, for which we use single-product producers within a sector that produce similar goods.

We define markets at the 5-digit product level according to the ASI Commodity Classification (ASICC). There are 5073 unique product codes between 2001 and 2008. These correspond to products such as Plywood, Cotton Yarn, Shoe Leather, Detergent, and Sugar. To keep our analysis tractable, we focus (when possible) on the two largest 5-digit products in each of the two main 3-digit sectors of each 2-digit industry in India (sectors and industries are defined following the 2- and 3-digit subdivisions of ASICC classification). This procedure leaves us with 22 product markets in 14 3-digit sectors; Table A.1 lists these products and their sectors. Defining markets at the product level instead of the sectoral or industrial level avoids grouping products like Sugar and Lactose, or Pesticides and Plant Hormones, that share the same industry but do not compete in the same markets. This is similar to what happens in the US retail market, as described

in Smith and Ocampo (2025).

The market share of the selected sectors within their corresponding industries ranges from 13% to 74%, while the market share of the top two products within their sectors ranges from 10% to 83%. On average, each product market contains about 970 establishment–product–year observations, and each 3-digit sector contains 2,865 establishment–product–year observations. For a complete overview, see Appendix A.1.

Production function estimation. We estimate quantity-based production functions at the sector-level assuming that products of the same 3-digit sector are produced using the same technology, regardless of their industry. For example, there is a common technology for Timber Planks and Plywood (sector 511), and a different technology for Pesticides and Insecticides (sector 344). Our production function estimation is therefore capturing the way inputs are used for the production of similar products, avoiding mixing between products in the same industries that are nevertheless produced in different manners—as is the case with Biscuits and Refined Sugar, both in industry 13 Food and Edible Products and Preparations but in different sectors.

We estimate these production functions building on the work of De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) and Ackerberg, Caves, and Frazer (2015). We use measures of establishment level quantities and prices of capital, labor, and intermediate material inputs contained in the ASI. In our baseline specification, we assume that capital and labor inputs are predetermined at the time of production, while intermediate material inputs are flexible in accordance to standard practice in the literature (see, for example, Gandhi, Navarro, and Rivers 2020). We obtain similar results when assuming that labor is also a flexible input.

A central challenge for the estimation is that we do not observe product-level input allocations for multi-product firms. We avoid this input-allocation bias by estimating production functions using only single-product firms for which input allocations are

known.[4] To control for unobserved input quality, we employ a control function as in De Loecker et al. (2016) exploiting the observations of output prices in our data. We effectively assume higher-quality products command higher prices and must be produced with higher-quality inputs. This creates a mapping from observed output prices to unobserved input prices, allowing us to account for quality differences. Finally, we include state fixed effects to control for geographic differences in input prices.

Markup measures. We use our estimates of quantity-based production functions to recover markups for all establishments in the products markets we study. We define markups as the ratio of prices to marginal costs, $\mu \equiv p/\lambda$. We do so by exploiting the optimal input choices that come from the cost minimization problem,

$$C\left(y|\left\{p_n\right\}_{n=1}^{N},\left\{K_h\right\}_{h=1}^{H}\right)=\min_{\{x_n\}_{n=1}^{N}}\sum_{n=1}^{N}p_n\cdot x_n \quad \text{s.t. } \overline{y}\leq zF\left(x_1,\ldots,x_N,K_1,\ldots,K_H\right), \quad (1)$$

where production requires variable inputs $\{x_n\}_{n=1}^{N}$ and fixed inputs $\left\{K_h\right\}_{h=1}^{H}$. The marginal cost of production is captured by the Lagrangian multiplier (λ) associated with the production constraint, as $\lambda = C'(y)$.

Simple manipulation of the first-order condition with respect to variable input x_n $\left(p_n = \lambda \partial F/\partial x_n\right)$ results in an expression of the markup in terms of the cost-to-revenue share of the input, $s_n \equiv p_n x_n/py$, and the output elasticity with respect to x_n, ϵ_n:

$$\mu = \frac{p}{\lambda} = \frac{py}{p_n x_n}\epsilon_{x_n} = \frac{\epsilon_{x_n}}{s_n}. \quad (2)$$

[4] Our procedure assumes that multi-product firms employ the same technology for each product, an approach that precludes any within-firm production complementarities. To maximize the number of single-product firms within each sector we complement the sample of single-product firms with those that produce 5-digit product identified as co-produced with the products in our main sample. These are products that are often produced together by multi-product firms and are hence likely to share a common production function. We estimate using data from 43,300 establishment-year observations obtaining separate estimates for 14 3-digit sectors. See Appendix A.2 for details.

This condition is entirely compatible with the demand system and the optimal pricing behavior we describe in Section 3. Our data provide us with direct measures of input shares, s_n, and we obtain the elasticity of output with respect to variable inputs, ϵ_{x_n}, from the estimation of quantity-based production functions. Together, these estimates allow us to obtain product-level markups for all the firms in our sample.[5]

In additional results using Colombian and US data, we use alternative approaches to measure the elasticity of output with respect to variable inputs ϵ_{x_n} and recover markups as in equation (2). For Colombian data, we follow Raval (2023) and exploit variation in cost shares under the assumption that production functions exhibit constant-returns-to-scale. For the US data, we follow De Loecker, Eeckhout, and Unger (2020). Both of these approaches rely on revenue data, as neither dataset contains information on prices. This introduces well-known biases described in Bond, Hashemi, Kaplan, and Zoch (2021). Nevertheless, our main results also hold under these alternative approaches—namely the primary role of within firm-size-groups markup variation and the importance of demand elasticity shifters. The robustness of our results is also in line with the findings of De Ridder, Grassi, and Morzenti (2026), who show that revenue-based markup estimates are biased on average but present the same properties in terms of their variability.

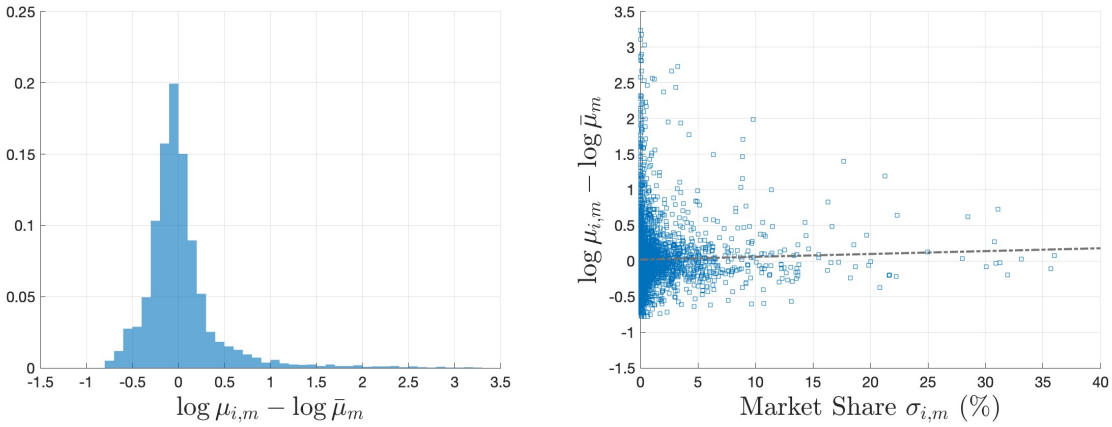
2.1. Markup dispersion: Between and within firm size

Before turning to our modeling of markup determination, we take a first look at the estimated distribution of markups. Figure 1 shows the distribution of log-markups relative to their market's average, computed as a sales-weighted harmonic mean of product-year markups (Edmond, Midrigan, and Xu 2015; Smith and Ocampo 2025;

[5]We assume multi-product firms allocate inputs proportionally to each product's within-firm revenue share, as we do not directly observe input allocations in these firms. For robustness, we also calculated the input allocations generated by the De Loecker et al. (2016) method. We find a correlation close to 0.7 between the revenue shares and the input allocations generated by their model.

FIGURE 1. Markups and Market Share Distribution in Indian Product Markets

(A) Markups Distribution (B) Markups and Market-Share



Notes: Panel A shows the histogram of markups relative to their market's average for the Indian data covering the period 2001–2007. Panel B shows the joint scatter plot of markups and market shares. Markets are defined at the 5-digit ASICC product level.

Hasenzagl and Pérez 2023),

$$\mu_m \equiv \left(\sum_{i=1}^{N_m} \frac{\sigma_i^m}{\mu_i^m} \right)^{-1} . \tag{3}$$

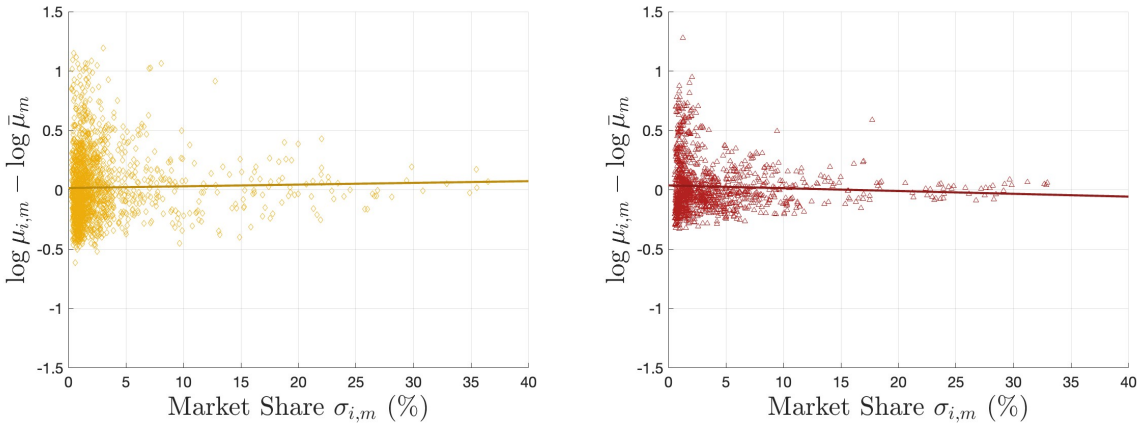
The (sales-weighted) average standard deviation of markups is 0.38 log-points.

Panel B shows this dispersion reflects both within- and between-size dispersion in markups.[6] In fact, the figure clearly shows large variation in markups between firms of similar size, particularly among small firms (establishments) that even have the largest measured markups. Conversely, firms with a large market share often have markups below their market's average. This is reflected in the correlation between markups and market shares, which is only 0.06 as shown in Table 1. The same patterns are present in other countries and when measuring markups using other methodologies, as we show in Figure 2, where we plot the distributions of markups and market shares for Colombia

[6]The analysis in Section 3 motivates our use of market sales-shares to measure firm size.

FIGURE 2. Markups and Market Share Across Countries

(A) Colombia (B) United States



Notes: Panel A shows the joint scatter plot of markups and market shares in Colombia. Markets are defined at the 3-digit ISIC rev 2 classification. Panel B shows the joint scatter plot of markups and market shares in US. Markets are defined at the 3-digit NAICS classification.

and the United States. Figure B.1 in Appendix B shows the marginal distribution of markups in each dataset.

To get an idea of the role markup variation between similarly-sized firms, we divide observations in each product-market according to their market share in equally sized bins of 5 percentage points and decompose the variance of log-markups into their within-bin and between-bin components. Table 1 presents the average results in the markets for our main Indian sample, as well as for Colombia and the United States. The variation in markups within bins of similar market share accounts on average for 77 percent of the total variance of markups in product-markets (74 percent in Colombia and 67 percent in the United States).

The share of variation of markups concentrated between similarly-sized firms is still substantial when we divide observations into 1 percentage point market share bins, with within-bin variation accounting for 56 percent of markup variation; see Table B.1 in Appendix B. The same pattern emerges when we average measured observed markups

TABLE 1. Markup Variance Decomposition

	Variance	Share within	Share between	Corr($\log \mu_i$, σ_i)
India	0.196	0.770	0.230	0.061
Colombia	0.059	0.739	0.261	0.030
United States	0.024	0.676	0.324	-0.075

Notes: The variance is computed from the log ratio $\log(\mu_i^m/\mu_m)$. Numbers are averages across markets. To decompose the variance of markups, we divide firms in each market into equally sized bins of 5 percentage point market share. The variance and its decomposition correspond to the average variance and shares across market-year pairs. The correlation between markups and market shares is the pooled correlation of the full sample.

at the firm level over time, as a way to reduce potential measurement error, and when we reproduce our exercise with the markups of single-product firms for which there is no imputation in the allocation of inputs, also reducing potential measurement error. Depending on the sample restrictions, the within-bin component accounts for at least 40 percent of the total variation in average firm markups; see table B.5.

Thus, accounting for the markup distribution requires accounting for the presence of a large number of firms with low market shares and high markups, as well as of larger firms with below-average markups. This points to a role for markup differences across the firm size distribution different from the one implied by standard theories of variable markups and imperfect competition, where firms' state is uni-dimensional resulting in firms with large market shares charging high markups and firms with low market shares charging low markups. We turn to this next.

3. Model Setup

The model economy contains M product markets. Each market contains finitely many firms, each producing a differentiated good of a product type. These firms are

monopolists in their good variety and compete oligopolistically either *à la* Cournot or *à la* Bertrand as in Atkeson and Burstein (2008). The goods of each market are aggregated according to a Homothetic Direct Implicit Additivity type aggregator that results in residual demand curves with variable elasticity (see Matsuyama 2023, Sect. 7). A single representative consumer demands the final consumption good (the aggregate of product output). We present the main features of the model in the main text and provide all the proofs in Appendix C.

3.1. Final good production

Final-good producers behave competitively and operate a CES technology. Their problem is

$$\min_{\{Y_m\}} \sum_{m=1}^{M} P_m Y_m \qquad \text{s.t. } Y \leq \left(\sum_{m=1}^{M} \alpha_m Y_m^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}, \tag{4}$$

where $\sum_m \alpha_m = 1$ and $\gamma \geq 1$. The demand for the output of market m satisfies

$$\frac{P_m}{P} = \alpha_m \left(\frac{Y_m}{Y} \right)^{-\frac{1}{\gamma}}, \quad \text{where } P = \left(\sum_{m=1}^{M} \alpha_m^{\gamma} (P_m)^{1-\gamma} \right)^{\frac{1}{1-\gamma}}. \tag{5}$$

3.2. Market good production and demand for individual firms

There are N_m firms operating in each product market. The market's output, Y_m, is an aggregate of differentiated goods produced by these firms, $\left\{ y_i^m \right\}$. The technology of competitive market aggregators is implicitly defined by

$$1 = \sum_{i=1}^{N_m} \Upsilon_i^m \left(\frac{y_i^m}{Y_m} \right). \tag{6}$$

This is a more general form of the aggregator used in Kimball (1995) and Edmond, Midrigan, and Xu (2023), among others, in that it allows for firm-specific functions $\Upsilon_i(\cdot)$.[7] This distinction is crucial for our objectives, as it allows us to introduce idiosyncratic demand elasticity shifters that affect the elasticity of demand of any given good. These shifters allow small firms to charge high markups and large firms to charge low markups, introducing a new force that interacts with market concentration and firm scale in equilibrium to determine markups.

Although most of our results hold for a general function Υ_i, we later parameterize them for estimation. We follow the literature by focusing on two-parameter functional forms so that $\Upsilon_i(x) = \Upsilon\left(x; \nu_i^m, \theta_m\right)$, but we let the parameters $\left\{\nu_i^m\right\}$ vary by firm, allowing us to capture idiosyncratic differences in the elasticity of demand of each differentiated good. The remaining parameter, $\{\theta_m\}$, is constant for each market. We return to these parameters in Section 4.[8]

The problem of an aggregator is to minimize expenditure $\sum_i p_i^m y_i^m$ subject to (6) and a desired level of output Y_m, where p_i^m and y_i^m are the price and quantity of good i in market m. Not all goods are necessarily demanded in equilibrium. The solution of this problem implies a demand for variety i characterized by

$$\frac{p_i^m}{P_m} = \frac{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}{\sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}}, \tag{7}$$

where P_m is market m' ideal price index, that is, $P_m Y_m = \sum_i p_i^m y_i^m$. We provide details for these derivations in Appendix C.1.

[7] Baqaee, Farhi, and Sangani (2023) also consider this type of aggregator in a setting without strategic interactions between firms.

[8] An alternative parameterization, offered in Baqaee, Farhi, and Sangani (2023), imposes conditions on equilibrium market shares, $\sigma_i^m \equiv p_i^m y_i^m / P_m Y_m$, as a function of relative prices, $\sigma_i^m = s\left(\nu_i^m p_i^m / P_m\right)$. The parameterizations differ on their implications for how shifters $\left\{\nu_i^m\right\}$ affect markups.

The homotheticity of the demand system imposes restrictions on how the market's output and price react to changes in firms' output and prices. These restrictions shape the behavior of equilibrium market shares. We highlight these implications in the following lemma (see also Matsuyama 2023, Sect. 5.2).

Lemma 1. ***(Market Shares)*** *The change in market output Y_m to firm i's output and of market price P_m to firm i's price satisfy*

$$\frac{\partial Y_m}{\partial y_i^m} = \frac{p_i^m}{P_m} \qquad \text{and} \qquad \frac{\partial P_m}{\partial p_i^m} = \frac{y_i^m}{Y_m}. \tag{8}$$

So, the market share of firm i in market m, σ_i^m, gives the elasticity of market output and market price to changes in the firm's output and price, respectively,

$$\sigma_i^m \equiv \frac{p_i^m y_i^m}{P_m Y_m} = \frac{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}{\sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}} = \frac{y_i^m}{Y_m}\frac{\partial Y_m}{\partial y_i^m} = \frac{p_i^m}{P_m}\frac{\partial P_m}{\partial p_i^m}. \tag{9}$$

3.3. Firm's problem: Profit maximization

The problem of firm i in market m is to maximize its profits, taking as given the demand schedule for its goods in (7) and the actions of other firms. The firms' problem is

$$\max \; p_i^m y_i^m - C_i\left(y_i^m\right) \qquad \text{s.t. (5), (6), and (7);} \tag{10}$$

where we take as given the firm's optimal input demand choices so that the cost of production is described by an increasing function C_i, as discussed in Section 2. The firm maximizes over its price p_i^m if competition is *à la* Bertrand, or over its quantity y_i^m if competition is *à la* Cournot. In either case, the firm optimally chooses to set a

markup μ_i^m over its marginal cost so that

$$p_i^m = \underbrace{\frac{1}{1-\frac{1}{\eta_i^m}}}_{\text{Markup: } \mu_i^m} C_i'\left(y_i^m\right), \tag{11}$$

where $\eta_i^m \equiv -\left(\frac{\partial \log p_i^m}{\partial \log y_i^m}\right)^{-1}$ is the elasticity of demand and μ_i^m the markup.

The role of the demand system and the nature of competition between firms are therefore captured by the firms' elasticities of demand, $\left\{\eta_i^m\right\}$. When a firm changes its output or price, it affects its own demand, the aggregate market output, and the demand faced by other firms. Crucially, firms internalize how their actions affect other firms' demand and, in turn, their own elasticity. These effects correspond to *(i)* the firm's own-elasticity (taking aggregates and the behavior of other firms as given), *(ii)* the granular effect of the firm on market output Y_m, and *(iii)* the substitution effects across firms that depend on their elasticities.

For instance, under Cournot competition, where firms take each other's output as given, we can write the elasticity of demand using (5) and (7) as[9]

$$\left(\eta_i^m\right)^{-1} = \underbrace{\left(-\frac{\partial \log \Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}{\partial \log y_i^m}\right)}_{\text{Own-Elasticity}} + \underbrace{\left(-\frac{\partial \log \left(\frac{Y_m}{Y}\right)^{-\frac{1}{\gamma}}}{\partial \log y_i^m}\right)}_{\text{Granular Effect}} - \underbrace{\left(\frac{\partial \log \sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_i^m}{Y_m}}{\partial \log y_i^m}\right)}_{\text{Substitution Effects}}. \tag{12}$$

The elasticity of demand under Bertrand competition can be similarly expressed as a function of three terms capturing the firm's own-elasticity, its granular and substitution effects, see equation (C.33). In both cases, the effect on the firm's demand comes from

[9] We assume throughout that firms internalize the effect of their actions on their market's output and price, but not on the aggregate output and price, Y, P. This is the same assumption used in Atkeson and Burstein (2008) where there is a continuum of markets, or in Grassi (2017) and Burstein, Carvalho, and Grassi (2025) where they have finitely many markets.

two sources. First, a change in market output changes the level of demand; larger firms can take more advantage of this effect, as they have a larger impact on sectoral output. Second, a change in market output changes the elasticity of demand of all firms. If the aggregator Υ_i is concave, a higher relative output implies a lower elasticity of demand, and firms facing lower elasticity charge higher markups.

Under monopolistic competition, only the first term emerges as the firm has no effect on the aggregate market output, and therefore no effect on other firms. It is useful to then define the *own-elasticity of demand*—the elasticity of demand for the variety produced by firm i holding the aggregate market output constant—as

$$\varepsilon_i^m \equiv \left(-y_i^m \frac{\partial \log \Upsilon_i' \left(\frac{y_i^m}{Y_m} \right)}{\partial y_i^m} \right)^{-1} = -\frac{\Upsilon_i' \left(\frac{y_i^m}{Y_m} \right)}{\frac{y_i^m}{Y_m} \Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right)}. \tag{13}$$

This is the first term in (12) and corresponds to the measure of elasticity in Kimball (1995), Klenow and Willis (2016), Edmond, Midrigan, and Xu (2023), and many others.

The rate at which the firm's own-elasticity of demand changes plays a crucial role in the estimation of the model in Section 4. This rate depends on the super-elasticity of demand, which measures how much the elasticity changes with output and thus captures the shape of the function Υ_i. The super-elasticity of demand is

$$\xi_i^m \equiv -\frac{p_i^m}{P_m} \frac{\partial \log \varepsilon_i^m}{\partial \left(\frac{p_i^m}{P_m} \right)} = 1 + \varepsilon_i^m + \varepsilon_i^m \frac{\frac{y_i^m}{Y_m} \Upsilon_i''' \left(\frac{y_i^m}{Y_m} \right)}{\Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right)}. \tag{14}$$

Finally, the elasticity of substitution between products i and j in market m is captured

by the Allen-Uzawa elasticity as in Baqaee and Farhi (2020),

$$\varphi_{ij}^{m} = \frac{\partial \log y_j^m / \partial \log p_i^m \big|_{Y_m}}{\sigma_j^m} = \begin{cases} \frac{\varepsilon_i^m \varepsilon_j^m}{E_\sigma[\varepsilon_h]} & \text{if } j \neq i \\ -\frac{\varepsilon_i^m}{\sigma_i^m}\left(1 - \frac{\sigma_i^m \varepsilon_i^m}{E_\sigma[\varepsilon_h]}\right) & \text{if } j = i \end{cases} \tag{15}$$

where $E_\sigma\left[x_h\right] = \sum_h^{N_m} x_h \sigma_h^m$ is the sales-weighted average.

3.4. Equilibrium markups

We solve analytically for the firms' elasticity of demand and, therefore, their markups. The result is a characterization of the firm's elasticity, η_i^m, as a sales-weighted average of the elasticity of market output, γ, and the elasticity of demand for the firm's variety inside the market. This *variety elasticity* is itself a function of the firm's own-elasticity, ε_i^m, and the own-elasticities of its competitors. The form of competition (Bertrand or Cournot) determines how competitors affect the firm's elasticity of demand.

Proposition 1. ***(Elasticities)*** *The elasticity of firm i in market m satisfies*

$$\textbf{\textit{Cournot}} \quad \frac{1}{\eta_i^m} = \underbrace{\frac{1}{\gamma}}_{\text{Market Elasticity}} \sigma_i^m + \underbrace{\left(\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right) + E_\sigma\left[\frac{1}{\varepsilon_j^m}\middle| j \neq i\right]\sigma_i^m\right)}_{\text{Variety Elasticity}}\left(1-\sigma_i^m\right) \; ; \tag{16}$$

$$\textbf{\textit{Bertrand}} \quad \eta_i^m = \underbrace{\gamma}_{\text{Market Elasticity}} \sigma_i^m + \underbrace{\varepsilon_i^m \frac{E_\sigma\left[\varepsilon_j^m | j \neq i\right]}{E_\sigma\left[\varepsilon_j^m\right]}}_{\text{Variety Elasticity}}\left(1-\sigma_i^m\right) \; ; \tag{17}$$

where $\sigma_i^m = p_i^m y_i^m / P_m Y_m$ is firm i's market share, ε_i^m is its own-elasticity as in (13), $E_\sigma\left[x_j\right] = \sum_j^{N_m} x_j \sigma_j^m$ is the average with respect to expenditure in market m, and $E_\sigma\left[x_j | j \neq i\right] = \sum_{j \neq i} x_j \frac{\sigma_j^m}{1-\sigma_i^m}$ is the average with respect to firm i's competitors.

The characterization of the elasticity in Proposition 1 makes clear the role of market

concentration. As a firm increases its market share, its elasticity reflects more the market's elasticity and less the elasticity of the firm's variety. When firms are granular, they care more about the demand for their market's aggregate output the larger they are. Similarly, larger firms are less affected by the elasticity of their own variety and weight more the elasticity of their competitors, as that becomes the main source of changes in their demand. Crucially, the Proposition also makes it clear that even though firms care about the elasticity of their competitors, these only matter through a sufficient statistic defined as an appropriate average of the competitors' elasticities.

Proposition 1 leads to analytical expressions for the firms' markups in terms of market shares and the set of own-elasticities of demand, $\left\{\sigma_i^m, \varepsilon_i^m\right\}$. We summarize the results in Proposition 2. Its proof follows from manipulating the elasticities in Proposition 1 along with equation (11).

Proposition 2. ***(Markups)*** *The equilibrium markups under Cournot competition are:*

$$\textbf{\textit{Cournot Markup}} \quad \frac{1}{\mu_i^m} = \frac{\gamma - 1}{\gamma} + \left(\frac{1}{\gamma} - \frac{1}{\varepsilon_i^m}\right)\left(1 - \sigma_i^m\right) + \sigma_i^m \left(\frac{1}{\varepsilon_i^m} - E_\sigma\left[\frac{1}{\varepsilon_j^m}\right]\right) \; ; \tag{18}$$

$$\textbf{\textit{Bertrand Markup}} \quad \frac{1}{\mu_i^m} = 1 - \frac{1}{\gamma\sigma_i^m + \varepsilon_i^m\left[1 - \frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]}\sigma_i^m\right]} \; ; \tag{19}$$

where ε_i^m is firm i's own-elasticity of demand as in (13), $\sigma_i^m \equiv p_i^m y_i^m / P_m Y_m$ its market share, and $E_\sigma\left[x_j\right] = \sum_{j=1}^{N_m} x_j \sigma_j^m$ is the average with respect to expenditure in market m, in this case giving an harmonic mean of the elasticities, ε_j^m.

Remark The Bertrand and Cournot markups converge to the monopolistic competition markups $\left(\mu_i^m = \frac{\varepsilon_i^m}{\varepsilon_i^m - 1}\right)$ as σ_i^m tends to 0, where elasticities vary according to (13). Similarly, they converge to the oligopolistically competitive markups of Atkeson and Burstein (2008) when the own-demand elasticity is constant and common across firms, $\varepsilon_i^m = \varepsilon^m$.

3.5. Aggregation: Market level variables

The average markup in market m, μ_m, is defined as the ratio between the market's price P_m and the market's marginal cost λ_m, $\mu_m \equiv P_m/\lambda_m$.[10] Because the market aggregators (6) have constant-returns-to-scale, each market's marginal cost is equal to the output-weighted average of the individual marginal costs, that we label $\lambda_i^m \equiv C_i'\left(y_i^m\right)$,

$$\lambda_m = \sum_{i=1}^{N_m} \lambda_i^m \frac{y_i^m}{Y_m}. \tag{20}$$

Then, the market's markup is the sales-weighted harmonic mean of individual markups

$$\mu_m = \left[\sum_{i=1}^{N_m} \lambda_i^m \frac{y_i^m}{P_m Y_m}\right]^{-1} = \left[\sum_{i=1}^{N_m} \frac{1}{\mu_i^m} \sigma_i^m\right]^{-1} = \left[\sum_{i=1}^{N_m} \left(1 - \frac{1}{\eta_i^m}\right) \sigma_i^m\right]^{-1} = \frac{1}{1 - \frac{1}{\eta_m}}. \tag{21}$$

This is the same expression as for the individual markup μ_i^m in (11), but replacing firm i's elasticity η_i^m with the corresponding market elasticity,

$$\eta_m \equiv \left[\sum_{i=1}^{N_m} \frac{1}{\eta_i^m} \sigma_i^m\right]^{-1}. \tag{22}$$

We can therefore express market markups in terms of the underlying own-elasticities and market shares using the results in Proposition 1.

Proposition 3. ***(Market Elasticities and Markups)*** *Market elasticities satisfy*

$$\textbf{\textit{Cournot}} \qquad \frac{1}{\eta_m} = HHI_m \frac{1}{\gamma} + (1 - HHI_m)\, E_\sigma\left[\frac{1}{\varepsilon_i^m}\right] - 2Cov_\sigma\left(\sigma_i^m, \frac{1}{\varepsilon_i^m}\right); \tag{23}$$

$$\textbf{\textit{Bertrand}} \qquad \frac{1}{\eta_m} = E_\sigma\left[\frac{1}{\varepsilon_i^m}\right] + \frac{HHI_m}{E_\sigma\left[\varepsilon_i^m\right]} \tag{24}$$

[10] Equivalently, Edmond, Midrigan, and Xu (2023) define it as the ratio of the market's revenue $P_m Y_m$ and the market's total cost, which, under constant returns-to-scale, is $\sum_i \lambda_i^m y_i^m$.

$$- \gamma \left(HHI_m E_\sigma \left[\left(\frac{1}{\varepsilon_i^m} \right)^2 \right] + Cov_\sigma \left(\sigma_i^m, \left(\frac{1}{\varepsilon_i^m} \right)^2 \right) \right) + \Omega$$

where $HHI_m = \sum_i \left(\sigma_i^m \right)^2$ *is the Herfindahl-Hirschman index,* $E_\sigma \left[x_j \right] = \sum_{j=1}^{N_m} x_j \, \sigma_j^m$ *is the expectation with respect to sales shares and* $Cov_\sigma \left(x_j, y_j \right) = \sum_{j=1}^{N_m} \left(x_j \right) \left(y_j - E_\sigma \left[y_j \right] \right) \sigma_j^m$ *is the covariance with respect to sales in the market. Finally,* $\Omega \equiv \sum_i^{N_m} \frac{\sigma_i^m}{\varepsilon_i^m} \left[\sum_{k=2}^{\infty} \left(\sigma_i^m \left(\frac{\varepsilon_i^m}{E_\sigma \left[\varepsilon_j^m \right]} - \frac{\gamma}{\varepsilon_i^m} \right) \right)^k \right]$ *contains higher-order moments.*

The average markup satisfies

$$\textbf{Cournot} \quad \frac{1}{\mu_m} = \underbrace{\left(1 - \frac{1}{\gamma} \right)}_{\text{Monopoly Markup}} + \underbrace{\left(\frac{1}{\gamma} - E_\sigma \left[\frac{1}{\varepsilon_i^m} \right] \right) (1 - HHI_m)}_{\text{Concentration}} + \underbrace{2 Cov_\sigma \left(\sigma_i^m, \frac{1}{\varepsilon_i^m} \right)}_{\text{Distribution}} ; \qquad (25)$$

$$\textbf{Bertrand} \quad \frac{1}{\mu_m} = \underbrace{\left(1 - \frac{1}{\gamma} \right)}_{\text{Monopoly Markup}} + \underbrace{\left(\frac{1}{\gamma} - E_\sigma \left[\frac{1}{\varepsilon_i^m} \right] \right) + \frac{HHI_m}{\gamma} \left(E_\sigma \left[\left(\frac{\gamma}{\varepsilon_i^m} \right)^2 \right] - \frac{\gamma}{E_\sigma \left[\varepsilon_i^m \right]} \right)}_{\text{Concentration}}$$
$$+ \underbrace{Cov_\sigma \left(\sigma_i^m, \left(\frac{1}{\varepsilon_i^m} \right)^2 \right)}_{\text{Distribution}} - \Omega \, . \qquad (26)$$

These results show how market markups come from three sources. This is more easily seen under Cournot competition. First, there is a monopoly markup. That is, the markup that would arise if the market were a monopoly. Second, the markup increases with concentration (as measured by the Herfindahl-Hirschman Index) when the average elasticity of varieties in the market is higher than the market's demand reflecting varieties being more substitutable among themselves than products, $E_\sigma \left[1/\varepsilon_i^m \right] < 1/\gamma$. This is the same force as in Edmond, Midrigan, and Xu (2015). Finally, the dispersion in elasticities introduces a third source of markups coming from the

covariance between market shares and the own-elasticity of varieties in the market. This source reduces average markups when sales are concentrated in firms with a low ε_i^m. This is because these large firms care more about the elasticity of the market's demand (Y_m) relative to smaller firms. In this way, it is the small (niche) firms who increase average markups when their varieties are less elastic, whereas larger firms respond more to the market's elasticity.

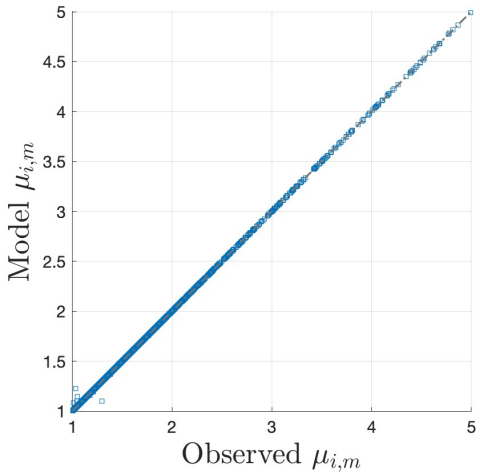
Under Bertrand competition, the higher-order moments of the joint distribution of market shares and elasticities also matter, as captured by the term Ω. These moments capture how concentration interacts with the elasticities of demand and are most relevant in highly concentrated markets, as they depend on higher-order concentration indexes of the type developed in Hannah and Kay (1977). We present more details on these higher-order terms in equation C.62. Crucially, these terms are present even when there are no differences in elasticities between firms, as shown in Grassi (2017, Proposition 4), something we verify in Corollary A1.

4. Accounting for the Distribution of Markups

We now account for the distribution of firm markups and firm sizes in Figure 1. We proceed as follows: First, the optimal pricing choices of firms in Proposition 2 allow us to recover own-demand elasticities that match the joint distribution of market shares and markups. This step does not depend on the functional form imposed on the demand aggregator Υ_i. We then estimate demand parameters after imposing a parametric form for Υ_i following Klenow and Willis (2016).[11] Finally, we recover relative output, prices, and marginal costs. For space considerations, we present results only for Bertrand competition on a sample covering 2005–2008. See Appendix B.3 for additional results under Bertrand competition and Appendix B.4 for results under Cournot competition.

[11] We focus on two families of aggregators presented in Dotsey and King (2005) and Klenow and Willis (2016). Appendix E shows these families encompass all the functional forms used in the literature.

FIGURE 3. Markup Fit in Indian Data (2005–2008)



Notes: The figure shows the scatterplot of measured and model-implied markups under Bertrand competition, obtained by solving for $\{\varepsilon_i^m\}$ in equation 19 given $\{\mu_i^m, \sigma_i^m\}$. We set $\gamma = 1.3$. A value in the 45 degree line corresponds to a perfect match of the model for a given firm.

4.1. Recovering own-demand elasticities

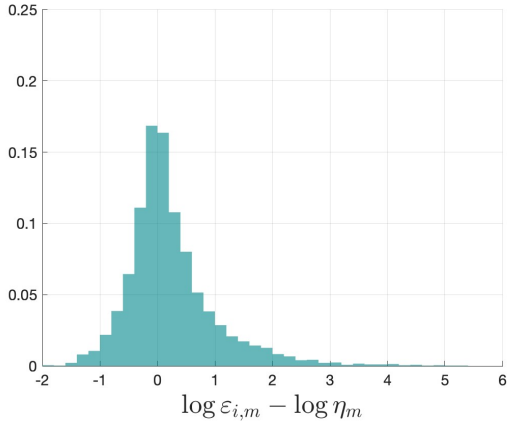
We use the analytical expressions for markups in Proposition 2 to recover the own-elasticities of demand implied by firms' market shares and markups $\left\{\sigma_i^m, \mu_i^m\right\}$. We recover the own-elasticities by minimizing the distance between measured and model-implied. This amounts to solving a system of non-linear equations on $\left\{\varepsilon_i^m\right\}$.[12] In doing this we set the elasticity of substitution between the output of different markets in (4) to $\gamma = 1.3$, consistent with the values used in Edmond, Midrigan, and Xu (2023). Figure 3 plots measured and model-implied markups. The model matches the distribution of markups and market shares almost exactly.

There is substantial variation in own-elasticities of demand, in line with the large variation in markups that generate them, see Figure 4. For instance, a log-difference of 2 implies an elasticity 7.4 times the market average, this goes up to 20 times for

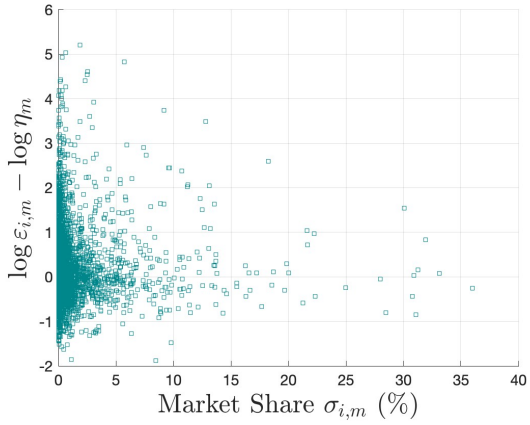
[12]We show in Appendix F.1 that the system is actually linear under Cournot competition and that it can be solved exactly if $1/\bar{\varepsilon}_i^m > 1/\gamma\,\sigma_i^m$ for all firms, with $1/\bar{\varepsilon}_i^m = 1 - 1/\mu_i^m$. Under Bertrand competition the system is nonlinear but it admits a unique solution and can match markups exactly in many of our samples.

FIGURE 4. Distribution of Elasticities (2005–2008)

(A) Own-Elasticity Distribution



(B) Elasticity and Market-Share



Notes: Panel A shows the histogram of own-elasticity of demand relative to their market's average for the Indian data covering the period 2005–2007. Panel B shows the joint scatter plot of own-elasticities and market shares. Markets are defined at the 5-digit ASICC product level.

a log-difference of 3. This large dispersion in the firms' own-elasticity of demand is present across the firm size distribution, as Figure 4B clearly shows. Even among large firms, that tend to have lower elasticities, some exhibit higher-than-average elasticities ($\log \varepsilon_i^m - \log \eta_m \approx 2$). The dispersion is much larger among small firms. In particular, there are many small firms with lower-than-average elasticities, which is necessary to match the observed dispersion in markups and the presence of high-markup firms with small market shares.

On a technical note, although the value of the elasticities of demand does not depend on the functional form of Υ_i, it does impose conditions on it. In particular, it must be that the demand firms face exhibit a variable super-elasticity of demand in order to be compatible with the distribution of market shares and markups. A constant super-elasticity of demand implies a one-to-one link between elasticities and market shares, which is generically rejected by the data. See Proposition A3 in Appendix F.2

4.2. Estimation of demand parameters

We estimate the shape of the demand faced by firms, captured by the functions Υ_i, using variation in firms' size and elasticity over time. The characteristic feature of homothetic demand aggregators with variable elasticity is that the elasticity varies with firm size (measured by the firms' relative output, y_i^m/Y_m). We use the analytical characterization of equilibrium markups to extend this relationship to one between a change in the firm's observed market share and the firm's elasticity of demand that does not depend on the functional form of Υ. Instead, the change is captured by the super-elasticity of demand, ξ_i^m, the relevant measure of how much the own-elasticity of demand changes as the firm changes in size. We present the details of the derivation in Appendix F.2.

Proposition 4. *The equilibrium relationship between changes in a firm's market share* $\left(\sigma_i^m\right)$ *and its own-elasticity of demand* $\left(\varepsilon_i^m\right)$ *is:*

$$d\log\varepsilon_i^m = -\left(\frac{\xi_i^m}{\varepsilon_i^m}\right)\left(\frac{\varepsilon_i^m}{1+\varepsilon_i^m}\right)d\log\sigma_i^m, \tag{27}$$

where ε_i^m *is the own-elasticity of demand of firm* i *in market* m*, defined as in* (13) *and* ξ_i^m *is its super-elasticity of demand, defined as in* (14).

We implement this relationship by regressing the (log) change of the own-elasticities we recovered above on the (log) changes of market shares we observe in the data.

$$\Delta\log\varepsilon_i^{st} = \hat{\beta}\cdot\Delta\log\sigma_i^{st}, \tag{28}$$

and estimate the aggregator by indirect inference, minimizing the distance between $\hat{\beta}$ and the coefficient implied by Proposition 4:

$$\min\left|\hat{\beta} - \beta\left(\Upsilon\right)\right|. \tag{29}$$

We show in Appendix F.2 that minimizing this objective is equivalent to minimizing the sales-share-weighted difference of changes in elasticities because

$$\left|\hat{\beta} - \beta\left(\Upsilon\right)\right| \propto \left|\sum \left(\Delta \log \varepsilon_i^{mt} - \left(\frac{\xi_i^{mt}}{\varepsilon_i^{mt}}\right)\left(\frac{\varepsilon_i^{mt}}{1+\varepsilon_i^{mt}}\right)\Delta \log \sigma_i^{mt}\right) \cdot \Delta \log \sigma_i^{mt}\right|. \tag{30}$$

That is, the estimation places more weight on the elasticity changes of firms with the largest changes in market shares. Changes in elasticity are exactly what is needed to identify super-elasticity ξ, and the changes of firms whose market share changed the most are the most informative.

This process takes advantage of the panel dimension of the data. The introduction of firm-specific demand shifters makes it unfeasible to estimate the shape of the aggregator solely from cross-sectional data, as in Baqaee, Farhi, and Sangani (2023) or Edmond, Midrigan, and Xu (2023). Instead, changes in elasticities and market shares provide the relevant variation to estimate demand parameters.

Functional form for demand. We estimate Υ_i for the leading functional form of the literature following Klenow and Willis (2016).[13] This functional form is parameterized by a pair of parameters so that $\Upsilon_i(x) = \Upsilon\left(x; \nu_i^m, \theta_m\right)$. The first parameter, ν_i^m, directly affects the elasticity of demand, and we make it firm-specific; this is our elasticity of demand shifter. The second parameter, θ_m, controls the shape of the function and is constant within a market. When $\theta = 0$ there is no variation in elasticities, $\varepsilon_i^m = \nu_i^m$, recovering the constant-elasticity-of-demand aggregator with firm-specific elasticities.

The functional form determines how demand depends on the firm's (relative) output.

[13] Papers like Edmond, Midrigan, and Xu (2023), Boar and Midrigan (2024), and Hubmer and Restrepo (2021) use the functional form of Klenow and Willis (2016). We also consider the functional form adopted by Dotsey and King (2005). We show in appendix E that Barde (2008), Levin, López-Salido, Nelson, and Yun (2008), Darracq Pariès and Loublier (2010), and Kurozumi and Zandweghe (2020) use functions that are equivalent to those in Dotsey and King (2005).

The elasticity and the super-elasticity of demand for firm i in market m become

$$\varepsilon_i^m = \nu_i^m \cdot \left(\frac{y_i^m}{Y_m}\right)^{-\frac{\theta_m}{\nu_i^m}} \quad \text{and} \quad \xi_i^m = \theta_m \cdot \left(\frac{y_i^m}{Y_m}\right)^{-\frac{\theta_m}{\nu_i^m}}. \tag{31}$$

The estimation takes advantage of the fact that the ratio of the super-elasticity to the elasticity in Proposition 4 and equation (30) is constant for a given firm,

$$\frac{\xi_i^m}{\varepsilon_i^m} = \frac{\theta_m}{\nu_i^m}. \tag{32}$$

Given a value for θ_m, we recover the value of the idiosyncratic shifters, $\left\{\nu_i^m\right\}$, by matching the observed market shares of firms. Equation (31) implies that

$$\frac{y_i^m}{Y_m} = \left(\frac{\varepsilon_i^m}{\nu_i^m}\right)^{-\frac{\nu_i^m}{\theta_m}}, \tag{33}$$

which allows us to evaluate σ_i^m directly as in equation (9) from Lemma 1. Matching the market shares amounts to solving another system of non-linear equations.[14]

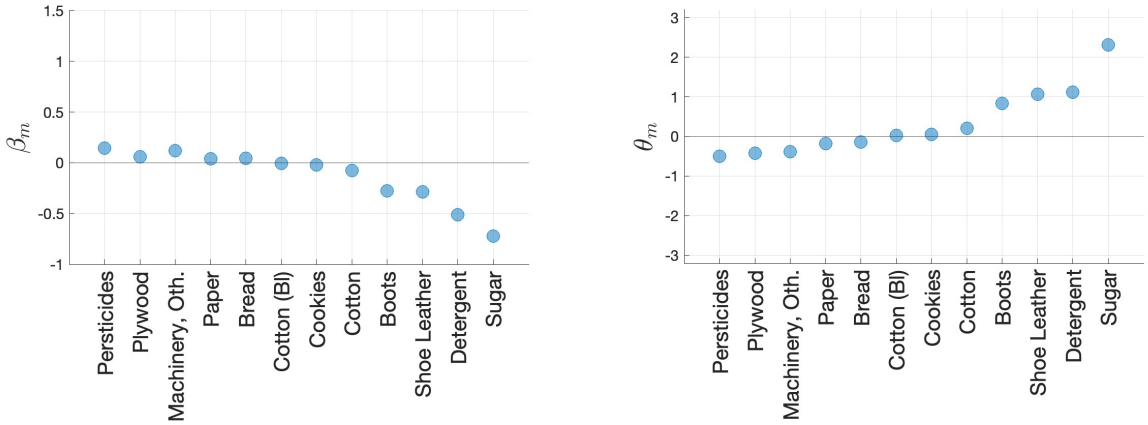
Demand estimates. Figure 5 presents the estimates of market-specific parameters $\left\{\hat{\beta}_m\right\}$ in Panel A and $\{\theta_m\}$ in Panel B. The model matches the estimated $\left\{\hat{\beta}_m\right\}$ precisely as we show in Figure B.5 in the Appendix. Some markets have firms with relatively stable elasticities, reflected in low values of both parameters that make the aggregator close to the standard CES with product-specific elasticities $\left(\varepsilon_i^m = \nu_i^m\right)$. However, the elasticities in these markets tend to increase when firms increase in size, resulting in negative estimates for θ_m, corresponding to varieties being gross-complements from

[14] The same process applies for alternative functional forms for Υ as that in Dotsey and King (2005). In this case, the relative output is expressed in terms of parameters as $y_i^m/Y_m = (1+\theta_m)\nu_i^m/\theta_m\left(\nu_i^m - \varepsilon_i^m\right)$. Then, the values of $\left\{\nu_i^m\right\}$ are recovered by matching the firms' market shares, $\left\{\sigma_i^m\right\}$, for a given θ_m. The ratio of super-elasticity to elasticity is $\xi_i^m/\varepsilon_i^m = -\frac{\theta_m}{1+\theta_m}/\left(\frac{y_i^m}{Y_m} - \frac{\theta_m}{1+\theta_m}\right)$.

FIGURE 5. Demand Parameters – Selected Markets (2005–2008)

(A) $\hat{\beta}_m$ by Market

(B) θ_m by Market



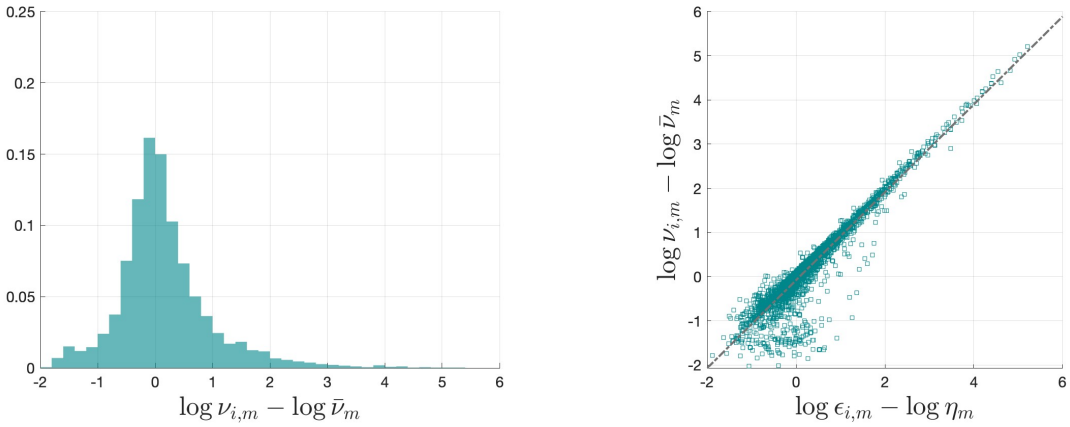
Notes: The left panel shows the $\hat{\beta}$'s estimated from the data using equation (28) and the market-specific β implied by the Klenow and Willis aggregator in Appendix F. The right panel shows the market-specific θ estimated that generate the β that minimize the distance with respect to the observed $\hat{\beta}$.

the point of view of the aggregator Υ.

Figure 6 presents the implied distribution of firm-specific elasticity shifters $\left\{\nu_i^m\right\}$ in Panel A and how it correlates with the distribution of own-elasticities of demand $\left\{\varepsilon_i^m\right\}$ in Panel B. These shifters are such that we exactly recover the observed market share of each firm. We find large variation in the distribution of shifters of comparable magnitude to that in markups (Figure 1A) and own-elasticities (Figure 4A). This reflects a limited role for differences in firm size to affect the elasticity of demand and, in turn, markups, as evidenced in the strong correlation between ν and ε in Figure 6B. We expand on this result in Section 5, where we show that idiosyncratic demand elasticity shifters are the key ingredient in accounting for the distribution of markups, particularly the high dispersion in markups among firms with low market shares.

FIGURE 6. Elasticity Shifter (ν) and Relationship with Elasticities

(A) ν Distribution (B) ν and Elasticities



Notes: Panel A shows the histogram of demand elasticity shifters relative to their market's average for the Indian data covering the period 2005–2007. Panel B shows the joint scatter plot of own-elasticities and demand elasticity shifters. Markets are defined at the 5-digit ASICC product level.

4.3. Recovering relative output, prices, and marginal costs

A byproduct of the estimation of the demand parameters is the set of relative outputs, $\{y_i^m/Y_m\}$, consistent with firms' choices in the model as in (33). Along with these relative outputs, we also recover relative prices and marginal costs (normalized by the sector's price level) using the measured market shares and markups as[15]

$$\frac{p_i^m}{P_m} = \sigma_i^m \frac{Y_m}{y_i^m} \quad \text{and} \quad \frac{\lambda_i^m}{P_m} = \frac{\sigma_i^m}{\mu_i^m} \frac{Y_m}{y_i^m}. \tag{34}$$

Finally, we express firms' marginal cost (relative to their market's marginal cost)

$$\frac{\lambda_i^m}{\lambda_m} = \frac{\frac{\sigma_i^m}{\mu_i^m}}{\frac{1}{N_m}\sum_j \frac{\sigma_j^m}{\mu_j^m}} \left(\frac{y_i^m}{Y_m}\right)^{-1}. \tag{35}$$

[15] Alternatively, we can obtain relative marginal costs as $\lambda_i^m/\lambda_j^m = {}^{p_i^m/P_m}/_{p_j^m/P_m} \cdot \mu_j^m/\mu_i^m$, and report marginal costs relative to, for instance, the lowest cost firm.

We find marginal costs are lower for high-markup firms, in line with the standard predictions of variable markup models (either coming from oligopolistic competition or variable elasticity of demand), as we show in Figure 7A. However, there is significant variation in marginal costs among firms with the same markups, particularly for firms with markups between 40 percent below and 65 percent above the market's average markup that make up the large majority of firms.[16]

By contrast, larger firms tend to have higher (and not lower!) marginal costs, as we show in Figure 7B. Although there is a lot of variation in marginal costs between small firms with near-zero market share, the correlation we find is present even among larger firms (some of which charge low markups despite their size). This fact opens up an additional source of misallocation coming from demand factors. This source interacts with the standard mechanism emphasized in the literature, where the most productive firms are simultaneously the largest and charge the highest markups, implying gains from reallocation economic activity towards them (Baqaee and Farhi 2020; Baqaee, Farhi, and Sangani 2023). We find that it is not the case that all large firms are productive, nor do they all charge high-markup. This pushes misallocation in opposite directions as there is a mass of highly-productive firms among the small firms of a market, but there are also unproductive firms whose size is kept in check because of their high markups.

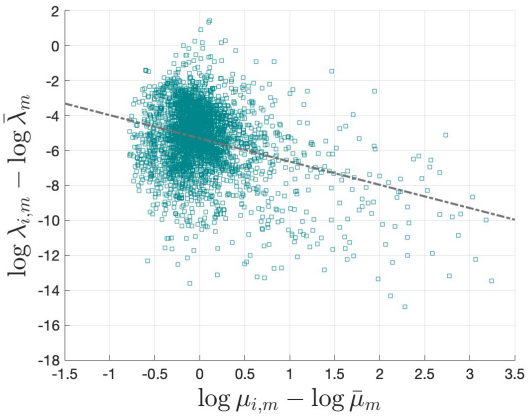
5. Market Structure, Firm Scale, and Demand Factors

Our results show that accounting for markup dispersion involves accounting for a large variation in firms' demand elasticities. These elasticities depend endogenously on the size of the firm (both in terms of its market share as in Atkeson and Burstein 2008 and

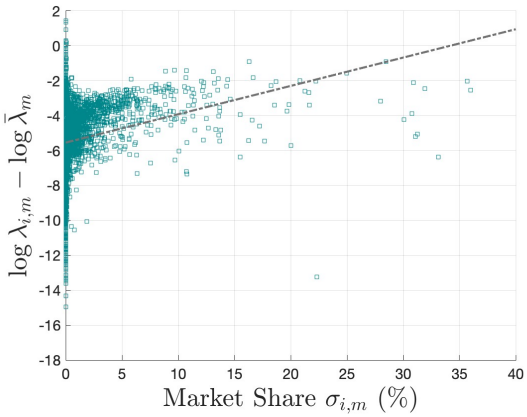
[16] Similarly, firms with lower marginal costs tend to produce more, as captured by higher relative output, but there is a lot of variation in the marginal costs given the size of the firm. Marginal costs do correlate strongly with relative prices despite the discrepancies between marginal costs and markups. We show these additional results in Figure B.6 in the Appendix. Figure B.7 compares these results to those coming from the production estimation described in Section 2.

FIGURE 7. Marginal Costs and Relationship with Markups and Market Shares

(A) λ and Markups



(B) λ and Market-Share



Notes: Panel A shows the scatter plot of firm log-marginal costs relative to their market's marginal cost and log-markups relative to their market's average. Panel B shows the scatter plot of firm log-marginal costs relative to their market share. Marginal costs are computed as in (35).

its scale of production as in Kimball 1995), the elasticities and sizes of competitors, and additional idiosyncratic factors that capture differences in elasticities between firms of the same size, as shown in Proposition 1. The role of market shares and market concentration in shaping elasticities reflects strategic interactions between granular firms in oligopolistic markets, while the scale of production of the firm matters when the firm's residual demand curve has variable elasticity. Our characterization of elasticities (and markups) merges these forces shaping firms' markups and extends them by introducing idiosyncratic factors that explain deviations in elasticities among firms with the same size (either in terms of their market share or scale of production).

To understand the role of these mechanisms, we now solve our model under three alternative specifications that differ in the form of competition between firms and in the properties of demand. Specifically, we consider: *(i)* a model of oligopolistic competition with a common constant-elasticity demand; *(ii)* a model of monopolistic competition with a common variable-elasticity demand; and *(iii)* a model of oligopolistic competition

with a common variable-elasticity demand. All three models differ from our full model in the absence of idiosyncratic demand elasticity shifters, setting $\nu_i^m = \nu_m$ in each.

Oligopolistic competition with constant elasticity of demand. Under these conditions, markups $\left\{\tilde{\mu}_i^m\right\}$ depend only on market concentration through market shares as in Atkeson and Burstein (2008), so that

$$\frac{1}{\tilde{\mu}_i^m} = \frac{\gamma - 1}{\gamma} + \left(\frac{1}{\gamma} - \frac{1}{\tilde{\varepsilon}_m}\right)\left(1 - \sigma_i^m\right), \tag{36}$$

where $\tilde{\varepsilon}_m = \nu_m$ is the (common) elasticity of demand in market m. We set $\tilde{\varepsilon}_m$ so that average markups in each market match the average measured markup. In this way, we make markups and elasticities comparable across models. We do this by choosing ν_m to solve $\sum_i \left(\mu_i^m\right)^{-1} \sigma_i^m = \sum_i \left(\tilde{\mu}_i^m\right)^{-1} \sigma_i^m$.

Monopolistic competition with variable elasticity of demand. Under these conditions, markups $\left\{\tilde{\mu}_i^m\right\}$ depend only on the firms' own-elasticity that varies depending on the firm's scale of production as in Kimball (1995),

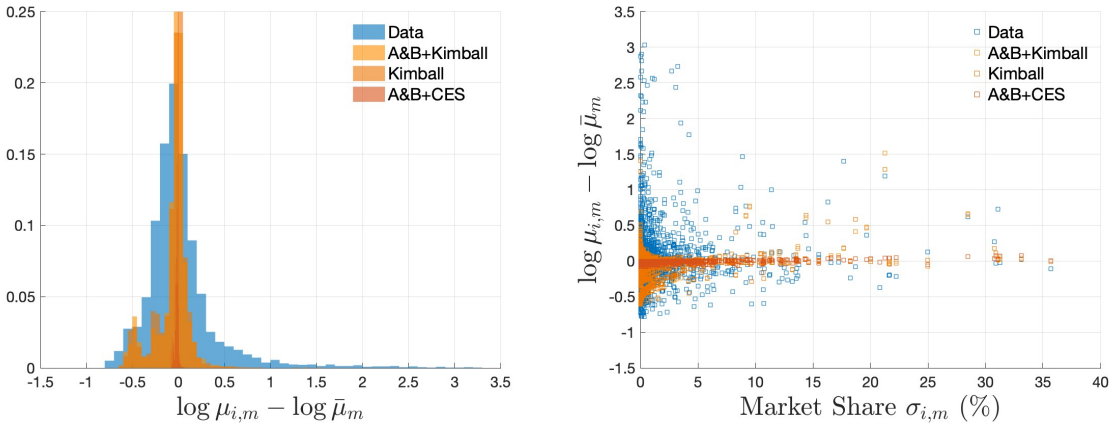
$$\frac{1}{\tilde{\mu}_i^m} = 1 - \frac{1}{\varepsilon_i^m}, \qquad \text{where } \varepsilon_i^m = \nu_m \left(\frac{y_i^m}{Y_m}\right)^{\frac{-\theta_m}{\nu_m}} \tag{37}$$

as in Klenow and Willis (2016). We set the demand parameters θ_m and ν_m and the relative outputs in each market $\left\{\frac{y_i}{Y}\right\}$ by minimizing the distance between the market shares implied by this demand following Lemma 1,

$$\tilde{\sigma}_i^m = \frac{\frac{y_i}{Y} \exp\left[\frac{1}{\theta_m}\left(1 - \left(\frac{y_i}{Y}\right)^{\frac{\theta_m}{\nu_m}}\right)\right]}{\sum_j \frac{y_j}{Y} \exp\left[\frac{1}{\theta_m}\left(1 - \left(\frac{y_j}{Y}\right)^{\frac{\theta_m}{\nu_m}}\right)\right]}. \tag{38}$$

FIGURE 8. Markups and Market Share Distribution Across Models (2005–2008)

(A) Markups Distribution (B) Markups and Market-Share



Notes: The left panel shows the histogram of markups relative to their market's average for the Indian data (blue bars) and the markups implied by the Atkeson and Burstein (2008) Bertrand markups taking as given the market shares of firms and adjusting the (common) elasticity of demand of each industry to match the market's average markup. The right panel shows shows the histogram the implied distribution of own-elasticities, ε_i^m, required to match the distribution of markups under Cournot competition and Cobb-Douglas demand for market output, $\gamma = 1.3$. The average elasticity in each market corresponds to the harmonic sales-weighted average of the individual elasticities.

Oligopolistic competition with variable elasticity of demand. Under these conditions, markups $\left\{\tilde{\mu}_i^m\right\}$ depend on both market concentration and on own elasticities as noted in Amiti, Itskhoki, and Konings (2019, Appendix D). In the case of Bertrand competition this results in

$$\tilde{\mu}_i^m = 1 - \left(\gamma\sigma_i^m + \varepsilon_i^m\left(1 - \frac{\varepsilon_i^m\sigma_i^m}{E_\sigma\left[\varepsilon_i^m\right]}\right)\right)^{-1}. \tag{39}$$

We set the own-elasticities and obtain the required parameters as in the previous case.

Markups across models. Figure 8 reproduces Figure 1 overlaying the distribution of markups under the alternative models without demand elasticity shifters. All models match, by construction, the market share of each individual firm and differ only in their

implications for markups. Panel A makes clear that variation in market shares alone is not enough to capture the observed dispersion in firm markups. Market concentration alone is not enough to drive significant variation in markups, even when it generates differences in markup levels.

Allowing for variable-elasticity demand generates four times as much variation in markups—with or without strategic interactions between firms. This follows from our results in Figure 4, establishing that dispersion in markups is tied to significant dispersion in demand elasticities. Nevertheless, without idiosyncratic elasticity shifters, the alternative models generate only one-fourth of the variance in markups in the data.

Panel B contrasts the joint distributions of markups and market shares across models. Once again, variation in the elasticity of demand is required to generate markup dispersion, but the constraint imposed by the common demand curve proves binding when accounting for the large dispersion in markups among small firms. In particular, the alternative models cannot explain small firms that charge high markups as well as the presence of large firms with below-average markups.

The patterns we describe above come together in the ability of alternative models to account for markup dispersion between firms of similar sizes. As we show in Table 2, the models without idiosyncratic demand elasticity shifters account for only a fraction of the variation in markups. Moreover, they counterfactually predict that at least two-thirds of the variation comes from differences between firms of different sizes, while these account for only 23 percent of the variance of markups in the data. This is further reflected in the implied correlation between markups and market shares (a common measure of market power), which is only 0.06 in the data but is no less than 0.20 in the alternative models.[17] We reproduce these exercises in our Colombian and US samples

[17]The correlation coefficient we report is computed for the pooled cross-section of all market-years; as opposed to the variance decomposition that treats each market-year separately to then average. Because of this, and the size of the bins, the correlation is not perfect even as the only source of markup variation in the alternative models is firm size.

TABLE 2. Markup Variance Decomposition India (2001-2008)

	Variance	Share within	Share between	Corr($\log \mu_i$, σ_i)
Data	0.196	0.770	0.230	0.061
Full Model	0.195	0.772	0.228	0.061
A&B + Kimball	0.047	0.388	0.612	0.211
Kimball	0.048	0.338	0.662	0.217
Atkeson & Burstein	0.001	0.175	0.825	0.337

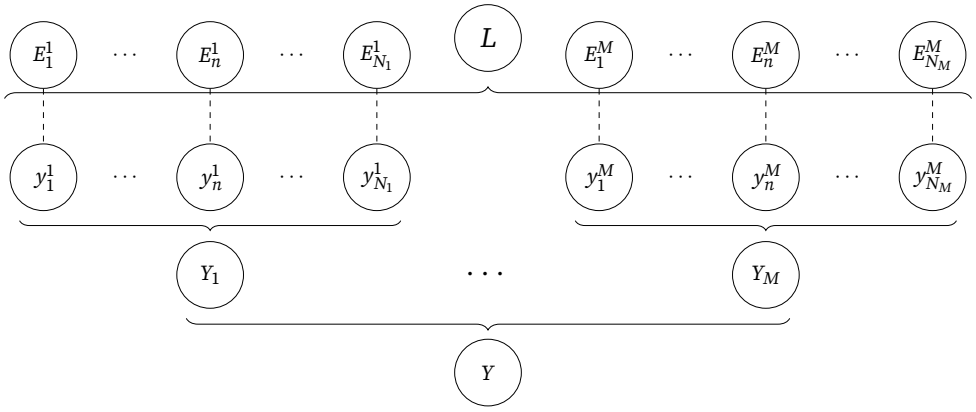
Notes: The table presents statistics for measured markups and market shares along with statistic from four models: our full model described in Section 3, a model of oligopolistic competition with constant elasticity of demand, a model of monopolistic competition with a common demand with variable elasticity, and a model of oligopolistic competition with variable elasticity of demand. The variance is computed from the log ratio $\log(\mu_i^m/\mu_m)$. Numbers are averages across markets. To decompose the variance of markups, we divide firms in equally sized bins of 5 percentage point market share. The variance and its decomposition correspond to the average variance and shares across market-year pairs. The correlation between markups and market shares is the pooled correlation of the full sample.

in Figure B.10 and Table B.2, verifying our results.

5.1. Markup dispersion and misallocation

Our results provide evidence that factors that affect firms' elasticity of demand play a central role in accounting for markups, in line with Blum, Claro, Horstmann, and Rivers (2024) and Eslava, Haltiwanger, and Urdaneta (2023). This has direct implications for the scope of misallocation stemming from markups, as large firms are not necessarily charging high-markups and are often less productive than smaller firms—even in narrowly-defined product markets. In this way, our results highlight the importance of mechanisms that endogenize the distribution of demand factors such as the presence of lock-in innovations that operate by capturing demand instead of raising productivity as in Casal (2026), customer acquisition and customer capital as in Afrouzi, Drenik, and Kim (2025); Hubmer and Nord (2025); Haddara (2026), among others.

FIGURE 9. Horizontal Economy



Notes: The figure describes the horizontal economy in Section 3 where final output (Y) is produced using products from M product-markets $\{Y_m\}$, each an aggregate of the N_m product varieties in the market $\{y_n^m\}$. Each product variety is produced using a stand-in inputs L and E_n^m. Input L stands for the physical inputs employed in production. Inputs E_n^m stand for the entrepreneurial input present in decreasing-returns-to-scale production functions.

We illustrate the role of demand factors in markup dispersion by computing the *distance to the frontier* in the horizontal economy of Section 3, described in Figure 9. That is, the (log-)gap between the output without markups, $Y^\star$, and the output in the economy, Y. Our objective is to show how different mechanisms that affect the dispersion of markups shape the efficiency cost of markups in a distorted economy. We provide the formulas for the distance to the frontier in Appendix D. They apply the results of Baqaee and Farhi (2020) to our economy with non-CES aggregation.

There are two margins of misallocation in our economy. Namely, there is misallocation in the output of the producers of product-varieties inside each market and misallocation in the use of inputs by those producers. Misallocation in output markets for product varieties depends not only on markup dispersion but also on the elasticities of substitution $\left\{\phi_{ij}^m\right\}$ between varieties implied by the market aggregator's technology defined in (6). Crucially, these elasticities are internally consistent with the markups and market shares of firms in the sample. The input-output linkages in the economy then connect this misallocation to market aggregators and the aggregator of

TABLE 3. Distance to the Frontier in India (2001-2008)

	$\log \frac{Y^\star}{Y}$	Output Misallocation	Input Misallocation
Full Model	0.245	0.227	0.018
A&B + Kimball	0.054	0.043	0.012
Kimball	0.055	0.046	0.009
Atkeson & Burstein	0.025	0.019	0.006

Notes: The table presents the distance to the frontier as described in Appendix D. Numbers correspond to the average statistic across sample years.

final output that operate competitively and charge no markups of their own.

Misallocation in input markets reflects the production choices of producers of product varieties. In particular, the scalability of their production technologies, captured by their returns-to-scale.[18] To highlight this dimension—often absent from analysis of misallocation—and to keep the exercise simple, we assume that there are only two types of inputs in the economy, a fixed managerial input used in product markets with decreasing returns to scale $\{E_n^m\}$, and a flexible stand-in input that we term L. All producers use of L, but each has its own type of managerial input.[19]

We calculate the misallocation cost of markups for our baseline model and the three alternative models described in Section 5 and Table 2. In each case, we report the total cost of misallocation and the terms corresponding to output and input misallocation. Table 3 presents the results. Our baseline model—that accounts for the joint distribution of markups and market shares—implies an output gain of 24.5 log-points (or 28 percent) from eliminating markups. This is over 5 times more than what is implied when not including idiosyncratic variation in elasticity shifters (as in

[18] Hubmer et al. (2025) show that there is a large variation in returns to scale across firms using Canadian and US administrative data and link them to differences in returns and profitability differences.

[19] The introduction of the managerial inputs makes production technologies be constant-returns-to-scale and allows us to incorporate sources of profit coming from scalability instead of markups.

the A&B plus Kimball and Kimball models in the second and third rows) and over 10 times more than what is implied with no variation in elasticities of substitution across products (as in the Atkerson & Burstein model in the fourth row). In all cases, most of the misallocation comes from output markets, reflecting the fact that most production technologies in our sample are close to being constant-returns-to-scale (see Appendix A.3). This leaves little room for reallocation between physical and managerial inputs.

Ultimately, markup dispersion leads to substantial losses from misallocation. This dispersion in markups is accounted for primarily by dispersion in firms' own-elasticities of demand $\{\varepsilon_i^m\}$, which we interpret as pointing towards demand factors that can shift these elasticities.[20] Crucially, all three models match, by construction, the same market concentration and the same average markup in each market. This highlights the critical role of differences in the dispersion of markups between firms of similar size, on which we have focused, in estimating the losses from markup dispersion.

6. Conclusions

We document that most markup dispersion arises among firms of similar size, a fact that cannot be explained by productivity differences or market concentration alone. To account for this pattern, we develop an analytically tractable model of oligopolistic competition with variable markups and non-CES demand that merges two of the main models of variable markups and extends them by introducing firm-specific demand elasticity shifters. Quantitatively, these shifters are essential for matching the joint distribution of markups and market shares observed in Indian manufacturing data and complementary evidence from the United States and Colombia. The model reconciles the presence of small, high-markup firms and large, low-markup firms, and implies a weak link between firm size, productivity, and market power, highlighting demand

[20] Our quantitative results for the product markets we study align with the findings in Baqaee and Farhi (2020), Baqaee, Farhi, and Sangani (2023) and Edmond, Midrigan, and Xu (2023), among others.

heterogeneity as a key driver of markup dispersion. Not accounting for this heterogeneity underestimates the cost of markups by a factor of at last 5.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6):2411–2451. URL https://doi.org/10.3982/ECTA13408.

Afrouzi, Hassan, Andres Drenik, and Ryan Kim. 2025. "Concentration, Market Power, and Misallocation: The Role of Endogenous Customer Acquisition." Working Paper 31415, National Bureau of Economic Research. URL https://doi.org/10.3386/w31415.

Albrecht, Brian C., Tom Phelan, and Nick Pretnar. 2023. "Time Use and the Efficiency of Heterogeneous Markups." Working Paper 23–28, Federal Reserve Bank of Cleveland. URL https://doi.org/10.26509/frbc-wp-202328.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2019. "International Shocks, Variable Markups, and Domestic Prices." *The Review of Economic Studies* 86 (6):2356–2402. URL https://doi.org/10.1093/restud/rdz005.

Atkeson, Andrew and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review* 98 (5):1998–2031. URL https://doi.org/10.1257/aer.98.5.1998.

Baqaee, David Rezza and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *The Quarterly Journal of Economics* 135 (1):105–163. URL https://doi.org/10.1093/qje/qjz030.

Baqaee, David Rezza, Emmanuel Farhi, and Kunal Sangani. 2023. "The Darwinian Returns to Scale." *The Review of Economic Studies* :1–33URL https://doi.org/10.1093/restud/rdad061.

Barde, Sylvain. 2008. "A Generalized Variable Elasticity of Substitution Model of New Economic Geography." Sciences Po publications 2008-33, Sciences Po. URL https://ideas.repec.org/p/spo/wpmain/infohdl2441-7348.html.

Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Suedekum. 2020. "Quantifying the Gap Between Equilibrium and Optimum under Monopolistic Competition." *The Quarterly Journal of Economics* 135 (4):2299–2360. URL https://doi.org/10.1093/qje/qjaa017.

Blum, Bernardo S., Sebastian Claro, Ignatius Horstmann, and David A. Rivers. 2024. "The ABCs of Firm Heterogeneity When Firms Sort into Markets: The Case of Exporters." *Journal of Political Economy* 132 (4):1162–1208. URL https://doi.org/10.1086/727289.

Boar, Corina and Virgiliu Midrigan. 2024. "Markups and Inequality." *The Review of Economic Studies* 92 (5):2828–2860. URL https://doi.org/10.1093/restud/rdae103.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data." *Journal of Monetary Economics* 121:1–14. URL https://doi.org/10.1016/j.jmoneco.2021.05.004.

Burdett, Kenneth and Kenneth L. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51 (4):955–970. URL https://www.jstor.org/stable/1912045.

Burstein, Ariel T, Vasco M Carvalho, and Basile Grassi. 2025. "Bottom-Up Markup Fluctuations." *The Quarterly Journal of Economics* 140 (4):2619–2684. URL https://doi.org/10.1093/qje/qjaf029.

Butters, Gerard R. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *The Review of Economic Studies* 44 (3):465–491. URL https://doi.org/10.2307/2296902.

Casal, Lucia. 2026. "Lock-In and Productive Innovations: Implications for Firm-to-Firm Innovation Pass-Through." URL https://www.luciacasal.com/research/Casal_JMP.pdf. Unpublished.

Chernoff, Alex, Allen Head, and Beverly Lapham. 2024. "Consumer Search, Productivity Heterogeneity, Prices, Markups, and Pass-through: Theory and Estimation." Staff Working Paper 2024–50, Bank of Canada. URL https://doi.org/10.34989/swp-2024-50.

Darracq Pariès, Matthieu and Alexis Loublier. 2010. "Epstein-Zin preferences and their use in macro-finance models: implications for optimal monetary policy." Working Paper Series 1209, European Central Bank. URL https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1209.pdf.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *The Quarterly Journal of Economics* 135 (2):561–644. URL https://doi.org/10.1093/qje/qjz041.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik. 2016. "Prices, Markups, and Trade Reform." *Econometrica* 84 (2):445–510. URL https://doi.org/10.3982/ECTA11042.

De Loecker, Jan and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review* 102 (6):2437–71. URL https://doi.org/10.1257/aer.102.6.2437.

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti. 2026. "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates from Financial Data." *Econometrica* 94 (1):137–168. URL https://doi.org/10.3982/ECTA22733.

Dhingra, Swati and John Morrow. 2019. "Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity." *Journal of Political Economy* 127 (1):196–232. URL https://doi.org/10.1086/700732.

Dixit, Avinash K. and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *The American Economic Review* 67 (3):297–308. URL http://www.jstor.org/stable/1831401.

Dotsey, Michael and Robert G. King. 2005. "Implications of state-dependent pricing for dynamic macroeconomic models." *Journal of Monetary Economics* 52 (1):213–242. URL https://doi.org/10.1016/j.jmoneco.2004.10.004.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015. "Competition, Markups, and the Gains from International Trade." *American Economic Review* 105 (10):3183–3221. URL https://doi.org/10.1257/aer.20120549.

———. 2023. "How Costly Are Markups?" *Journal of Political Economy* 131 (7):1619–1675. URL https://doi.org/10.1086/722986.

Eslava, Marcela, John Haltiwanger, and Nicolas Urdaneta. 2023. "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand, and Wedges." *The Review of Economic Studies* 91 (1):259–300. URL https://doi.org/10.1093/restud/rdad029.

Foster, Lucia, Cheryl Grim, and John Haltiwanger. 2016. "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics* 34 (S1):S293–S331. URL https://doi.org/10.1086/682397.

Gale, David and Hukukane Nikaido. 1965. "The Jacobian matrix and global univalence of mappings." *Mathematische Annalen* 159 (2):81–93. URL https://doi.org/10.1007/BF01360282.

Gandhi, Amit, Salvador Navarro, and David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." *Journal of Political Economy* 128 (8):2973–3016. URL https://doi.org/10.1086/707736.

Grassi, Basile. 2017. "IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important." URL https://sites.google.com/site/grassibasile/home/research/job-market-paper. Unpublished.

Haddara, Mahmood. 2026. "Customer Accumulation, Firm Size, and Aggregate Productivity." URL https://drive.google.com/file/d/12vQkzlrKA3bhGShLjw8d-D0OVwCvhs3K/view. Unpublished.

Hannah, Leslie and John A. Kay. 1977. *Concentration in Modern Industry: Theory, measurement and the U.K. experience*. London, UK: Palgrave Macmillan. URL https://doi.org/10.1007/978-1-349-02773-6.

Hasenzagl, Thomas and Luis Pérez. 2023. "The Micro-Aggregated Profit Share." Tech. rep., Arxiv, General Economics. URL https://doi.org/10.48550/arXiv.2309.12945.

Herreño, Juan, Noémie Pinardon-Touati, and Malte Thie. 2025. "How Steep is the Phillips Curve in Developing Economies? A Sufficient Statistics Approach and Estimates for India." URL https://juanherreno.github.io/hpt_india.pdf. Unpublished.

Hubmer, Joachim, Mons Chan, Serdar Ozkan, Sergio Salgado, and Guangbin Hong. 2025. "Scalable versus Productive Technologies." Tech. rep., The Wharton School Research Paper. URL http://dx.doi.org/10.2139/ssrn.4915780.

Hubmer, Joachim and Lukas Nord. 2025. "Investment in Demand and Dynamic Competition for Customers." Tech. rep., University of Pennsylvania. URL https://dx.doi.org/10.2139/ssrn.5491506.

Hubmer, Joachim and Pascual Restrepo. 2021. "Not a Typical Firm: The Joint Dynamics of Firms, Labor Shares, and Capital–Labor Substitution." Working Paper 28579, National Bureau of Economic Research. URL https://doi.org/10.3386/w28579.

Kimball, Miles S. 1995. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit and Banking* 27 (4):1241–1277. URL https://doi.org/10.2307/2078048.

Kirov, Ivan, Paolo Mengano, and James Traina. 2026. "Measuring Markups with Revenue Data." URL https://paolomengano.com/Markups_Revenues.pdf. Unpublished.

Klenow, Peter J. and Jonathan L. Willis. 2016. "Real Rigidities and Nominal Price Changes." *Economica* 83 (331):443–472. URL https://doi.org/10.1111/ecca.12191.

Kurozumi, Takushi and Willem Van Zandweghe. 2020. "Output-Inflation Trade-offs and the Optimal Inflation Rate." Working Papers 202020, Federal Reserve Bank of Cleveland. URL https://doi.org/10.26509/frbc-wp-202020.

Levin, Andrew T, J David López-Salido, Edward Nelson, and Tack Yun. 2008. "Macroeconometric equivalence, microeconomic dissonance, and the design of monetary policy." *Journal of Monetary Economics* 55:S48–S62. URL https://doi.org/10.1016/j.jmoneco.2008.07.013.

Levinsohn, James and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies* 70 (2):317–341. URL https://doi.org/10.1111/1467-937X.00246.

Lindé, Jesper and Mathias Trabandt. 2018. "Should we use linearized models to calculate fiscal

multipliers?" *Journal of Applied Econometrics* 33 (7):937–965. URL https://doi.org/10.1002/jae.2641.

Matsuyama, Kiminori. 2023. "Non-CES Aggregators: A Guided Tour." *Annual Review of Economics* 15 (Volume 15, 2023):235–265. URL https://doi.org/10.1146/annurev-economics-082322-013910.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6):1695–1725. URL https://doi.org/10.1111/1468-0262.00467.

Menzio, Guido. 2024. "Markups: A Search-Theoretic Perspective." Working Paper 32888, National Bureau of Economic Research. URL https://doi.org/10.3386/w32888.

Olley, G. Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6):1263–1297. URL https://doi.org/10.2307/2171831.

Raval, Devesh. 2023. "Testing the Production Approach to Markup Estimation." *The Review of Economic Studies* 90 (5):2592–2611. URL https://doi.org/10.1093/restud/rdad002.

Smith, Dominic A. and Sergio Ocampo. 2025. "The Evolution of US Retail Concentration." *American Economic Journal: Macroeconomics* 17 (1):71–101. URL https://doi.org/10.1257/mac.20220249.

Varian, Hal R. 1980. "A Model of Sales." *The American Economic Review* 70 (4):651–659. URL http://www.jstor.org/stable/1803562.

Wooldridge, Jeffrey M. 2009. "On estimating firm-level production functions using proxy variables to control for unobservables." *Economics Letters* 104 (3):112–114. URL https://www.sciencedirect.com/science/article/pii/S0165176509001487.

Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and Jacques-François Thisse. 2012. "Monopolistic Competition: Beyond the Constant Elasticity of Substitution." *Econometrica* 80 (6):2765–2784. URL https://doi.org/10.3982/ECTA9986.

Appendix A. Estimation of Production Functions and Markups

A.1. The Indian Annual Survey of Industries

TABLE A.1. Product Classification Codes

Sector			Product		
Code	**Name**	**Sales Share in Ind.**	**Code**	**Name**	**Sales Share in Sector**
131	Sugar & Molasses	0.43	13103	Refined Sugar	0.76
134	Bakery Products	0.24	13401	Biscuits & Cookies	0.83
			13402	Bread, Buns & Croissants	0.10
344	Pesticides & Insecticides	0.13	34422	Pesticides	0.59
			34423	Insecticides	0.23
363	Soaps & Detergents	0.61	36304	Detergent Powder	0.20
			36307	Toilet Soap	0.34
443	Leather Footwear	0.68	44302	Boots	0.50
			44306	Leather Shoe Uppers	0.25
511	Wood & Wood Products	0.29	51105	Timber & Wooden Planks	0.22
			51125	Plywood	0.48
551	Printing & Writing Paper	0.54	55102	Printing & Writing Paper	0.76
632	Cotton Yarn & Fiber	0.43	63216	Bleached Cotton Yarn	0.30
			63221	Unbleached Cotton Yarn	0.27
642	Man-Made Yarn & Fiber	0.74	64222	Polyester Yarn	0.21
732	Aluminum & Alloys	0.44	73208	Aluminum Sections, Plates, etc.	0.32
769	Misc. Non-Electrical Machinery	0.34	76904	Boiler Accessories	0.29
			76926	Other Machinery	0.34
772	Electrical Motors & Transformers	0.22	77232	Transformer	0.28
782	Audio/Video Apparatus	0.47	78234	Printed Circuit Boards (PCB)	0.06
			78256	Color TV Set	0.61
821	Motor Vehicles & Parts	0.53	82189	Other Motor Vehicles & Parts	0.08

Note: The sector shares in the left panel are calculated with respect to the total revenue in the 2-digit industry across all the sample (2001-2008). The product shares are calculated with respect to the total revenue in the 3-digit sector across all the sample.

A.2. Co-Production Methodology

We identify co-production patterns using multi-product firms to maximize the number of available single-product firms in 3-digit sectors. Specifically, we consider as part of a 3-digit sector, the single-product firms directly producing in it, and the single-product firms producing in 3-digit sectors identified as co-production sectors based

on multi-product firm data. Formally, we identify the 3-digit sectors where we find the highest number of single-product firms. Then, for each 3-digit sector k, we select the multi-product firms that produce at least one good in k. For each mult-iproduct firm m selected, we identify the set of 3-digit sectors k' different from k in which they produce. We then aggregate and calculate the number of multi-product firms that co-produce in k and k'. For each sector k, the top two k' sectors most frequently co-produced with k are retained. Finally, we combine single-product firms producing in k and in the top two k' sectors and estimate the production function for this group, assuming that k and the selected k' use similar technologies.

Example of Sector and Co-production Patterns. To illustrate our methodology, consider the following example from the **Food** industry. In this example, sector 113 (Fish) is one of the top 2 sectors in the *Animal Based Food Products* industry. Within this sector, the two products that generate the most revenue are *Shrimps* and *Frozen Fish*. To increase the number of single-product firms available for production function estimation, we consider co-production patterns. Based on multi-product firm data, we identify the most frequently co-produced sectors: 116 (Preparation used for animal feeding) and 112 (Meat). Single-product firms producing in these co-production sectors are pooled together with those in sector 113, assuming they use similar technologies. This subset of firms is used to estimate production function parameters of *Fish* sector. Our final unit of analysis to solve our model is the subset of 5-digit products.

Industry: 11. Animal Based Food Products	
Main Sector	**Main Products**
113. Fish	11314. Shrimps
	11316. Frozen Fish
Co-production Sectors	
116. Preparation used for animal feeding	
112. Meat	

TABLE A.2. Competition within 2-digit Industries and 3-digit sectors

2-Digit	3-Digit Sector	Included	Excluded Examples
13-Food	131-Sugar/Molasses	Refined Sugar	Lactose
	134-Bakery Products	Cookies, Bread, Buns	Cakes
	132-Macaroni/Noodles	—	Spaghetti
34-Fertilizer	344-Pesticides	Pesticides, Insecticides	Mosquito Repellent
	349-Misc. Agri Chem	—	Plant Hormones
36-Cleaning	363-Soaps/Deters.	Detergent, Toilet Soap	Shaving Cream
	365-Dental Materials	—	Dental Flox
44-Leather	443-Footwear	Boots	Shoes Uppers, Liners
	441-Leather Bags	—	Vanity Bags
51-Wood	511-Wood Products	Timber, Planks, Plywood	Wood Wool
	512-Wooden Furniture	—	Benches
55-Paper	551-Print/Writing	Print/Writing Paper	Art Paper
	553-Coated Paper	—	Aluminum Paper
63-Cotton	632-Yarn/Fiber	Bleached/Unbleach. Yarn	Linters
	633-Cotton Fabrics	—	Towels
64-Synthetic	642-Man-made Yarn	Polyester Yarn	Fibre Acrylic
	644-Made-up Articles	—	Underwear
73-Metals	732-Aluminum	Aluminum Sections	Aluminum Bottles
	737-Other Metals	—	Mercury
76-Non-Elec.	769-Machinery Misc.	Boiler Accessories	Incubators
	764-Construction	—	Cement Machinery
77-Elec. Mach.	772-Motors/Transf.	Transformer	Spin Motors, Air Heaters
	771-Elec. Machinery	—	Miscellaneous parts
78-Electronics	782-Audio/Video	Circuit Boards, TV-Sets	Amplifiers
	783-Computer	—	Laptops
82-Vehicles	821-Motor Vehicles	General Motor Vehicles	Racing Cars
	823-Cycles/Scooters	—	Scooters

Note: "—" denotes example sectors that were not included in the estimation.

A.3. Production Function Estimation

We estimate the quantity-based production function

$$q_{ijt} = f(m_{ijt}, l_{ijt}, k_{ijt}) + \omega_{ijt} + \zeta_{ijt}, \tag{A.1}$$

where q_{ijt} denotes product j *log-quantities* produced by establishment i at year t; m_{ijt}, l_{ijt}, and k_{ijt} denote product-specific *physical* intermediate inputs, labor and capital, respectively; ω_{ijt} denotes a productivity term; and ζ_{ijt} is measurement error.

Estimating a quantity-based production function uses separate price and quantity information, removing output-price bias, present in revenue-based production functions. However, simultaneity and selection biases remain because productivity is unobserved and firms' input choices depend on productivity. Two further issues arise. First, input allocations across products within multi-product firms are unobserved.

This matters, for instance, for firms using differentiated inputs for differentiated products so that input prices vary depending on the product they are used for. Although we have data on the number of workers and some input quantities, physical units of capital are not observed. We therefore use input expenditures for all inputs following De Loecker et al. (2016).

In general, without product-level input data for a variable input x, the cost-share in equation (2) has to be computed as

$$\mu_{ijt} = \epsilon_{x,ijt} \times \frac{P_{ijt}Q_{ijt}}{\rho_{x,ijt}R_{x,it}}, \tag{A.2}$$

where $\rho_{x,ijt}$ is the share of input x expenditures attributable to product j for establishment i in period t, and $R_{x,it}$ are total expenditures in input x. Not observing product-level inputs also means that there are no product-level input prices available, which makes the estimation of the elasticity $\epsilon_{x,ijt}$ challenging.

The relationship between product- and firm-level input demand is then

$$x_{ijt} = \rho_{x,ijt} + r_{x,it} - w_{x,ijt}, \tag{A.3}$$

where $\hat{w}_{x,ijt}$ denotes the deviation of the unobserved (log) firm–product-specific variable input price from the (log) industry-wide variable input price index used to deflate input expenditures.

Consequently, equation A.1 can be written as:

$$\begin{aligned} q_{ijt} = {} & f\left(r_{l,it}, r_{m,it}, r_{k,it}\right) + \omega_{ijt} + \zeta_{ijt} \\ & + A\left(\rho_{ijt}, r_{l,it}, r_{m,it}, r_{k,it}\right) + B\left(w_{ijt}, \rho_{ijt}, r_{l,it}, r_{m,it}, r_{k,it}\right), \end{aligned} \tag{A.4}$$

where we use expenditures $r_{x,it}$ to redefine the production function in terms of observables, and functions A and B are unobservable due to the absence of information about ρ_{ijt} and w_{ijt}, the vectors of (log) input expenditure shares and price deviations by product. The bias emerges from the correlation between the unobservable productivity ω and A and B via the deflated input expenditures.

We follow De Loecker et al. (2016) to address these issues. First, we control for the simultaneity bias by employing a control function as in Olley and Pakes (1996) and Levinsohn and Petrin (2003). We assume that the intermediate input demand is a monotone function in productivity, while labor and capital are chosen before the firm observes the productivity realization ω. Under this assumption, we can invert a control function for productivity that depends on input demands.

To address the unobserved input allocation issue, we limit estimation to single-product firms for which input allocation is known, that is, $\rho_{x,ijt} = 1$. For single-product firms, equation A.4 reduces to:

$$q_{ijt} = f\left(r_{l,it}, r_{m,it}, r_{k,it}\right) + B\left(w_{ijt}, r_{l,it}, r_{m,it}, r_{k,it}\right) + \omega_{ijt} + \zeta_{ijt}. \tag{A.5}$$

We assume multi-product firms use the same technologies in the manufacturing of each product as as single-product firms that make that product. This means production

technology is independent across product lines within a firm.

Finally, we have to correct for the term related to the firm-specific input prices $B(\cdot)$. To address this, we rely on a control function that exploits the theoretical relation between input prices, input quality, and output quality. We follow De Loecker et al. (2016) and argue that high-quality products produced using high-quality inputs command higher final prices. This theoretical connection establishes a positive relationship between unobserved product-specific input prices and observed product-specific output prices. In particular, we use a non-parametric control function using output prices and market shares to approximate for this relation. We complement this control function with state-level fixed effects.

For estimation, we measure the expenditure in labor services, $r_{l,it}$, using the wage bill and the expenditure in intermediate inputs, $r_{m,it}$, using the intermediate input total expenditure. Expenditure in capital services, $r_{k,it}$, is measured using the perpetual inventory method. All inputs are deflated using a 3-digit sector-year specific deflator.

We implement the estimation of A.4 following Wooldridge's 2009 method. We use first-order lags to instrument for intermediate input expenditure, output prices, and market shares. We assume that labor and capital are predetermined and are not correlated with productivity. We assume $f(\cdot)$ is Cobb-Douglas.

After estimating output elasticities at the 3-digit sector level using only single-product firms, we assign these elasticities to each product line in the sector for multi-product firms. We then assume multi-product firms allocate inputs in proportion to each product's revenue share within the firm as in Blum et al. (2024), since direct input allocations are unobserved.

For robustness, we also compute input allocations following De Loecker et al. (2016). Their approach assumes common productivity across products within the same establishment, which allows for the construction of a system of equations that determines input allocations. Using this method, we find a correlation of approximately 0.7 between revenue shares and the allocations generated by their model. We retain item-level revenue shares as our baseline because, although highly correlated, the De Loecker et al. (2016) routine often produces input allocations near the boundaries, which distort the resulting markup distributions leading to negative or implausibly large markups.

A.4. Revenue production functions: Cost-share approach

In additional exercises we use data from the Colombian manufacturing survey and the US Census of manufactures. These datasets do not report prices and quantities separately and thus force us to estimate revenue-based production functions. We do this following the cost-share approach This method relies on the cost-minimization (2) and imposes constant-returns-to-scale on variable inputs to recover the elasticity of output with respect to variable inputs from their cost-shares. We follow the exposition of Raval (2023) who builds on Foster, Grim, and Haltiwanger (2016). This is the approach followed, for instance, by Edmond, Midrigan, and Xu (2023).

TABLE A.3. Sector-Level Production Function Betas and Returns to Scale

Sector Code	**Sector Name**	β_M	β_K	β_L	**RTS**
131	Sugar & Molasses	0.85	0.01	0.11	0.97
134	Bakery Products	1.05	0.06	-0.12	0.99
344	Pesticides & Insecticides	0.63	0.20	0.01	0.84
363	Soaps & Detergents	0.66	0.17	0.30	1.13
443	Leather Footwear	0.83	-0.05	0.16	0.94
511	Wood & Wood Products	0.82	0.02	0.09	0.93
551	Printing & Writing Paper	0.90	0.07	-0.03	0.94
632	Cotton Yarn & Fiber	1.06	-0.06	-0.02	0.97
642	Man-Made Yarn & Fiber	0.59	0.01	0.05	0.65
732	Aluminum & Alloys	0.88	0.01	0.09	0.97
769	Misc. Non-Electrical Machinery	0.75	0.11	0.21	1.07
772	Electrical Motors & Transformers	0.26	0.16	0.45	0.87
782	Audio/Video Apparatus	0.82	-0.09	0.18	0.91
821	Motor Vehicles & Parts	0.69	-0.12	0.58	1.15

Note: Only the beta coefficients for material (β_M), capital (β_K), and labor (β_L) are shown. The "RTS" column provides an empirical estimate of returns to scale for each sector. The estimates correspond to the results described in section 2, for a Cobb-Douglas production function.

The key assumption is that firms operating constant-returns-to-scale technologies optimize over the level of all their inputs on average, so that the following expression holds in the cross-section of firms for each variable input x_h:

$$E\left[P_n X_{n,i}\right] \;=\; \epsilon_n E\left[\lambda_i Y_i\right] \;. \tag{A.6}$$

This is weaker than demanding that each firm's first order condition holds, and allows for firms to have adjustment costs, or time-to-build costs, as long as they average in the aggregate. This method also allows for non-neutral technology as opposed to only Hicks-neutral technology. In practice this amounts to separating firms into groups that (are expected to) differ in their elasticities, for instance because of differences in productivity (Raval 2023). However, this condition is clearly violated if not all inputs are perfectly flexible (as is the case for capital and labor) and there is an industry-wide shock.

Equation (A.6) coupled with the assumption of constant-returns-to-scale on the inputs over which the firm optimizes lets us recover the quantity $\lambda_i Y_i$ as the total cost of the firm. So, if there are N flexible inputs as in (1) we get the elasticity of output with respect to the n^{th} input as the input's cost-share:

$$\epsilon_n \;=\; \frac{E\left[P_n X_{n,i}\right]}{E\left[\sum_n^N P_n X_{n,i}\right]} \;. \tag{A.7}$$

Implementing (A.7) requires assuming that capital and labor are flexible on average and that the production function has constant returns to scale in all the observable inputs.

Remark This method also applies to cases where data on combined inputs is available. Say that one observes only the total expenditure in two inputs (labeled 1 and 2), then we can sum equation (A.6) applied to each input to obtain:

$$E\left[P_1 X_{1,i} \;+\; P_2 X_{2,i}\right] \;=\; (\epsilon_1 \;+\; \epsilon_2)\, E\left[\lambda_i Y_i\right] \;. \tag{A.8}$$

In this way, we get the sum of elasticities from the combined cost-share of the inputs:

$$\epsilon_1 \;+\; \epsilon_2 \;=\; \frac{E\left[P_1 X_{1,i} \;+\; P_2 X_{2,i}\right]}{E\left[\sum_n^N P_n X_{n,i}\right]}. \tag{A.9}$$

This is precisely the value needed to recover markups. Applying (2) to the two inputs and summing we can write:

$$\mu\,(\rho_1 + \rho_2) \;=\; \epsilon_1 + \epsilon_2 \;. \tag{A.10}$$

In practice this is useful when only data on cost-of-goods-sold is available, combining the costs of several inputs.

Assumptions and output-price bias Relying on cost shares to measure markups requires strong assumptions, namely constant-returns-to-scale and uniform output elasticities across firms. Moreover, cost shares only contain firms' revenues and expenditures. An alternative approach that avoids these strong assumptions is to estimate the output elasticity to variable inputs using a production function estimation approach as in De Loecker and Warzynski (2012).

However, the estimation of production functions when firm-level output without price data is also problematic, as it is only feasible to estimate revenue-based production functions instead of quantity-based production functions. Elasticity and markup estimates coming from revenue-based production functions suffer from "output-price bias" as shown in Bond et al. (2021). The bias arises because the price-setting behavior of firms in non-competitive markets make it impossible to establish a connection between revenue- and quantity-based production functions.

It is possible to counteract the output-price bias when prices are observable. Furthermore, under specific conditions, it becomes feasible to consistently derive markups from revenue data provided there is independent variation in variables associated with markups (Kirov, Mengano, and Traina 2026).

A.5. Markup Measures from Production

We consider several approaches to measure markups from firm level data. These approaches exploit the relationship between the firm's cost minimization problem and their markups, defined as the ratio of the price to the marginal cost of the firm, $\mu = p/\lambda$.

The cost minimization problem of the firm is:

$$C\left(y|\left\{p_n\right\}_{n=1}^{N},\{K_m\}_{m=1}^{M}\right)=\min_{\{x_n\}_{n=1}^{N}}\sum_{n=1}^{N}p_n\cdot x_n\quad\text{s.t. }\overline{y}\leq zF\left(x_1,\ldots,x_N,K_1,\ldots,K_M\right), \tag{A.11}$$

where production requires variable inputs $\{x_n\}_{n=1}^{N}$ and fixed inputs $\{K_m\}_{m=1}^{M}$. The marginal cost of production is captured by the Lagrangian multiplier (λ) of the production constraint. In equilibrium it holds that $\lambda = C'(y)$.

The first order condition with respect to variable input x_n is:

$$p_n=\lambda\frac{\partial F}{\partial x_n}. \tag{A.12}$$

Simple manipulation of this condition makes it possible to express the markup in terms of the cost-to-revenue share of input x_n and the elasticity of output with respect to the variable input x_n, ϵ_n:

$$\mu=\frac{p}{\lambda}=\frac{py}{p_nx_n}\epsilon_n=\frac{\epsilon_n}{s_n}. \tag{A.13}$$

Crucially, this condition is entirely compatible with the demand system and the optimal pricing behavior of firms in Section 3. Moreover, data on the cost-to-revenue share of inputs ($s_n \equiv p_n x_n / py$) is readily available from manufacturing surveys, economic censuses, or public firm data like compustat.

Appendix B. Additional Results

B.1. Data Description: Colombia and US

Colombia: We utilize the markup dataset from Raval (2023), which provides estimates derived via the cost-share approach. While Raval offers markups based on three flexible inputs—materials, labor, and a labor-materials composite—we specifically employ the estimates based on labor cost ratios. The underlying data originates from Colombia's Encuesta Anual Manufacturera (EAM). Our analysis focuses on the 1980–1989 subsample, comprising 12,848 firm-year observations across 2,551 unique firms.

US: For the U.S., we use markup estimates from De Loecker, Eeckhout, and Unger (2020), based on firm-level data from Compustat. These estimates utilize a production approach augmented by time-varying production function parameters to account for the long sample period. The variable input is defined as a bundle of all expenses attributable to the cost of goods sold (COGS), interpreted as a composite of labor, energy, and materials. We restrict the sample to the 1980–1989 period for our calibration, covering 5,192 firm-year observations and 790 unique firms.

B.2. Dispersion of measured markups

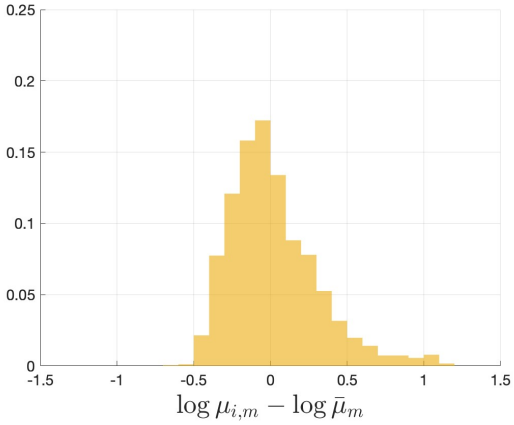
TABLE B.1. Markup Variance Decomposition – 1 percentage point bins

	Variance	Share within	Share between
India	0.196	0.564	0.436
Colombia	0.059	0.524	0.476
United States	0.024	0.386	0.614

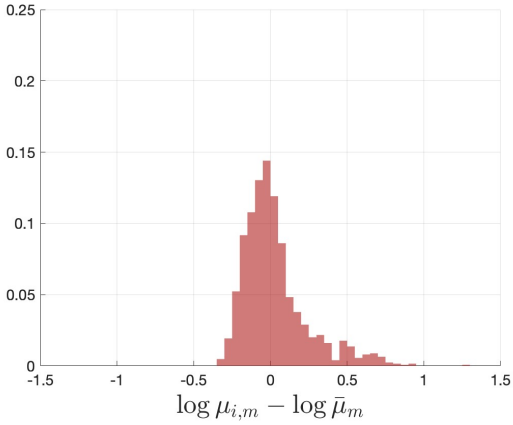
Notes: The variance is computed from the log ratio $\log(\mu_i^m/\mu_m)$. Numbers are averages across markets. To decompose the variance of markups, we divide firms in equally sized bins of 1 percentage point market share. The variance and its decomposition correspond to the average variance and shares across market-year pairs.

FIGURE B.1. Distribution of Markups in Colombian and US markets

(A) Colombia



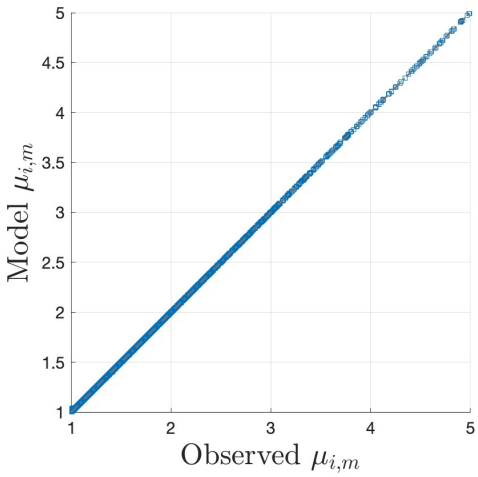
(B) United States



Notes: Panel A shows the histogram of markups relative to their market's average in Colombia. Markets are defined at the 3-digit ISIC rev 2 classification. Panel B shows the histogram of markups relative to their market's average for the Indian data in the US. Markets are defined at the 3-digit NAICS classification.

B.3. Bertrand competition

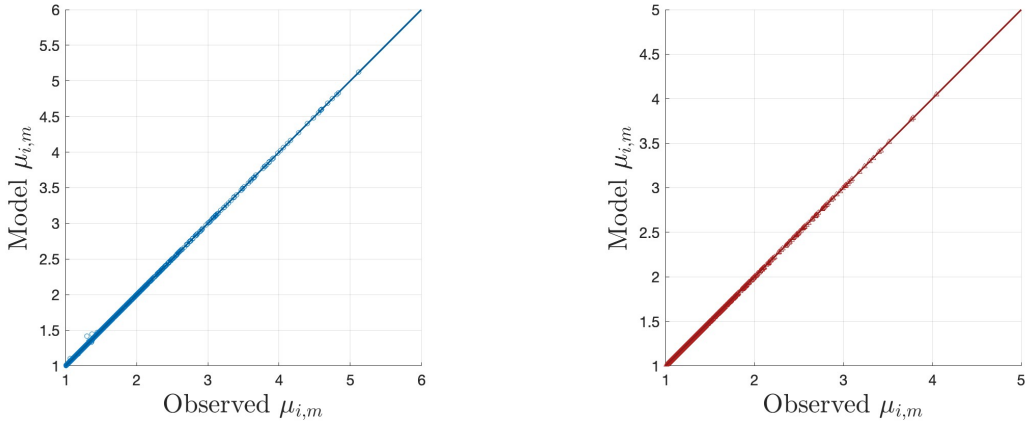
FIGURE B.2. Markup Fit in Indian Data (2001-2004)



Notes: The figure shows the scatterplot of measured and model implied markups under Bertrand competition

FIGURE B.3. Markups Fit in Colombian and US markets

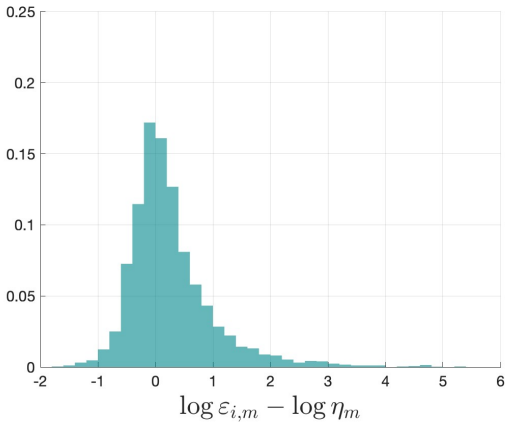
(A) Colombia (B) United States



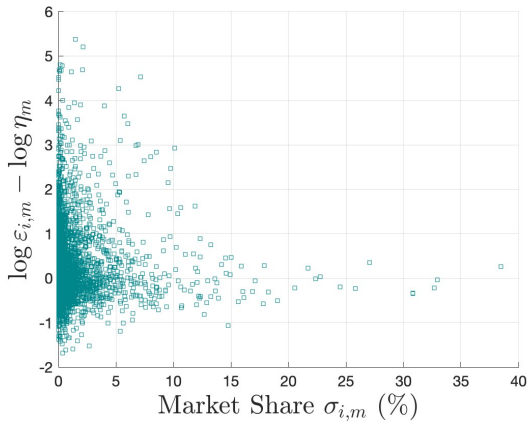
Notes: Panel A shows the scatterplot of measured and model-implied markups under Bertrand competition for Colombia and Panel B shows the scatterplot of measured and model-implied markups under Bertrand competition for US.

FIGURE B.4. Distribution of Own-Elasticities (ε) and Elasticity Shifter (ν) (2001–2004)

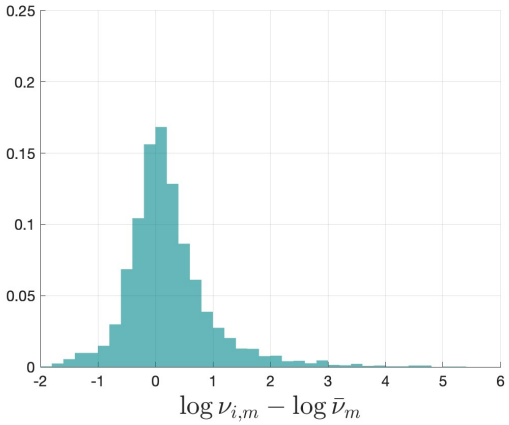
(A) Own-Elasticity Distribution



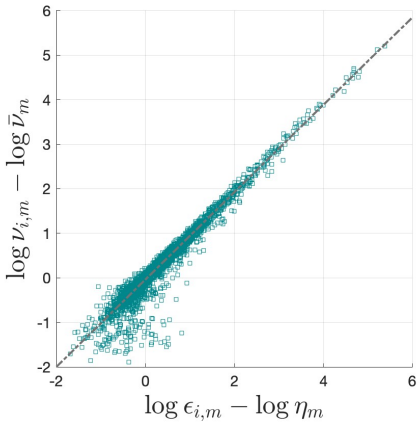
(B) Elasticity and Market-Share



(C) ν Distribution



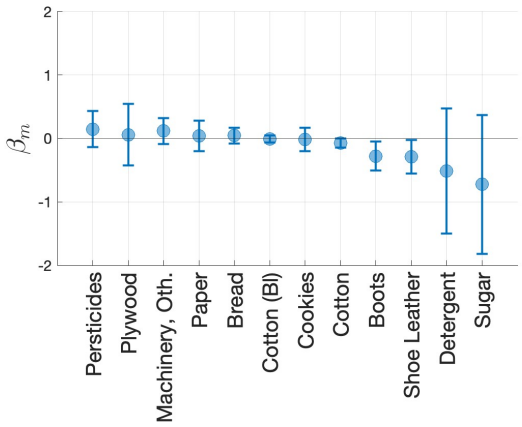
(D) ν and Elasticities



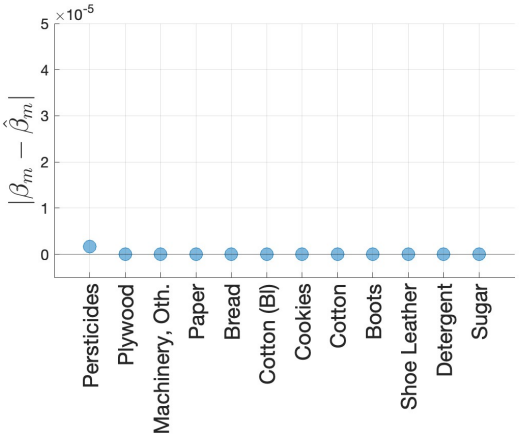
Notes: Panel A shows the histogram of own-elasticity of demand relative to their market's average for the Indian data covering the period 2005–2007. Panel B shows the joint scatter plot of own-elasticities and market shares. Panel C shows the histogram of demand elasticity shifters relative to their market's average for the Indian data covering the period 2005–2007. Panel D shows the joint scatter plot of own-elasticities and demand elasticity shifters. Markets are defined at the 5-digit ASICC product level.

FIGURE B.5. Demand Estimation – Selected Products (2005–2007)

(A) Standard Errors for β



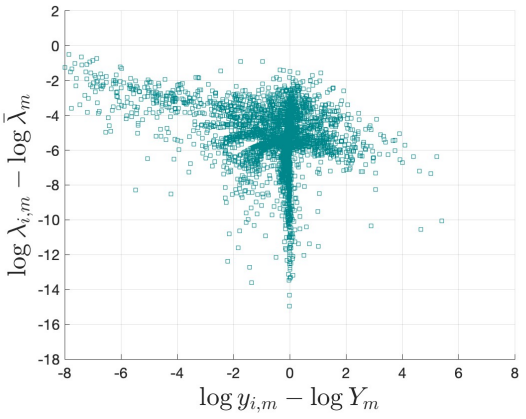
(B) Indirect Inference Residuals



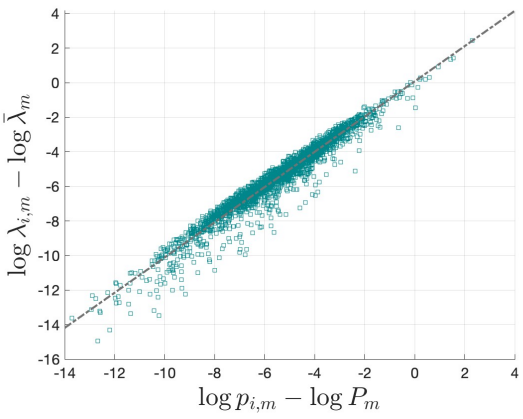
Notes: Panel A reproduces Figure 5A along with standard errors. Panel B plots the residuals of the demand estimation captured by the absolute difference between the coefficient β in equation (28) and its model counterpart in Proposition 4. We impose the Klenow and Willis aggregator in both cases.

FIGURE B.6. Marginal Costs (λ) and Relationship with Output and Prices

(A) λ and Output y_i/y_m



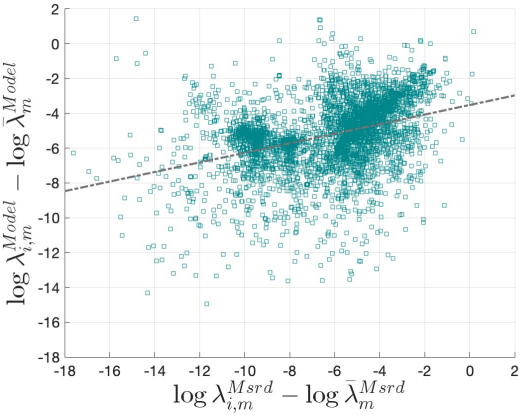
(B) λ and Prices P_i/P_m



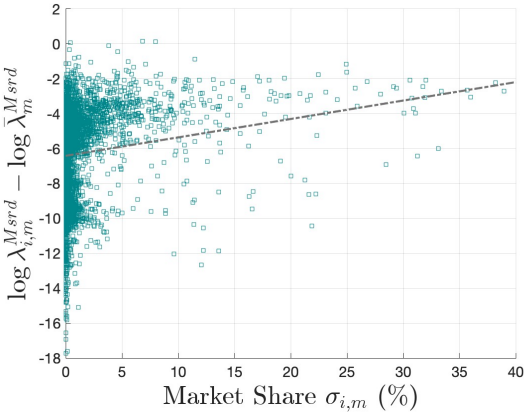
Notes: Panel A shows the scatter plot of firm log-marginal costs relative to their market's marginal cost and log-markups relative to their market's average. Panel B shows the scatter plot of firm log-marginal costs relative to their market share. Marginal costs are computed as in (35), relative output $\frac{y_i}{Y}$ obtained as in (33), and relative prices as in (34).

FIGURE B.7. Marginal Costs Fit and Relationship with Market Shares (2005-2007)

(A) Measured λ and Model λ



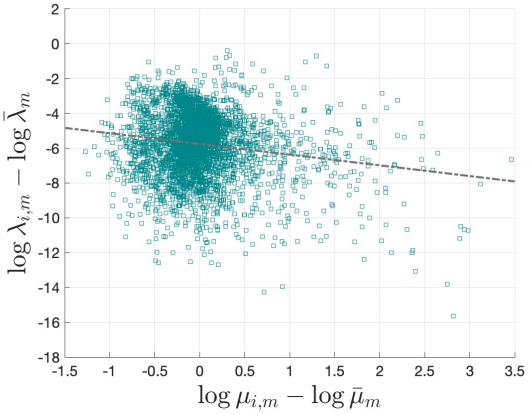
(B) Measured λ and Market-Share



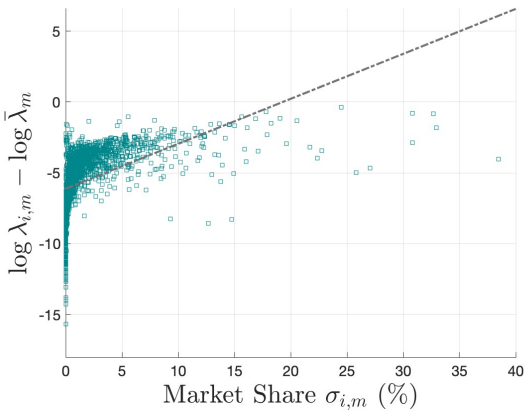
Notes: Panel A shows the scatter of the model's marginal costs λ from (35) and the measured marginal costs, recovered as the ratio between observed output prices and the estimated markup from Section 2. Panel B shows the scatter of the measured marginal costs and market shares.

FIGURE B.8. Marginal Costs and Relationship with Markups and Market Shares (2001-2004)

(A) λ and Markups



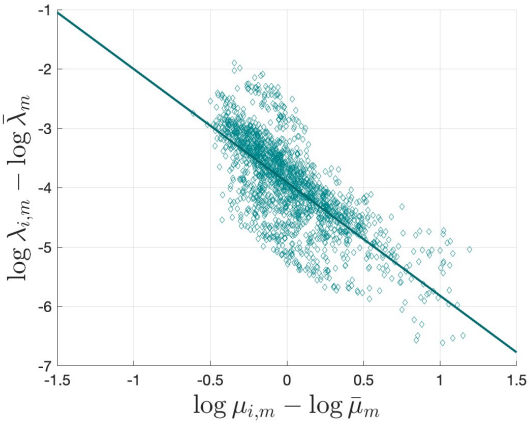
(B) λ and Market-Share



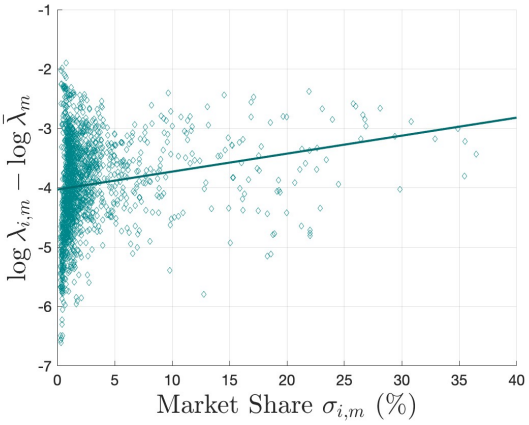
Notes: Panel A shows the scatter plot of firm log-marginal costs relative to their market's marginal cost and log-markups relative to their market's average. Panel B shows the scatter plot of firm log-marginal costs relative to their market share. Marginal costs are computed as in (35).

FIGURE B.9. Marginal Costs and Relationship with Markups and Market Shares

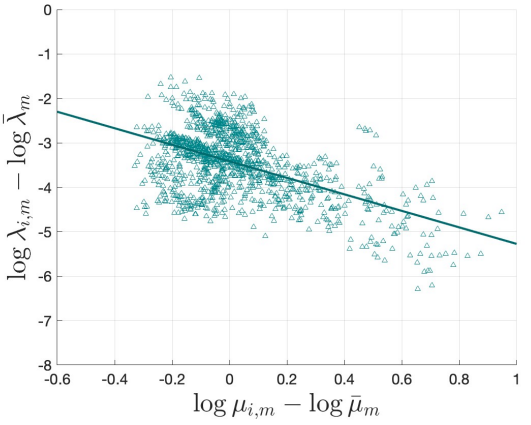
(A) **Colombia**: λ and Markups



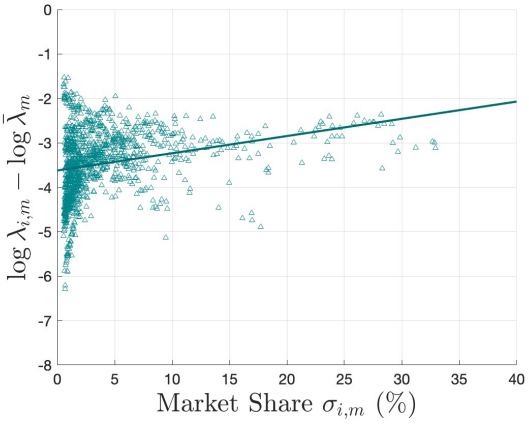
(B) **Colombia**: λ and Market-Share



(C) **US**: λ and Markups



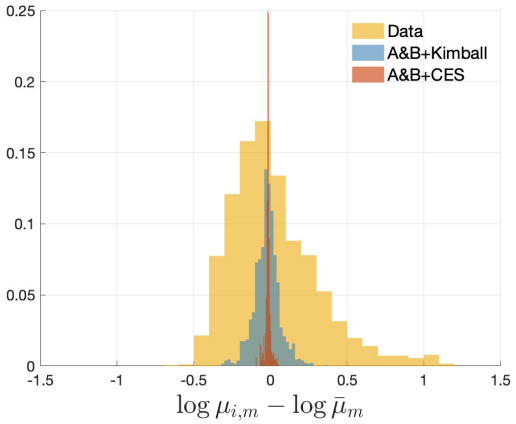
(D) **US**: λ and Market-Share



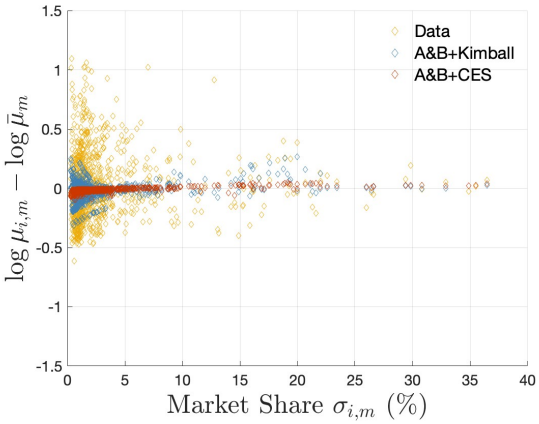
Notes: Panels (a) and (c) show the scatter plot of firm log-marginal costs relative to their market's marginal cost and log-markups relative to their market's average, for Colombia and the US, respectively. Panels (b) and (d) show the scatter plot of firm log-marginal costs relative to their market share, for Colombia and the US, respectively. Marginal costs are computed as in (35).

FIGURE B.10. Markups and Market Share Distribution Across Models

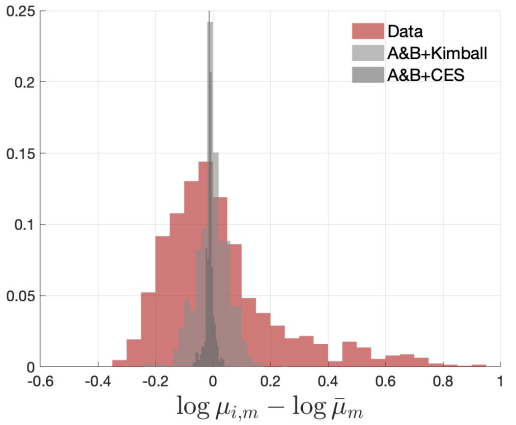
(A) Colombia: Markups Distribution



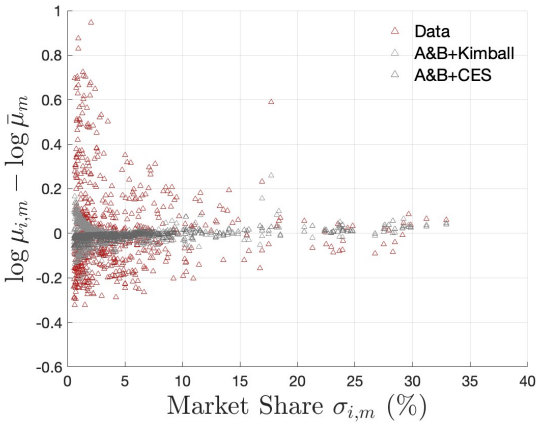
(B) Colombia: Markups and Market-Share



(C) United States: Markups Distribution



(D) United States: Markups and Market-Share



Notes: The left panels show the histogram of markups relative to their market's average for the data (blue bars) and the markups implied by the Atkeson and Burstein (2008) Bertrand model, taking market shares as given and adjusting each industry's (common) elasticity of demand to match the market's average markup. The right panels show the relationship between markups and market shares.

TABLE B.2. Markup Variance Decomposition Colombia and US (1 pp bins)

(a) Colombia

	Variance	Share within	Share between	Corr($\log \mu_i$, σ_i)
Data	0.059	0.524	0.476	0.029
Full Model	0.059	0.525	0.475	0.030
Atkeson & Burstein	0.001	0.001	0.999	0.608
Kimball	0.006	0.012	0.988	0.213
A&B + Kimball	0.006	0.021	0.979	0.183

(b) United States

	Variance	Share within	Share between	Corr($\log \mu_i$, σ_i)
Data	0.024	0.386	0.614	-0.075
Full Model	0.024	0.386	0.614	-0.074
Atkeson & Burstein	0.000	0.001	0.999	0.628
Kimball	0.002	0.008	0.992	-0.018
A&B + Kimball	0.002	0.014	0.986	-0.061

Notes: The table presents statistics for measured markups and market shares along with statistic from four models: our full model described in Section 3, a model of oligopolistic competition with constant elasticity of demand, a model of monopolistic competition with a common demand with variable elasticity, and a model of oligopolistic competition with variable elasticity of demand. The variance is computed from the log ratio $\log(\mu_i^m/\mu_m)$. Numbers are averages across markets. To decompose the variance of markups, we divide firms in equally sized bins of 1 percentage point market share. The variance and its decomposition correspond to the average variance and shares across market-year pairs. The correlation between markups and market shares is the pooled correlation of the full sample.

TABLE B.3. Markup Variance Decomposition India - 1pp Bins (2001-2008)

	Variance	Share within	Share between	Corr(log μ_i, σ_i)
Data	0.196	0.564	0.436	0.061
Full Model	0.196	0.564	0.436	0.061
A&B + Kimball	0.047	0.110	0.890	0.211
Kimball	0.048	0.083	0.917	0.217
Atkeson & Burstein	0.001	0.015	0.985	0.337

Notes: The table presents statistics for measured markups and market shares along with statistic from four models: our full model described in Section 3, a model of oligopolistic competition with constant elasticity of demand, a model of monopolistic competition with a common demand with variable elasticity, and a model of oligopolistic competition with variable elasticity of demand. The variance is computed from the log ratio $\log(\mu_i^m/\mu_m)$. Numbers are averages across markets. To decompose the variance of markups, we divide firms in equally sized bins of 1 percentage point market share. The variance and its decomposition correspond to the average variance and shares across market-year pairs. The correlation between markups and market shares is the pooled correlation of the full sample.

TABLE B.4. Markup Variance Decomposition India - Single Product Firms (2001-2008)

	Variance	Share within	Share between	Corr(log μ_i, σ_i)	Mean Log Deviation
Data All	0.196	0.770	0.230	0.061	0.078
Data Single Product	0.184	0.549	0.451	0.076	0.072

Notes: The table presents statistics for measured markups and market shares for the entire sample and for only single product firms. The variance is computed from the log ratio $\log(\mu_i^m/\mu_m)$. Numbers are averages across markets. To decompose the variance of markups, we divide firms in equally sized bins of 5 percentage point market share. The variance and its decomposition correspond to the average variance and shares across market-year pairs. The correlation between markups and market shares is the pooled correlation of the full sample.

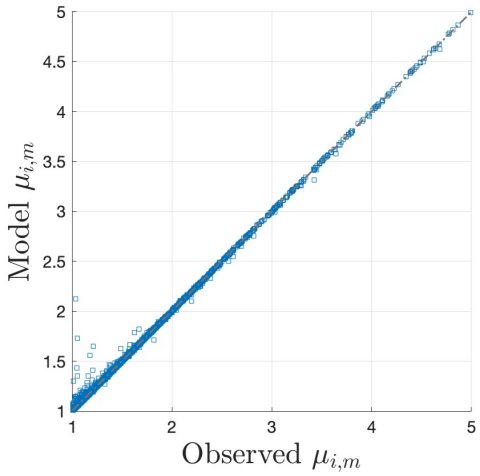
TABLE B.5. Markup Variance Decomposition India (2001-2008) - Robustness

	Baseline	Avg ac. years	at least 2	at least 3	at least 4
Variance	0.196	0.232	0.178	0.169	0.115
% Within (1pp)	0.564	0.833	0.704	0.550	0.400
% Within (5pp)	0.770	0.961	0.860	0.783	0.692
Avg Number firms	386	386	142	66	32

Notes: The table reports the decomposition of observed markup variance under several sample specifications. Column 2 presents the decomposition using the simple firm-level average markup across all years in which the firm is observed. Column 3 restricts the sample to firms observed producing the same product for at least two years and computes the simple average over those years. Columns 4 and 5 further restrict the sample to firms observed for at least three and four years, respectively.

B.4. Cournot competition

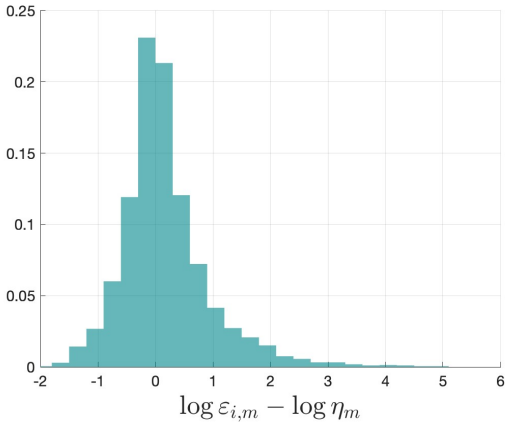
FIGURE B.11. Markup Fit in Indian Data Under Cournot Competition (2005–2007)



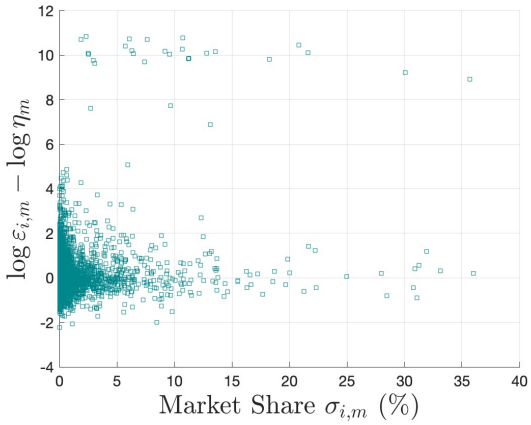
Notes: The figure shows the scatterplot of measured and model-implied markups under Cournot competition, obtained by solving for $\{\varepsilon_i^m\}$ in equation 18 given $\{\mu_i^m, \sigma_i^m\}$. We set $\gamma = 1.3$. A value in the 45 degree line corresponds to a perfect match of the model for a given firm.

FIGURE B.12. Distribution of Elasticities - Cournot (2005–2008)

(A) Own-Elasticity Distribution



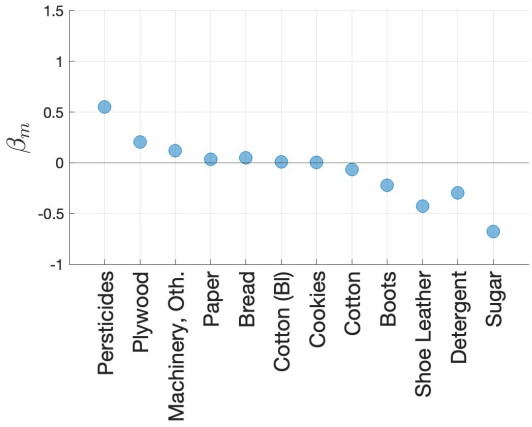
(B) Elasticity and Market-Share



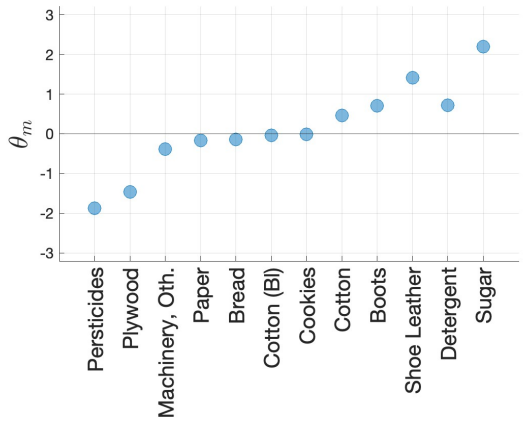
Notes: Panel A shows the histogram of own-elasticity of demand relative to their market's average for the Indian data covering the period 2005–2007. Panel B shows the joint scatter plot of own-elasticities and market shares. Markets are defined at the 5-digit ASICC product level.

FIGURE B.13. Demand Parameters - Cournot - Selected Markets (2005–2008)

(A) $\hat{\beta}_m$ by Market



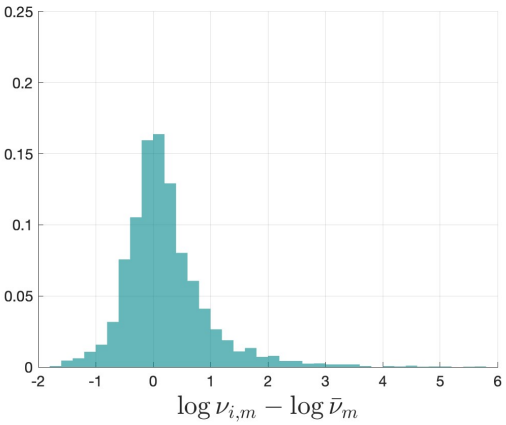
(B) θ_m by Market



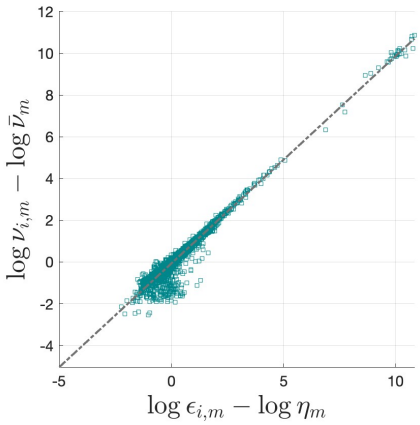
Notes: The left panel shows the $\hat{\beta}$'s estimated from the data using equation (28) and the market-specific β implied by the Klenow and Willis aggregator in Appendix F. The right panel shows the market-specific θ estimated that generate the β that minimize the distance with respect to the observed $\hat{\beta}$.

FIGURE B.14. Elasticity Shifter (ν) and Relationship with Elasticities - Cournot

(A) ν Distribution



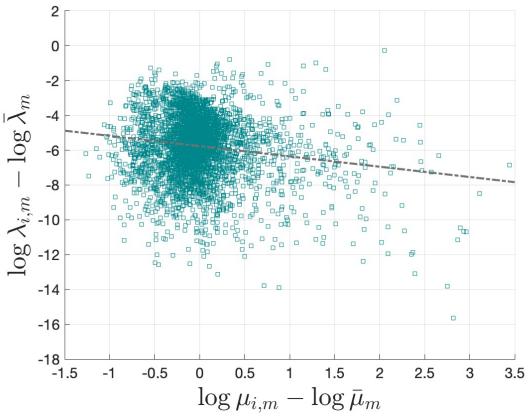
(B) ν and Elasticities



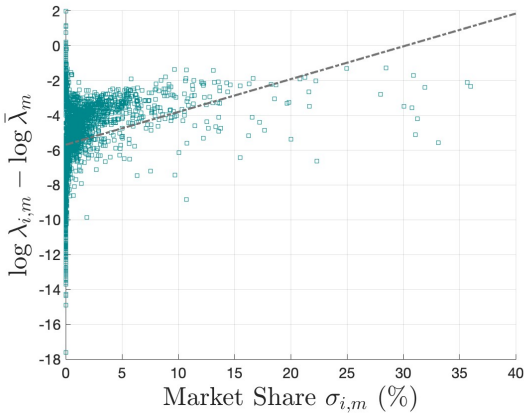
Notes: Panel A shows the histogram of demand elasticity shifters relative to their market's average for the Indian data covering the period 2005–2007. Panel B shows the joint scatter plot of own-elasticities and demand elasticity shifters. Markets are defined at the 5-digit ASICC product level.

FIGURE B.15. Marginal Costs and Relationship with Markups and Market Shares - Cournot

(A) λ and Markups

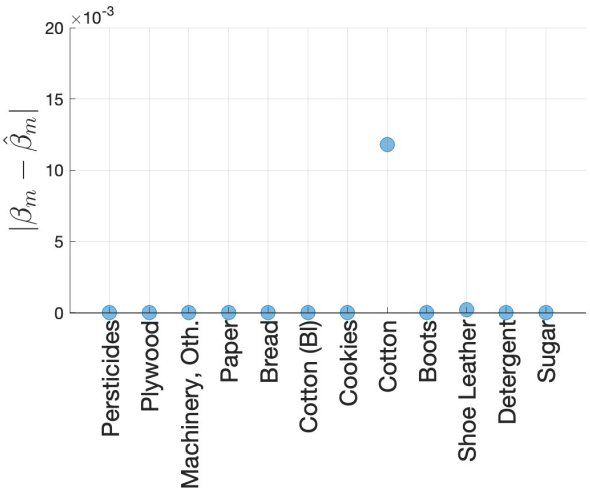


(B) λ and Market-Share



Notes: Panel A shows the scatter plot of firm log-marginal costs relative to their market's marginal cost and log-markups relative to their market's average. Panel B shows the scatter plot of firm log-marginal costs relative to their market share. Marginal costs are computed as in (35).

FIGURE B.16. Demand Estimation Residuals – Selected Products (2005–2007)



Notes: The figures plot the residuals of the demand estimation captured by the absolute difference between the coefficient β in equation (28) and its model counterpart in Proposition 4. We impose the Klenow and Willis aggregator in both cases.

Appendix C. Granular competition with HDIA aggregators

C.1. Oligopolistic Competition Setup

The aggregate good in a market m is produced using N differentiated inputs, each produced by a monopolist. Market aggregators behave competitively and aggregate intermediate inputs operating a technology implicitly defined by the Kimball (1995) aggregator. The problem of the good-m producer is:

$$\min_{y_i^m} \sum_{i=1}^{N_m} p_i^m y_i^m \qquad \text{s.t.} \quad 1 = \sum_{i=1}^{N_m} \Upsilon\left(\frac{y_i^m}{Y_m}\right), \tag{C.1}$$

where y_i^m is the quantity of variety i purchased. The first order conditions are

$$p_i^m = \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{\Lambda_m}{Y_m} \tag{C.2}$$

where Λ_m is the Lagrange multiplier associated with the constraint.

To arrive to the demand for variety i in (7) we need to solve for the Lagrange Multiplier, Λ_m. To do this, multiply both sides of (C.2) by y_i^m and sum across varieties to obtain

$$\Lambda_m = \frac{\sum_i p_i^m y_i^m}{\sum_i \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}. \tag{C.3}$$

Then, define the price of the sectoral good, P_m so that it satisfies

$$P_m \cdot Y_m = \sum_{i=1}^{N_m} p_i^m \cdot y_i^m. \tag{C.4}$$

Finally, replace Λ_m into (C.2) and divide both sides by P_m to obtain (7):

$$\frac{p_i^m}{P_m} = \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{\Lambda_m}{P_m Y_m} = \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{\sum_i \frac{p_i^m y_i^m}{P_m Y_m}}{\sum_i \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}} = \frac{\Upsilon'\left(\frac{y_i^m}{Y_m}\right)}{\sum_i \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}. \tag{C.5}$$

Having derived the demand for variety i we can further characterize the P_m following Boar and Midrigan (2024). It is useful to define the auxiliary variable D_m,

$$D_m = \sum_{i=1}^{N_m} \Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}. \tag{C.6}$$

Then, the sectoral price is obtained implicitly as:

$$P_m = \frac{\sum_i p_i^m y_i^m}{Y_m} = \frac{\Lambda_m \cdot D_m}{Y_m}. \tag{C.7}$$

This allow us to characterize P_m by inverting (C.2) to get

$$\frac{y_i^m}{Y_m} = \left(\Upsilon'\right)^{-1}\left(Y_m \frac{p_i^m}{\Lambda_m}\right) = \left(\Upsilon'\right)^{-1}\left(D_m \frac{p_i^m}{P_m}\right). \tag{C.8}$$

Replacing on the definition of the price index gives

$$P_m = \sum_{i=1}^{N} p_i^m \frac{y_i^m}{Y_m} = \sum_{i=1}^{N} p_i^m \left(\Upsilon'\right)^{-1}\left(D_m \frac{p_i^m}{P_m}\right). \tag{C.9}$$

Then, from the definition of the aggregator (6) we get

$$1 = \sum_{i=1}^{N_m} \Upsilon\left(\left(\Upsilon'\right)^{-1}\left(D_m \frac{p_i^m}{P_m}\right)\right). \tag{C.10}$$

This delivers a pair of equations that jointly determine P_m and D_m given prices.

C.2. Proofs for market shares, elasticities, and markups

Lemma A1. ***(Market Shares)*** *The change in market output Y_m to firm i's output and of market price P_m to firm i's price satisfy*

$$\frac{\partial Y_m}{\partial y_i^m} = \frac{p_i^m}{P_m} \qquad \text{and} \qquad \frac{\partial P_m}{\partial p_i^m} = \frac{y_i^m}{Y_m}. \tag{C.11}$$

So, the market share of firm i in market m gives the elasticity of market output and market price to changes in the firm's output and price, respectively,

$$\sigma_i^m \equiv \frac{p_i^m y_i^m}{P_m Y_m} = \frac{\Upsilon_i'\left(\frac{y_i^s}{Y_s}\right)\frac{y_i^s}{Y_s}}{\sum_j \Upsilon_j'\left(\frac{y_j^s}{Y_s}\right)\frac{y_j^s}{Y_s}} = \frac{y_i^m}{Y_m}\frac{\partial Y_m}{\partial y_i^m} = \frac{p_i^m}{P_m}\frac{\partial P_m}{\partial p_i^m}. \tag{C.12}$$

Proof. Equation (9) comes from the definition of market shares and the (inverse) demand for good i in (7).

We now prove that $\partial Y_m/\partial y_i^m = p_i^m/P_m$. From the aggregation technology (6), we obtain by totally differentiating,

$$0 = \sum_{j=1}^{N_m} \Upsilon'\left(\frac{y_j^m}{Y_m}\right)\left(\frac{1}{Y_m} dy_j^m - \frac{y_j^m}{Y_m^2} dY_m\right). \tag{C.13}$$

So that the derivative with respect to firm i's output, y_i^m is

$$\frac{dY_m}{dy_i^m} = \frac{\Upsilon'\left(\frac{y_i^m}{Y_m}\right)}{\sum_{i=1}^{N}\Upsilon'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}} = \frac{p_i^m}{P_m}. \tag{C.14}$$

We now prove that $\partial P_m/\partial p_i^m = y_i^m/Y_m$. The proof works in three steps. First define an

auxiliary variable $D_m \equiv \sum_{h=1}^{N_m} \Upsilon'_h \left(y_h^m/Y_m\right) \cdot y_h^m/Y_m$, this will make the remaining steps more tractable. Second establish restrictions on the changes in total demand from a change in prices. For this consider equations (6) and (7) to get

$$1 = \sum_{j=1}^{N_m} \Upsilon\left(\left(\Upsilon'\right)^{-1}\left(D_m \frac{p_j^m}{P_m}\right)\right). \tag{C.15}$$

Differentiate with respect to p_i^m we get

$$0 = \sum_{j=1}^{N_m} \frac{\Upsilon'\left(\frac{y_j^m}{Y_m}\right)}{\Upsilon''\left(\frac{y_j^m}{Y_m}\right)} \left[\frac{\partial}{\partial p_i^m} D_m \frac{p_j^m}{P_m}\right]. \tag{C.16}$$

Finally, from the definition of the price index we get

$$1 = \sum_{j=1}^{N_m} \left(\Upsilon'\right)^{-1}\left(D_m \frac{p_j^m}{P_m}\right) \frac{p_j^m}{P_m}. \tag{C.17}$$

Differentiate with respect to p_i^m we get

$$0 = \sum_{j=1}^{N_m} \left[\frac{p_j^m}{P_m} \frac{\partial}{\partial p_i^m} \left(\Upsilon'\right)^{-1}\left(D_m \frac{p_j^m}{P_m}\right) + \frac{y_j^m}{Y_m} \frac{\partial}{\partial p_i^m} \frac{p_j^m}{P_m}\right] \tag{C.18}$$

$$0 = \sum_{j=1}^{N_m} \frac{\frac{p_j^m}{P_m}}{\Upsilon''\left(\frac{y_j^m}{Y_m}\right)} \left[\frac{\partial}{\partial p_i^m} D_m \frac{p_j^m}{P_m}\right] - \sum_{j=1}^{N_m} \frac{y_j^m}{Y_m} \frac{p_j^m}{P_m} \frac{1}{P_m} \frac{\partial P_m}{\partial p_i^m} + \frac{y_i^m}{Y_m} \frac{1}{P_m}$$

$$0 = \underbrace{\frac{1}{D_m} \sum_{j=1}^{N_m} \frac{\Upsilon'\left(\frac{y_j^m}{Y_m}\right)}{\Upsilon''\left(\frac{y_j^m}{Y_m}\right)} \left[\frac{\partial}{\partial p_i^m} D_m \frac{p_j^m}{P_m}\right]}_{=0 \text{ by (C.16)}} - \frac{1}{P_m}\left(\underbrace{\left(\sum_{j=1}^{N_m} \sigma_j^m\right)}_{=1} \frac{\partial P_m}{\partial p_i^m} - \frac{y_i^m}{Y_m}\right)$$

$$0 = -\frac{\partial P_m}{\partial p_i^m} + \frac{y_i^m}{Y_m},$$

which gives the result. See Matsuyama (2023) for a general proof for homothetic demand. □

Proposition A1. ***(Elasticities)*** *The elasticity of firm i in market m satisfies*

$$\textbf{\textit{Cournot}} \quad \frac{1}{\eta_i^m} = \underbrace{\frac{1}{\gamma}}_{\text{Market Elasticity}} \sigma_i^m + \underbrace{\left(\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right) + E_\sigma\left[\frac{1}{\varepsilon_j^m}\middle| j \neq i\right]\sigma_i^m\right)}_{\text{Variety Elasticity}}\left(1-\sigma_i^m\right) \; ; \tag{C.19}$$

$$\textbf{\textit{Bertrand}} \quad \eta_i^m = \underbrace{\gamma}_{\text{Market Elasticity}} \sigma_i^m + \underbrace{\varepsilon_i^m \frac{E_\sigma\left[\varepsilon_j^m | j \neq i\right]}{E_\sigma\left[\varepsilon_j^m\right]}}_{\text{Variety Elasticity}}\left(1-\sigma_i^m\right) \; ; \tag{C.20}$$

where $\sigma_i^m = p_i^m y_i^m / P_m Y_m$ is firm i's market share, ε_i^m is its own-elasticity as in (13), $E_\sigma\left[x_j\right] = \sum_j^{N_m} x_j \sigma_j^m$ is the average with respect to expenditure in market m, and $E_\sigma\left[x_j | j \neq i\right] = \sum_{j\neq i} x_j \frac{\sigma_j^m}{1-\sigma_i^m}$ is the average with respect to firm i's competitors.

Proof. We start by tackling Cournot competition. We take as given the output of competitors. The total (or average) elasticity of demand is

$$\left(\eta_i^m\right)^{-1} = -y_i^m \frac{\partial \log p_i^m}{\partial y_i^m} = -y_i \frac{\partial \log\left(\frac{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)\alpha_m\left(\frac{Y_m}{Y}\right)^{-\frac{1}{\gamma}} P}{\sum_j \Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\right)}{\partial y_i} \tag{C.21}$$

So that we can express the elasticity in terms of three terms and solve using Lemma 1 and equation (13).

$$\frac{1}{\eta_i^m} = -y_i^m\left[\frac{\partial \log \Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}{\partial y_i^m} + \frac{\partial \log \alpha_m\left(\frac{Y_m}{Y}\right)^{-\frac{1}{\gamma}}}{\partial y_i} - \frac{\partial \log \sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}}{\partial y_i^m}\right] \tag{C.22}$$

$$= -y_i^m\left[\frac{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)}{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}\frac{\partial\left(\frac{y_i^m}{Y_m}\right)}{\partial y_i^m} - \frac{1}{\gamma}\frac{1}{Y_m}\frac{\partial Y_m}{\partial y_i^m}\right] \tag{C.23}$$

$$+ \frac{y_i^m}{\sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}}\left(\frac{\partial \Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}{\partial y_i^m} + \sum_{j\neq i}\frac{\partial \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_{mj}}{Y_m}}{\partial y_i^m}\right)$$

$$= -\frac{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}\left(1 - \frac{y_i^m}{Y_m}\frac{\partial Y_m}{\partial y_i^m}\right) + \frac{1}{\gamma}\frac{y_i^m}{Y_m}\frac{\partial Y_m}{\partial y_i^m} \tag{C.24}$$

$$+\frac{\frac{y_i^m}{Y_m}}{\sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}}\left(\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}+\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)\right)\left(1-\frac{y_i^m}{Y_m}\frac{\partial Y_m}{\partial y_i^m}\right)$$
$$-\frac{\frac{y_i^m}{Y_m}}{\sum_j \Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}}\sum_{j\neq i}\left(\Upsilon_j''\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}+\Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\right)\left(\frac{y_j^m}{Y_m}\frac{\partial Y_m}{\partial y_i^m}\right)$$

At this point we suppress the arguments of Υ and we take advantage of Lemma 1 that gives us $\frac{\partial Y_m}{\partial y_i^m}=\frac{p_i^m}{P_m}$ to group terms and obtain market shares σ_i^m.

$$\frac{1}{\eta_i^m}=-\frac{\Upsilon_i''\cdot\frac{y_i^m}{Y_m}}{\Upsilon_i'}\left(1-\sigma_i^m\right)+\frac{1}{\gamma}\sigma_i^m+\frac{\frac{y_i^m}{Y_m}}{\sum_j \Upsilon_j'\cdot\frac{y_j^m}{Y_m}}\left(\Upsilon_i''\cdot\frac{y_i^m}{Y_m}+\Upsilon_i'\right)\left(1-\sigma_i^m\right) \tag{C.25}$$
$$-\frac{1}{\sum_j \Upsilon_j'\cdot\frac{y_j^m}{Y_m}}\sum_{j\neq i}\left(\Upsilon_j''\cdot\frac{y_j^m}{Y_m}+\Upsilon_j'\right)\left(\frac{y_j^m}{Y_m}\sigma_i^m\right)$$
$$=-\frac{\Upsilon_i''\cdot\frac{y_i^m}{Y_m}}{\Upsilon_i'}\left(1-\sigma_i^m\right)+\frac{1}{\gamma}\sigma_i^m+\frac{\Upsilon_i'\cdot\frac{y_i^m}{Y_m}}{\sum_j \Upsilon_j'\cdot\frac{y_j^m}{Y_m}}\left(\frac{\Upsilon_i''\cdot\frac{y_i^m}{Y_m}}{\Upsilon_i'}+1\right)\left(1-\sigma_i^m\right) \tag{C.26}$$
$$-\sum_{j\neq i}\left(\frac{\Upsilon_j''\cdot\frac{y_j^m}{Y_m}}{\Upsilon_j'}+1\right)\frac{\Upsilon_j'\cdot\frac{y_j^m}{Y_m}}{\sum_j \Upsilon_j'\cdot\frac{y_j^m}{Y_m}}\sigma_i^m$$
$$=\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right)+\frac{1}{\gamma}\sigma_i^m+\sigma_i^m\left(1-\frac{1}{\varepsilon_i^m}\right)\left(1-\sigma_i^m\right)-\sum_{j\neq i}\left(1-\frac{1}{\varepsilon_j^m}\right)\sigma_j^m\sigma_i^m, \tag{C.27}$$

where we use the definition of the own-elasticity ε_i^m in (13) and equation (9) from Lemma 1. The result is obtained after some manipulation:

$$\frac{1}{\eta_i^m}=\frac{1}{\gamma}\sigma_i^m+\left[\frac{1}{\varepsilon_i^m}+\left(1-\frac{1}{\varepsilon_i^m}\right)\cdot\sigma_i^m-\sigma_i^m\sum_{j\neq i}\left(1-\frac{1}{\varepsilon_j^m}\right)\cdot\frac{\sigma_j^m}{1-\sigma_i^m}\right]\left(1-\sigma_i^m\right) \tag{C.28}$$
$$=\frac{1}{\gamma}\sigma_i^m+\left[\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right)+\sigma_i^m\sum_{j\neq i}\frac{1}{\varepsilon_j^m}\cdot\frac{\sigma_j^m}{1-\sigma_i^m}\right]\left(1-\sigma_i^m\right)$$
$$=\frac{1}{\gamma}\sigma_i^m+\left[\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right)+E_\sigma\left[\frac{1}{\varepsilon_j^m}\middle|\, j\neq i\right]\sigma_i^m\right]\left(1-\sigma_i^m\right),$$

where

$$E_\sigma\left[\frac{1}{\varepsilon_j^m}\middle| j \neq i\right] \equiv \sum_{j\neq i}\frac{1}{\varepsilon_j^m}\frac{\sigma_j^m}{1-\sigma_i^m}. \tag{C.29}$$

We can also write the elasticity as

$$\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right)+\frac{1}{\eta_{-i}^m}\sigma_i^m = \sum_{j\neq i}\frac{\sigma_j^m}{1-\sigma_i^m}\left[\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right)+\sigma_i^m\frac{1}{\varepsilon_j^m}\right]. \tag{C.30}$$

So that

$$\frac{1}{\eta_i^m} = \sigma_i^m\frac{1}{\gamma}+\left(1-\sigma_i^m\right)\sum_{j\neq i}\frac{\sigma_j^m}{1-\sigma_i^m}\left[\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right)+\sigma_i^m\frac{1}{\varepsilon_j^m}\right]. \tag{C.31}$$

We now tackle Bertrand competition. We take as given the prices of competitors. The total (or average) elasticity of demand is

$$\eta_i^m = -p_i^m\frac{\partial\log y_i^m}{\partial p_i^m} = -p_i^m\frac{\partial\log\left(Y_m\cdot\left(\Upsilon_i'\right)^{-1}\left(\left(\sum_j\Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\cdot\frac{y_j^m}{Y_m}\right)\frac{p_i^m}{P_m}\right)\right)}{\partial p_i^m}. \tag{C.32}$$

This allows us to break down the elasticity into three terms

$$\eta_i^m = \underbrace{-p_i^m\frac{\partial\log Y_m}{\partial p_i^m}}_{\text{Granular Effect}}\underbrace{-\frac{p_i^m\sum_j\Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\cdot\frac{y_j^m}{Y_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\partial p_i^m/P_m}{\partial p_i^m}}_{\text{Own-Elasticity}}\underbrace{-\frac{p_i^m\frac{p_i^m}{P_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\partial\sum_j\Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\cdot\frac{y_j^m}{Y_m}}{\partial p_i^m}}_{\text{Substitution Effects}}. \tag{C.33}$$

We proceed by tackling each effect separately.

$$\underbrace{-p_i^m\frac{\partial\log Y_m}{\partial p_i^m}}_{\text{Granular Effect}} = -p_i^m\frac{\partial\log Y\left(\frac{1}{\alpha_m}\frac{P_m}{P}\right)^{-\gamma}}{\partial p_i^m} = \gamma\frac{p_i^m}{P_m}\frac{\partial P_m}{\partial p_i^m} = \gamma\sigma_i^m \tag{C.34}$$

where the last step follows from Lemma 1.

Now to the own-elasticity. We solve for this only partially as it interacts with the substitution effects term.

$$\underbrace{-\frac{p_i^m\sum_j\Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\cdot\frac{y_j^m}{Y_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\partial p_i^m/P_m}{\partial p_i^m}}_{\text{Own-Elasticity}} = -\frac{\frac{p_i^m}{P_m}\sum_j\Upsilon_j'\left(\frac{y_j^m}{Y_m}\right)\cdot\frac{y_j^m}{Y_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\left(1-\frac{p_i^m}{P_m}\frac{\partial P_m}{\partial p_i^m}\right) \tag{C.35}$$

$$= -\frac{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\left(1-\sigma_i^m\right) \;=\; \varepsilon_i^m\left(1-\sigma_i^m\right)$$

Now the substitution term. As in Lemma 1 we define $D_m \equiv \sum_h \Upsilon_h'\left(y_h^m/Y_m\right)\cdot y_h^m/Y_m$ and consider first its derivative with respect to the price of good i. We start from (C.16)

$$\begin{aligned}
0 &= \sum_{j=1}^{N_m} \frac{\Upsilon'\left(\frac{y_j^m}{Y_m}\right)}{\Upsilon''\left(\frac{y_j^m}{Y_m}\right)}\left[\frac{\partial}{\partial p_i^m} D_m \frac{p_j^m}{P_m}\right] \qquad \text{(C.36)}\\
&= \sum_{j=1}^{N_m} \frac{\Upsilon'\left(\frac{y_j^m}{Y_m}\right)}{\Upsilon''\left(\frac{y_j^m}{Y_m}\right)}\left[\frac{\partial D_m}{\partial p_i^m}\frac{p_j^m}{P_m} + \frac{D_m}{P_m^2}\left(P_m\frac{\partial p_j^m}{\partial p_i^m} - p_j^m\frac{\partial P_m}{\partial p_i^m}\right)\right]\\
&= \sum_{j=1}^{N_m} \frac{\Upsilon'\left(\frac{y_j^m}{Y_m}\right)}{\Upsilon''\left(\frac{y_j^m}{Y_m}\right)\frac{y_j^m}{Y_m}}\frac{y_j^m}{Y_m}\left[\frac{1}{D_m}\frac{\partial D_m}{\partial p_i^m}p_j^m - \frac{p_j^m}{p_i^m}\sigma_i^m\right] + \frac{\Upsilon'\left(\frac{y_i^m}{Y_m}\right)}{\Upsilon''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{y_i^m}{Y_m}\\
&= -\sum_{j=1}^{N_m}\varepsilon_j^m\sigma_j^m\left[\frac{P_m}{D_m}\frac{\partial D_m}{\partial p_i^m} - \frac{P_m}{p_i^m}\sigma_i^m\right] - \varepsilon_i^m\frac{y_i^m}{Y_m}\\
&= -\left(\sum_{j=1}^{N_m}\varepsilon_j^m\sigma_j^m\right)\frac{P_m}{D_m}\frac{\partial D_m}{\partial p_i^m} + \left[\left(\sum_{j=1}^{N_m}\varepsilon_j^m\sigma_j^m\right) - \varepsilon_i^m\right]\frac{y_i^m}{Y_m}
\end{aligned}$$

This lets us get the derivative of D_m

$$\frac{\partial D_m}{\partial p_i^m} = \left(1 - \frac{\varepsilon_i^m}{\sum_j \varepsilon_j^m\sigma_j^m}\right)\sigma_i^m\frac{D_m}{p_i^m}. \qquad \text{(C.37)}$$

Now we turn to the substitution effects

$$\underbrace{-\frac{p_i^m\frac{p_i^m}{P_m}}{\Upsilon_i''\frac{y_i^m}{Y_m}}\frac{\partial D_m}{\partial p_i^m}}_{\text{Substitution Effects}} = -\frac{\Upsilon_i'}{\Upsilon_i''\frac{y_i^m}{Y_m}}\left(1-\frac{\varepsilon_i^m}{\sum_j\varepsilon_j^m\sigma_j^m}\right)\sigma_i^m\frac{\frac{p_i^m}{P_m}D_m}{\Upsilon_i'} = \varepsilon_i^m\left(1-\frac{\varepsilon_i^m}{\sum_j\varepsilon_j^m\sigma_j^m}\right)\sigma_i^m$$

Finally, we get the elasticity

$$\eta_i^m = \gamma\sigma_i^m + \varepsilon_i^m\left(1-\sigma_i^m\right) + \varepsilon_i^m\left(1-\frac{\varepsilon_i^m}{\sum_j\varepsilon_j^m\sigma_j^m}\right)\sigma_i^m. \qquad \text{(C.38)}$$

We can also write down as

$$\eta_i^m = \gamma\sigma_i^m + \varepsilon_i^m\left(1 - \frac{\varepsilon_i^m\sigma_i^m}{\sum_j \varepsilon_j^m\sigma_j^m}\right) = \gamma\sigma_i^m + \varepsilon_i^m\frac{E_\sigma\left[\varepsilon_j^m | j \neq i\right]}{E_\sigma\left[\varepsilon_j^m\right]}\left(1-\sigma_i^m\right).$$

□

C.3. Allen-Uzawa elasticity of substitution

We now tackle derive the Allen-Uzawa elasticity of substitution between products i and j in market m. We take as given prices and market output. The elasticity when $i \neq j$ is

$$\varphi_{ij}^m = \frac{1}{\sigma_j^m}\frac{\partial \log y_i^m}{\partial \log p_j^m} = \frac{1}{\sigma_j^m}\frac{p_j^m}{y_i^m}\frac{\partial}{\partial p_j^m}\left[Y_m\left(\Upsilon_i'\right)^{-1}\left(\left(\sum_h \Upsilon_h'\left(\frac{y_h^m}{Y_m}\right)\frac{y_h^m}{Y_m}\right)\frac{p_i^m}{P_m}\right)\right] \tag{C.39}$$

This allows us to break down the elasticity into two terms

$$\varphi_{ij}^m = \frac{p_j^m p_i^m}{P_m\sigma_j^m}\left[\underbrace{-\frac{D_m\frac{1}{P_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\partial P_m}{\partial p_j^m}}_{\text{Cost Effect}} + \underbrace{\frac{1}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\partial D_m}{\partial p_j^m}}_{\text{Substitution Effects}}\right]; \tag{C.40}$$

where we define D_m as in (C.6). From Lemma 1 we know that $\frac{\partial P_m}{\partial p_j^m} = \frac{y_j^m}{Y_m}$ and from the proof of Proposition 1 we know that

$$\frac{\partial D_m}{\partial p_j^m} = \left(1 - \frac{\varepsilon_j^m}{\sum_h \varepsilon_h^m\sigma_h^m}\right)\sigma_j^m\frac{D_m}{p_j^m}. \tag{C.41}$$

Replacing,

$$\varphi_{ij}^m = \frac{p_j^m p_i^m}{P_m\sigma_j^m}\left[-\frac{D_m\frac{1}{P_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{y_j^m}{Y_m} + \frac{1}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\left(1 - \frac{\varepsilon_j^m}{\sum_h \varepsilon_h^m\sigma_h^m}\right)\sigma_j^m\frac{D_m}{p_j^m}\right]; \tag{C.42}$$

$$\varphi_{ij}^m = -\frac{D_m\frac{p_i^m}{P_m}}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\varepsilon_j^m}{\sum_h \varepsilon_h^m\sigma_h^m}; \tag{C.43}$$

$$\varphi_{ij}^m = -\frac{\Upsilon_i'\left(\frac{y_i^m}{Y_m}\right)}{\Upsilon_i''\left(\frac{y_i^m}{Y_m}\right)\frac{y_i^m}{Y_m}}\frac{\varepsilon_j^m}{\sum_h \varepsilon_h^m\sigma_h^m}; \tag{C.44}$$

$$\varphi_{ij}^m = \frac{\varepsilon_i^m \varepsilon_j^m}{\sum_h \varepsilon_h^m \sigma_h^m} \,. \tag{C.45}$$

The final step uses $D_m \frac{p_i^m}{P_m} = \Upsilon_i' \left(\frac{y_i^m}{Y_m} \right)$ from the proof of Lemma 1, equation (C.14).

Finally, the own-elasticity is

$$\varphi_{ii}^m = \frac{1}{\sigma_i^m} \frac{\partial \log y_i^m}{\partial \log p_i^m} = \frac{1}{\sigma_i^m} \frac{p_i^m}{y_i^m} \frac{\partial}{\partial p_i^m} \left[Y_m \left(\Upsilon_i' \right)^{-1} \left(D_m \frac{p_i^m}{P_m} \right) \right] \tag{C.46}$$

This allows us to break down the elasticity into two terms

$$\varphi_{ij}^m = \frac{Y_m}{\sigma_i^m} \frac{p_i^m}{y_i^m} \left[\frac{D_m}{\Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right)} \frac{\partial \, p_i^m / P_m}{\partial \, p_i^m} + \frac{\frac{p_i^m}{P_m}}{\Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right)} \frac{\partial D_m}{\partial \, p_i^m} \right] ; \tag{C.47}$$

$$\varphi_{ij}^m = \frac{p_i^m}{\sigma_i^m \Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right) \frac{y_i^m}{Y_m}} \left[D_m \left(\frac{1}{P_m} - \frac{p_i^m}{P_m^2} \frac{\partial P_m}{\partial \, p_i^m} \right) + \frac{p_i^m}{P_m} \frac{\partial D_m}{\partial \, p_i^m} \right] ; \tag{C.48}$$

$$\varphi_{ij}^m = \frac{p_i^m}{\sigma_i^m \Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right) \frac{y_i^m}{Y_m}} \left[D_m \left(\frac{1}{P_m} - \frac{p_i^m}{P_m^2} \frac{y_i^m}{Y_m} \right) + \frac{p_i^m}{P_m} \left(1 - \frac{\varepsilon_i^m}{\sum_h \varepsilon_h^m \sigma_h^m} \right) \sigma_i^m \frac{D_m}{p_i^m} \right] ; \tag{C.49}$$

$$\varphi_{ij}^m = \frac{D_m \frac{p_i^m}{P_m}}{\sigma_i^m \Upsilon_i'' \left(\frac{y_i^m}{Y_m} \right) \frac{y_i^m}{Y_m}} \left[1 - \sigma_i^m + \left(1 - \frac{\varepsilon_i^m}{\sum_h \varepsilon_h^m \sigma_h^m} \right) \sigma_i^m \right] ; \tag{C.50}$$

$$\varphi_{ij}^m = -\frac{\varepsilon_i^m}{\sigma_i^m} \left[1 - \frac{\sigma_i^m \varepsilon_i^m}{\sum_h \varepsilon_h^m \sigma_h^m} \right] . \tag{C.51}$$

$$\tag{C.52}$$

We verify explicitly that

$$\sum_{j=1}^{N_m} \sigma_j^m \varphi_{ij} = -\varepsilon_i^m \left[1 - \frac{\sigma_i^m \varepsilon_i^m}{\sum_h \varepsilon_h^m \sigma_h^m} \right] + \sum_{j \neq i} \sigma_j^m \frac{\varepsilon_i^m \varepsilon_j^m}{\sum_h \varepsilon_h^m \sigma_h^m} \tag{C.53}$$

$$= \varepsilon_i^m \left(1 - \frac{\sigma_i^m \varepsilon_i^m}{\sum_h \varepsilon_h^m \sigma_h^m} + \sum_{j \neq i} \frac{\sigma_j^m \varepsilon_j^m}{\sum_h \varepsilon_h^m \sigma_h^m} \right) = 0 \,. \tag{C.54}$$

C.4. Market level elasticities and markups

PROPOSITION A2. ***(Market Elasticities and Markups)*** *Market elasticities satisfy*

$$\textit{Cournot} \qquad \frac{1}{\eta_m} = HHI_m \frac{1}{\gamma} + (1 - HHI_m) \, E_\sigma \left[\frac{1}{\varepsilon_i^m} \right] - 2 Cov_\sigma \left(\sigma_i^m, \frac{1}{\varepsilon_i^m} \right) ; \tag{C.55}$$

Bertrand $$\frac{1}{\eta_m} = E_\sigma\left[\frac{1}{\varepsilon_i^m}\right] + \frac{HHI_m}{E_\sigma\left[\varepsilon_i^m\right]} - \gamma\left(HHI_m E_\sigma\left[\left(\frac{1}{\varepsilon_i^m}\right)^2\right] + Cov_\sigma\left(\sigma_i^m, \left(\frac{1}{\varepsilon_i^m}\right)^2\right)\right) + \Omega \tag{C.56}$$

where $HHI_m = \sum_i \left(\sigma_i^m\right)^2$ *is the Herfindahl-Hirschman index,* $E_\sigma\left[x_j\right] = \sum_{j=1}^{N_m} x_j \sigma_j^m$ *is the expectation with respect to sales shares, and* $Cov_\sigma\left(x_j, y_j\right) = \sum_{j=1}^{N_m}\left(x_j\right)\left(y_j - E_\sigma\left[y_j\right]\right)\sigma_j^m$ *is the covariance with respect to sales in the market. Finally,* $\Omega \equiv \sum_i^{N_m} \frac{\sigma_i^m}{\varepsilon_i^m}\left[\sum_{k=2}^{\infty}\left(\sigma_i^m\left(\frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \frac{\gamma}{\varepsilon_i^m}\right)\right)^k\right]$ *contains higher order moments.*

The average markup satisfy

Cournot $$\frac{1}{\mu_m} = \underbrace{\left(1 - \frac{1}{\gamma}\right)}_{\text{Monopoly Markup}} + \underbrace{\left(\frac{1}{\gamma} - E_\sigma\left[\frac{1}{\varepsilon_i^m}\right]\right)(1 - HHI_m)}_{\text{Concentration}} + \underbrace{2Cov_\sigma\left(\sigma_i^m, \frac{1}{\varepsilon_i^m}\right)}_{\text{Distribution}}; \tag{C.57}$$

Bertrand $$\frac{1}{\mu_m} = \underbrace{\left(1 - \frac{1}{\gamma}\right)}_{\text{Monopoly Markup}} + \underbrace{\left(\frac{1}{\gamma} - E_\sigma\left[\frac{1}{\varepsilon_i^m}\right]\right) + \frac{HHI_m}{\gamma}\left(E_\sigma\left[\left(\frac{\gamma}{\varepsilon_i^m}\right)^2\right] - \frac{\gamma}{E_\sigma\left[\varepsilon_i^m\right]}\right)}_{\text{Concentration}} + \underbrace{Cov_\sigma\left(\sigma_i^m, \left(\frac{1}{\varepsilon_i^m}\right)^2\right)}_{\text{Distribution}} - \Omega\,. \tag{C.58}$$

Proof. We start with Cournot competition. The market elasticity is

$$\begin{aligned}
\frac{1}{\eta_m} &= \sum_{i=1}^{N_m}\left(\frac{1}{\gamma}\sigma_i^m + \left[\frac{1}{\varepsilon_i^m}\left(1 - \sigma_i^m\right) + \sigma_i^m \sum_{j\neq i}\frac{1}{\varepsilon_j^m}\cdot\frac{\sigma_j^m}{1-\sigma_i^m}\right]\left(1 - \sigma_i^m\right)\right)\sigma_i^m \qquad \text{(C.59)}\\
&= \frac{1}{\gamma}\text{HHI} + \sum_{i=1}^{N_m}\left(\frac{1}{\varepsilon_i^m}\left(1 - 2\sigma_i^m\right) + \sigma_i^m\left(\sum_{j=1}^{N_m}\frac{1}{\varepsilon_j^m}\cdot\sigma_j^m\right)\right)\sigma_i^m\\
&= \frac{1}{\gamma}\text{HHI} + \sum_{i=1}^{N_m}\frac{1}{\varepsilon_i^m}\left(1 - \sigma_i^m\right)\sigma_i^m - \sum_{i=1}^{N_m}\sigma_i^m\left(\frac{1}{\varepsilon_i^m} - E_\sigma\left[\frac{1}{\varepsilon_i^m}\right]\right)\sigma_i^m\\
&= \frac{1}{\gamma}\text{HHI} + \sum_{i=1}^{N_m}\left[\frac{1}{\varepsilon_i^m}\left(1 - \sigma_i^m\right) + \sigma_i^m\left(\frac{1}{\varepsilon_i^m} - E_\sigma\left[\frac{1}{\varepsilon_i^m}\right]\right)\right]\sigma_i^m - 2\text{Cov}_\sigma\left(\sigma_i^m, \frac{1}{\varepsilon_i^m}\right)
\end{aligned}$$

$$= \text{HHI}\frac{1}{\gamma} + (1 - \text{HHI})\, E_\sigma\left[\frac{1}{\varepsilon_i^m}\right] - 2\text{Cov}_\sigma\left(\sigma_i^m, \frac{1}{\varepsilon_i^m}\right)$$

The markup follows from grouping terms in the expression $1/\mu_m = 1 - 1/\eta_m$.

For Bertrand competition we first express firm i's effective elasticity as

$$\eta_i^m = \gamma\sigma_i^m + \varepsilon_i^m\left(1 - \sigma_i^m\right) + \varepsilon_i^m\left(1 - \frac{\varepsilon_i^m}{\sum_j \varepsilon_j^m \sigma_j^m}\right)\sigma_i^m = \varepsilon_i^m\left(1 - \sigma_i^m\left(\frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \frac{\gamma}{\varepsilon_i^m}\right)\right) \tag{C.60}$$

Average elasticity of the market is

$$\begin{aligned}
\frac{1}{\eta_m} &= \sum_i^{N_m} \frac{\sigma_i^m}{\varepsilon_i^m} \frac{1}{1 - \sigma_i^m\left(\frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \frac{\gamma}{\varepsilon_i^m}\right)} \qquad \text{(C.61)} \\
&= \sum_i^{N_m} \frac{\sigma_i^m}{\varepsilon_i^m} \sum_{k=0}^{\infty}\left(\sigma_i^m\left(\frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \frac{\gamma}{\varepsilon_i^m}\right)\right)^k \\
&= \sum_i^{N_m} \frac{\sigma_i^m}{\varepsilon_i^m}\left[1 + \left(\frac{\sigma_i^m \varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \gamma\frac{\sigma_i^m}{\varepsilon_i^m}\right) + \sum_{k=2}^{\infty}\left(\sigma_i^m\left(\frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \frac{\gamma}{\varepsilon_i^m}\right)\right)^k\right] \\
&= E_\sigma\left[\frac{1}{\varepsilon_i^m}\right] + \frac{\text{HHI}_m}{E_\sigma\left[\varepsilon_i^m\right]} - \gamma E_\sigma\left[\sigma_i^m\left(\frac{1}{\varepsilon_i^m}\right)^2\right] + \Omega \\
&= E_\sigma\left[\frac{1}{\varepsilon_i^m}\right] + \frac{\text{HHI}_m}{E_\sigma\left[\varepsilon_i^m\right]} - \gamma\left(\text{HHI}_m E_\sigma\left[\left(\frac{1}{\varepsilon_i^m}\right)^2\right] + \text{Cov}_\sigma\left(\sigma_i^m, \left(\frac{1}{\varepsilon_i^m}\right)^2\right)\right) + \Omega
\end{aligned}$$

where $\Omega \equiv \sum_i^{N_m} \frac{\sigma_i^m}{\varepsilon_i^m}\left[\sum_{k=2}^{\infty}\left(\sigma_i^m\left(\frac{\varepsilon_i^m}{E_\sigma\left[\varepsilon_j^m\right]} - \frac{\gamma}{\varepsilon_i^m}\right)\right)^k\right]$ contains higher order moments and $\text{Cov}_\sigma\left(\sigma_i^m, X_i^m\right) = E_\sigma\left[\sigma_i^m X_i^m\right] - E_\sigma\left[\sigma_i^m\right] E_\sigma\left[X_i^m\right] = E_\sigma\left[\sigma_i^m X_i^m\right] - \text{HHI}_m E_\sigma\left[X_i^m\right]$.

The third-order higher order terms are

$$\begin{aligned}
\Omega_2 = \gamma^2 &\left(\text{HK}_m(3)^3 E_\sigma\left[\left(\frac{1}{\varepsilon_i^m}\right)^3\right] + \text{Cov}_\sigma\left(\left(\sigma_i^m\right)^2, \left(\frac{1}{\varepsilon_i^m}\right)^3\right)\right) \qquad \text{(C.62)} \\
&- 2\frac{\gamma}{E_\sigma\left[\varepsilon_i^m\right]}\left(\text{HK}_m(3)^3 E_\sigma\left[\left(\frac{1}{\varepsilon_i^m}\right)\right] + \text{Cov}_\sigma\left(\left(\sigma_i^m\right)^2, \left(\frac{1}{\varepsilon_i^m}\right)\right)\right)
\end{aligned}$$

$$+ \frac{1}{\left(E_\sigma\left[\varepsilon_i^m\right]\right)^2}\left(\mathrm{HK}_m(3)^3 E_\sigma\left[\varepsilon_i^m\right] + \mathrm{Cov}_\sigma\left(\left(\sigma_i^m\right)^2, \varepsilon_i^m\right)\right) ;$$

where $\mathrm{HK}_m(k) \equiv \left(\sum_i^{N_m}\left(\sigma_i^m\right)^k\right)^{\frac{1}{k}}$ is the Hannah and Kay index of order k. □

COROLLARY A1. ***(Markups with Constant Elasticity)*** *If all firms have the same common own-elasticity, $\varepsilon_i^m = \varepsilon_m$, markups are*

$$\textbf{\textit{Cournot}} \qquad \frac{1}{\mu_m} = \left(1 - \frac{1}{\gamma}\right) + \left(\frac{1}{\gamma} - \frac{1}{\varepsilon_m}\right)(1 - HHI_m) ; \tag{C.63}$$

$$\textbf{\textit{Bertrand}} \qquad \frac{1}{\mu_m} = \frac{\varepsilon - 1}{\varepsilon} - \frac{1}{\varepsilon}\sum_{k=2}^{\infty}\left(1 - \frac{\gamma}{\varepsilon}\right)^{k-1}\left(HK_m(k)\right)^k ; \tag{C.64}$$

where $HHI_m = \sum_i\left(\sigma_i^m\right)^2$ is the Herfindahl-Hirschman index and $HK_m(k) \equiv \left(\sum_i^{N_m}\left(\sigma_i^m\right)^k\right)^{\frac{1}{k}}$ is the Hannah and Kay index of order k. This corresponds to Grassi (2017, Proposition 4).

PROOF. The proof for Cournot competition is immediate as the covariance term is zero with common elasticities. For Bertrand we express the markup in terms of he Hannah and Kay (1977) concentration index, $\mathrm{HK}_m(k) \equiv \left(\sum_i^{N_m}\left(\sigma_i^m\right)^k\right)^{\frac{1}{k}}$,

$$\begin{aligned}
\frac{1}{\mu_m} &= 1 - \frac{1}{\varepsilon}\sum_i^{N_m}\frac{\sigma_i^m}{1 - \sigma_i^m\left(1 - \frac{\gamma}{\varepsilon}\right)} = 1 - \frac{1}{\varepsilon}\sum_i^{N_m}\sigma_i^m\sum_{k=0}^{\infty}\left(\sigma_i^m\left(1 - \frac{\gamma}{\varepsilon}\right)\right)^k \\
&= 1 - \frac{1}{\varepsilon}\sum_{k=0}^{\infty}\left(1 - \frac{\gamma}{\varepsilon}\right)^k\sum_i^{N_m}\left(\sigma_i^m\right)^{k+1} = \frac{\varepsilon - 1}{\varepsilon} - \frac{1}{\varepsilon}\sum_{k=2}^{\infty}\left(1 - \frac{\gamma}{\varepsilon}\right)^{k-1}\left(\sum_i^{N_m}\left(\sigma_i^m\right)^k\right) \\
&= \frac{\varepsilon - 1}{\varepsilon} - \frac{1}{\varepsilon}\sum_{k=2}^{\infty}\left(1 - \frac{\gamma}{\varepsilon}\right)^{k-1}\left(\mathrm{HK}_m(k)\right)^k .
\end{aligned}$$

□

Appendix D. Distance to the frontier

We start by making some required definitions. We let $N \equiv 1+M+\sum_{m=1}^{M} N_m$ be the total number of producers in all M markets inclusive of the market aggregators and the final good aggregator. Therefore, there are N total firms in the economy and $F \equiv 1+\sum_{m=1}^{M} N_m$ factors. The following definitions take an i from either N or $N+F$. The nominal GDP in this economy is $PY = \sum_{i=1}^{M} P_m Y_m$ and the share of each market in total expenditures is $\sigma_m = \frac{P_m Y_m}{PY}$.

Input-Output Revenue Matrix: The $N+F \times N+F$ matrix Ω encodes the input-output structure of the economy with a typical element $\Omega_{ij} = \frac{p_j x_j^i}{p_i y_i}$ being the input expenditure on producer j's goods by producer i relative to the total sales of producer i. This is a measure of the importance of other producers in the production of revenue by producer i. We denote by $\Omega^{(i)}$ the i^{th} row of Ω corresponding to the linkages of producer i with its suppliers. We denote the Leontief inverse by $\Psi = (I - \Omega)^{-1}$ and its j^{th} column by $\Psi_{(j)}$.

Input-Output Cost Matrix: The $N+2 \times N+2$ matrix $\widetilde{\Omega}$ encodes the input-output structure captured by the elasticities of marginal costs to prices of inputs. Its typical element is $\widetilde{\Omega}_{ij} = \frac{p_j x_j^i}{\sum_{h=1}^{N+2} p_h x_h^i}$. This is a measure of the importance of j in i's costs. The rows of $\widetilde{\Omega}$ sum to 1. The Leontief inverse is $\widetilde{\Psi}$ and its k^{th} column by $\widetilde{\Psi}_{(k)}$.

Domar weights: The Domar weight $\omega_i \equiv \frac{p_i y_i}{PY}$ measures producer i's sales relative to nominal GDP. We refer to the Domar weights of inputs as $\Lambda_f = \omega_{N+f}$.

The organization of the economy requires the first columns of Ω and $\widetilde{\Omega}$ be zero as the final good producer's output is not an input for any producer. The first row is only populated in columns 2 to $1+M$ for the aggregators of the M markets. The next M rows have a block-diagonal structure with non-zero entries only for the corresponding N_m producers. The next rows up-until the last F have only entries for labor and the fixed factor of the corresponding producer. The last F rows are all zeros.

In order to get the revenue and cost shares for inputs proceed as follows: The production function estimated for every product-market in Appendix A.3 is a Cobb-Douglas on capital, materials, and labor with potentially decreasing-returns-to-scale, $\beta_m \equiv \beta_K^m + \beta_M^m + \beta_L^m$ where β_f^m is the cost-share of input $f \in \{K, M, L\}$. We simplify the problem by aggregating all physical inputs into a single composite factor that we call L with output-elasticity β_m in the production of goods in market m. We add a stand-in managerial or entrepreneurial input that is in fixed supply E_n^m for each producer (n) in each market (m) with output-elasticity $1 - \beta_m$. These elasticities also give their cost shares for $\widetilde{\Omega}$ and their output shares for Ω as β_m/μ_n^m and

$(1-\beta_m)/\mu_n^m$. The elasticities of substitution between these factors are 1, and the own-elasticities are $-\beta_m/(1-\beta_m)$ and $-(1-\beta_m)/\beta_m$, for L and E_n^m, repsectively.

Now, we approximate the distance to the frontier up to a second order following Baqaee and Farhi (2020, Prop. 5),

$$\begin{aligned}\mathcal{L} \equiv \log\frac{Y^\star}{Y} \approx & \frac{1}{2}\sum_{i=1}^{N}\omega_i\Phi_{\Omega^{(i)}}\left(\sum_{j=1}^{N}\Psi_{(j)}\log\mu_j\,,\,\sum_{j=1}^{N}\Psi_{(j)}\log\mu_j\right) \\ & + \frac{1}{2}\sum_{i=1}^{N}\omega_i\mathrm{Var}_{\Omega^{(i)}}\left(\sum_{j=1}^{N}\Psi_{(j)}\log\mu_j\right) \\ & + \frac{1}{2}\sum_{i=1}^{N}\omega_i\Phi_{\Omega^{(i)}}\left(\sum_{f=1}^{F}\Psi_{(f)}\Delta\log\Lambda_f\,,\,\sum_{h=1}^{N}\Psi_{(h)}\log\mu_h\right) \\ & + \frac{1}{2}\sum_{i=1}^{N}\omega_i\mathrm{Cov}_{\Omega^{(i)}}\left(\sum_{f=1}^{F}\Psi_{(f)}\Delta\log\Lambda_f\,,\,\sum_{h=1}^{N}\Psi_{(h)}\log\mu_h\right)\end{aligned} \tag{D.1}$$

where the sum is over all producers in the economy and

$$\Delta\log\Lambda_f = \sum_{k=1}^{N}\frac{d\log\Lambda_f}{d\log\mu_k}\log\mu_k \tag{D.2}$$

where $\frac{d\log\Lambda_f}{d\log\mu_k}$ is the solution to

$$\frac{d\log\Lambda_f}{d\log\mu_k} = \sum_{j=1}^{N}\frac{\omega_j}{\mu_j}\Phi_{\widetilde{\Omega}^{(j)}}\left(\widetilde{\Psi}_{(k)}-\sum_{g=1}^{F}\widetilde{\Psi}_{(g)}\frac{d\log\Lambda_g}{d\log\mu_k}\,,\,\frac{\Psi_{(f)}}{\Lambda_f}\right)-\omega_k\frac{\Psi_{kf}}{\Lambda_f}\,. \tag{D.3}$$

The covariance operator weighted by the i^{th} row of matrix A between two h^{th} and k^{th} columns of matrices B and C, all of size $N+F$, is defined as

$$\mathrm{Cov}_{A^{(i)}}\left(B_{(h)}\,,\,C_{(k)}\right) = \sum_{j=1}^{N+F}A_{ij}B_{jh}C_{jk}-\left(\sum_{j=1}^{N+F}A_{ij}B_{jh}\right)\left(\sum_{j=1}^{N+F}A_{ij}C_{jk}\right)\,. \tag{D.4}$$

The operator $\Phi_a(b,c)$ is defined for the i^{th} row of matrix A and the h^{th} and k^{th} columns of matrices B and C, all of size $N+F$, as

$$\Phi_{A^{(i)}}\left(B_{(h)}\,,\,C_{(k)}\right) = -\sum_{j=1}^{N+F}\sum_{g=1}^{N+F}A_{ij}\left(\mathbb{1}_{j=g}+A_{ig}\left(\varphi^i_{jg}-1\right)\right)B_{jh}C_{gk}\,. \tag{D.5}$$

where φ is the elasticity of substitution defined in (15).

Appendix E. Functional forms for Kimball's aggregator

Several functional forms for the Kimball aggregator have been employed in the literature. The main functional forms are those introduced by Dotsey and King (2005) and Klenow and Willis (2016). We show how other functional forms are equivalent to the Dotsey and King (2005) aggregator after a suitable change of variables. Tables E.1 and E.2 summarize these results.

We also propose a new functional form that implies a constant super elasticity of demand. We show how this has direct implications for the relationship between markups and market shares in equilibrium. This relationship turns out to be one-to-one, so that knowing either markups or market shares implies the other. The data rejects this relationship.

TABLE E.1. Kimball Aggregators I

Function	CES	D&K(2005)	K&W(2016)
$\Upsilon(x)$	$x^{\frac{\epsilon-1}{\epsilon}}$	$\frac{1}{(1+\theta)\rho}[(1+\theta)x-\theta]^{\rho}-\left[1+\frac{1}{(1+\theta)\rho}\right]$	$1+(\epsilon-1)exp\left(\frac{1}{\theta}\right)\theta^{\frac{\epsilon}{\theta}-1}\left[\Gamma\left(\frac{\epsilon}{\theta},\frac{1}{\theta}\right)-\Gamma\left(\frac{\epsilon}{\theta},\frac{x^{\frac{\theta}{\epsilon}}}{\theta}\right)\right]$
$\Upsilon'(x)$	$\frac{\epsilon-1}{\epsilon}x^{-\frac{1}{\epsilon}}$	$\left[(1+\theta)x-\theta\right]^{\rho-1}$	$\frac{\epsilon-1}{\epsilon}exp\left(\frac{1-x^{\frac{\theta}{\epsilon}}}{\theta}\right)$
$\Upsilon''(x)$	$\frac{1-\epsilon}{\epsilon^2}x^{-\frac{(1+\epsilon)}{\epsilon}}$	$(\rho-1)(1+\theta)\left[(1+\theta)x-\theta\right]^{\rho-2}$	$\frac{1-\epsilon}{\epsilon^2}exp\left(\frac{1-x^{\frac{\theta}{\epsilon}}}{\theta}\right)x^{\frac{\theta-\epsilon}{\epsilon}}$
$[\Upsilon']^{-1}(z)$	$\left(\frac{\epsilon}{\epsilon-1}z\right)^{-\epsilon}$	$\frac{z^{\frac{1}{\rho-1}}+\theta}{(1+\theta)}$	$\left[1-\theta\left(log\left(\frac{\epsilon}{\epsilon-1}\right)+log(z)\right)\right]^{\frac{\epsilon}{\theta}}$
$\left([\Upsilon']^{-1}\right)'(x)$	$\frac{-\epsilon^2}{\epsilon-1}x^{\frac{1+\epsilon}{\epsilon}}$	$\frac{1}{(\rho-1)(1+\theta)}[(1+\theta)x-\theta]^{2-\rho}$	$\frac{\epsilon^2}{(1-\epsilon)exp\left(\frac{1-x^{\frac{\theta}{\epsilon}}}{\theta}\right)x^{\frac{\theta-\epsilon}{\epsilon}}}$
$\varepsilon(x)=-\frac{1}{x}\frac{\Upsilon'(x)}{\Upsilon''(x)}$	ϵ	$\frac{1}{1-\rho}\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$	$\epsilon\cdot x^{-\frac{\theta}{\epsilon}}$
Change of parameters		$\epsilon=\frac{1}{1-\rho}$	
(Simplified) Elasticity of demand	ϵ	$\epsilon\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$	$\epsilon\cdot x^{-\frac{\theta}{\epsilon}}$
Super-elasticity of demand (ξ)	0	$-\frac{\epsilon\theta}{1+\theta}x^{-1}$	$\theta\cdot x^{-\frac{\theta}{\epsilon}}$

E.1. Equivalence results

In this Section we highlight 4 additional Kimball aggregators used in the literature which are equivalent to the Dotsey and King (2005) aggregator.

Barde (2008) Aggregator. The functional form for the aggregator is:

$$\Upsilon(x) = \frac{(\tilde{\theta}x - (\tilde{\theta} - 1))^{\rho}}{\tilde{\theta}} - \frac{(1 - \tilde{\theta})^{\rho}}{\tilde{\theta}} \tag{E.1}$$

where $\rho = \frac{\epsilon - 1}{\epsilon}$, ϵ is the elasticity of substitution between varieties, $\tilde{\theta}$ controls the curvature of the function. If $\tilde{\theta} = 1$, this specification becomes the CES aggregator.

The first derivative is:

$$\Upsilon'(x) = \rho(\theta x - (\theta - 1))^{\rho - 1}$$

The second derivative is:

$$\Upsilon''(x) = \theta\rho(\rho - 1)(\theta x - (\theta - 1))^{\rho - 2}.$$

The elasticity of demand is:

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = -\frac{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)}{\frac{y_i^s}{Y_s}\Upsilon''\left(\frac{y_i^s}{Y_s}\right)} = \frac{1}{1 - \rho}\left[1 - \frac{\tilde{\theta} - 1}{\tilde{\theta}}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Changing variables to $\epsilon = \frac{1}{1-\rho}$ and $\theta = \tilde{\theta} - 1$ we get the same expression as in Dotsey and King (2005):

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = \epsilon\left[1 - \frac{\theta}{1 + \theta}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Levin et al. (2008) and Lindé and Trabandt (2018) Aggregator. The functional form for the Kimball Aggregator is:

$$\Upsilon(x) = \frac{\phi}{1 + \theta}\left[(1 + \theta)x - \theta\right]^{\frac{1}{\phi}} - \left[\frac{\phi}{1 + \theta} - 1\right] \tag{E.2}$$

where $\phi = \frac{\epsilon(1+\theta)}{\epsilon(1+\theta)-1}$, ϵ is the elasticity of substitution between varieties, and θ controls the curvature of the function. If $\theta = 0$ this specification becomes the CES aggregator.

The first derivative is:

$$\Upsilon'(x) = \left[(1 + \theta)x - \theta\right]^{\frac{1-\phi}{\phi}}$$

The second derivative is:

$$\Upsilon''(x) = \left(\frac{1 - \phi}{\phi}\right)(1 + \theta)\left[(1 + \theta)x - \theta\right]^{\frac{1}{\phi} - 2}$$

The elasticity of demand is:

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = -\frac{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)}{\frac{y_i^s}{Y_s}\Upsilon''\left(\frac{y_i^s}{Y_s}\right)} = \frac{\phi}{\phi - 1}\left[1 - \frac{\theta}{1 + \theta}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Changing variables to $\epsilon = \frac{\phi}{\phi - 1}$ we get the same expression as in Dotsey and King (2005):

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = \epsilon\left[1 - \frac{\theta}{1+\theta}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Darracq Pariès and Loublier (2010) Aggregator. The functional form for the Kimball Aggregator is:

$$\Upsilon(x) = \frac{\tilde{\epsilon}}{\tilde{\epsilon}(1+\theta) - 1}\left[(1+\theta)x - \theta\right]^{\frac{\tilde{\epsilon}(1+\theta)-1}{\tilde{\epsilon}(1+\theta)}} - \left[\frac{\tilde{\epsilon}}{\tilde{\epsilon}(1+\theta) - 1} - 1\right] \tag{E.3}$$

where $\tilde{\epsilon}$ is the elasticity of substitution between varieties, θ controls the curvature of the function. If $\theta = 0$ this specification becomes the CES aggregator.
The first derivative is:

$$\Upsilon'(x) = \left[(1+\theta)x - \theta\right]^{-\frac{1}{\tilde{\epsilon}(1+\theta)}}$$

The second derivative is:

$$\Upsilon''(x) = -\frac{1}{\tilde{\epsilon}}\left[(1+\theta)x - \theta\right]^{-\left(\frac{1+\tilde{\epsilon}(1+\theta)}{\tilde{\epsilon}(1+\theta)}\right)}$$

The elasticity of demand is:

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = -\frac{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)}{\frac{y_i^s}{Y_s}\Upsilon''\left(\frac{y_i^s}{Y_s}\right)} = \tilde{\epsilon}\cdot\left[(1+\theta) - \theta\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Changing variables to $\epsilon = \tilde{\epsilon}\,(1+\theta)$ we get the same expression as in Dotsey and King (2005):

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = \epsilon\left[1 - \frac{\theta}{1+\theta}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Kurozumi and Zandweghe (2020) Aggregator. The functional form for the Kimball Aggregator is:

$$\Upsilon(x) = \frac{\phi}{(1+\theta)(\phi - 1)}\left[(1+\theta)x - \theta\right]^{\frac{\phi-1}{\phi}} - \left[\frac{\phi}{(1+\theta)(\phi - 1)} - 1\right] \tag{E.4}$$

where $\phi = \epsilon(1+\theta)$, ϵ is the elasticity of substitution between varieties, $-(\theta\epsilon)$ controls the curvature of the function. If $\theta = 0$ this specification becomes the CES aggregator.
The first derivative is:

$$\Upsilon'(x) = \left[(1+\theta)x - \theta\right]^{-\frac{1}{\phi}}$$

The second derivative is:

$$\Upsilon''(x) = \left(-\frac{1}{\phi}\right)(1+\theta)\left[(1+\theta)x-\theta\right]^{-\frac{(1+\phi)}{\phi}}$$

The elasticity of demand is:

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = -\frac{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)}{\frac{y_i^s}{Y_s}\Upsilon''\left(\frac{y_i^s}{Y_s}\right)} = \phi\left[1-\frac{\theta}{(1+\theta)}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

Changing variables to $\epsilon = \phi$ we get the same expression as in Dotsey and King (2005):

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = \epsilon\left[1-\frac{\theta}{1+\theta}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right]$$

TABLE E.2. Kimball Aggregators II

Function	Barde (2008)	Levin et al (2008)	DP&L(2010)	K&VZ(2020)
$\Upsilon(x)$	$\frac{(\theta x-(\theta-1))^{\rho}}{\theta}-\frac{(1-\theta)^{\rho}}{\theta}$	$\frac{\phi}{1+\theta}\left[(1+\theta)x-\theta\right]^{\frac{1}{\phi}}-\left[\frac{\phi}{1+\theta}-1\right]$	$\frac{\epsilon}{\epsilon(1+\theta)-1}\left[(1+\theta)x-\theta\right]^{\frac{\epsilon(1+\theta)-1}{\epsilon(1+\theta)}}-\left[\frac{\epsilon}{\epsilon(1+\theta)-1}-1\right]$	$\frac{\phi}{(1+\theta)(\phi-1)}\left[(1+\theta)x-\theta\right]^{\frac{\phi-1}{\phi}}-\left[\frac{\phi}{(1+\theta)(\phi-1)}-1\right]$
$\Upsilon'(x)$	$\rho(\theta x-(\theta-1))^{\rho-1}$	$\left[(1+\theta)x-\theta\right]^{\frac{1-\phi}{\phi}}$	$\left[(1+\theta)x-\theta\right]^{-\frac{1}{\epsilon(1+\theta)}}$	$\left[(1+\theta)x-\theta\right]^{-\frac{1}{\phi}}$
$\Upsilon''(x)$	$\theta\rho(\rho-1)(\theta x-(\theta-1))^{\rho-2}$	$\left(\frac{1-\phi}{\phi}\right)(1+\theta)\left[(1+\theta)x-\theta\right]^{\frac{1}{\phi}-2}$	$-\left(\frac{1}{\epsilon}\right)\left[(1+\theta)x-\theta\right]^{-\left(\frac{1+\epsilon(1+\theta)}{\epsilon(1+\theta)}\right)}$	$\left(-\frac{1}{\phi}\right)(1+\theta)\left[(1+\theta)x-\theta\right]^{-\frac{(1+\phi)}{\phi}}$
$[\Upsilon']^{-1}(z)$	$\left[\left(\frac{z}{\rho}\right)^{\frac{1}{\rho-1}}+(\theta-1)\right]\frac{1}{\theta}$	$\frac{\phi}{z^{\frac{1-\phi}{\phi}}+\theta}$	$\frac{z^{-\epsilon(1+\theta)}+\theta}{(1+\theta)}$	$\frac{z^{-\phi}+\theta}{1+\theta}$
$\left([\Upsilon']^{-1}\right)'(x)$	$\frac{(\theta x-(\theta-1))^{2-\rho}}{\theta(\rho-1)}$	$\frac{\phi}{(1-\phi)(1+\theta)}[(1+\theta)x-\theta]^{2-\frac{1}{\phi}}$	$-\epsilon\left[(1+\theta)x-\theta\right)\right]^{\frac{1+\epsilon(1+\theta)}{\epsilon(1+\theta)}}$	$-\frac{\phi[(1+\theta)x-\theta]^{\frac{(1+\phi)}{\phi}}}{(1+\theta)}$
$\varepsilon(x)=-\frac{1}{x}\frac{\Upsilon'(x)}{\Upsilon''(x)}$	$\frac{1}{1-\rho}\left[1-\frac{\theta-1}{\theta}x^{-1}\right]$	$\frac{\phi}{\phi-1}\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$	$\epsilon\left[(1+\theta)-\theta x^{-1}\right]$	$\phi\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$
Change of parameters	$\epsilon=\frac{1}{1-\rho}$	$\epsilon=\frac{\phi}{\phi-1}$	$\epsilon=\epsilon(1+\theta)$	$\epsilon=\phi$
(Simplified) Elasticity of demand	$\epsilon\left[1-\frac{\theta-1}{\theta}x^{-1}\right]$	$\epsilon\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$	$\epsilon\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$	$\epsilon\left[1-\frac{\theta}{1+\theta}x^{-1}\right]$

E.2. Constant super-elasticity of demand (CSE) aggregator

The super-elasticity of demand for an aggregator $\Upsilon(\cdot)$ is

$$\xi_i = -\frac{p_i}{P}\frac{\partial \log \varepsilon_i}{\partial\left(\frac{p_i}{P}\right)} = 1+\varepsilon_i+\varepsilon_i\frac{\frac{y_i}{Y}\Upsilon_i'''\left(\frac{y_i}{Y}\right)}{\Upsilon_i''\left(\frac{y_i}{Y}\right)} = 1-\frac{\Upsilon_i'\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon_i''\left(\frac{y_i}{Y}\right)}-\frac{\Upsilon_i'\left(\frac{y_i}{Y}\right)\Upsilon_i'''\left(\frac{y_i}{Y}\right)}{\left(\Upsilon_i''\left(\frac{y_i}{Y}\right)\right)^2}. \tag{E.5}$$

$\Upsilon(\cdot)$ exhibits a constant super-elasticity of θ if it solves the differential equation:

$$\theta = 1-\frac{\Upsilon'\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon''\left(\frac{y_i}{Y}\right)}-\frac{\Upsilon'\left(\frac{y_i}{Y}\right)\Upsilon'''\left(\frac{y_i}{Y}\right)}{\left(\Upsilon''\left(\frac{y_i}{Y}\right)\right)^2}. \tag{E.6}$$

The solution is:

$$\Upsilon\left(\frac{y_i}{Y}\right) \propto \Gamma\left[1+\frac{1}{\theta}, \frac{\epsilon_i}{\theta} - \log\frac{y_i}{Y}\right] = \int_{\frac{\epsilon_i}{\theta}-\log\frac{y_i}{Y}}^{\infty} \left(t^{\frac{1}{\theta}} e^{-t}\right) dt\,, \tag{E.7}$$

with derivatives:

$$\Upsilon'\left(\frac{y_i}{Y}\right) \propto e^{-\frac{\epsilon_i}{\theta}} \left(\frac{\epsilon_i}{\theta} - \log\frac{y_i}{Y}\right)^{\frac{1}{\theta}} \tag{E.8}$$

$$\Upsilon''\left(\frac{y_i}{Y}\right) \propto -\frac{e^{-\frac{\epsilon_i}{\theta}} \left(\frac{\epsilon_i}{\theta} - \log\frac{y_i}{Y}\right)^{\frac{1}{\theta}}}{\frac{y_i}{Y}\left(\epsilon_i - \theta\log\frac{y_i}{Y}\right)} \tag{E.9}$$

$$\Upsilon'''\left(\frac{y_i}{Y}\right) \propto -\frac{e^{-\frac{\epsilon_i}{\theta}} \left(\frac{\epsilon_i}{\theta} - \log\frac{y_i}{Y}\right)^{\frac{1}{\theta}}}{\left(\frac{y_i}{Y}\right)^2 \left(\epsilon_i - \theta\log\frac{y_i}{Y}\right)^2} \left(1 - \theta + \epsilon_i - \theta\log\frac{y_i}{Y}\right) \tag{E.10}$$

which imply an elasticity of demand of:

$$\varepsilon_i = \frac{-\Upsilon'\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon''\left(\frac{y_i}{Y}\right)} = \epsilon_i - \theta\log\frac{y_i}{Y}\,. \tag{E.11}$$

Implications for market shares. If the super-elasticity of demand is constant it is possible to recover market shares as a function of elasticities, ε, and the super-elasticity, θ as in the Proposition below. The implication of this is that, having recovered elasticities as in Section 4.1, the estimate is generically not compatible with the observed market shares for any given super-elasticity θ. Under this aggregator we cannot jointly match markup and sales-share data.

Proposition A3. *If the super-elasticity of demand is constant, the market shares satisfy*

$$\sigma_i = \frac{e^{-\frac{\varepsilon_i}{\theta}} \left(\frac{\varepsilon_i}{\theta}\right)^{\frac{1}{\theta}}}{\sum_j e^{-\frac{\varepsilon_j}{\theta}} \left(\frac{\varepsilon_j}{\theta}\right)^{\frac{1}{\theta}}}. \tag{E.12}$$

Proof. The elasticity satisfies $\varepsilon_i = \epsilon_i - \theta\log\frac{y_i}{Y}$. From this we obtain

$$\frac{\varepsilon_i}{\theta} = \frac{\epsilon_i}{\theta} - \log\frac{y_i}{Y} \qquad \text{and} \qquad e^{-\frac{\epsilon_i}{\theta}} = \frac{e^{-\frac{\varepsilon_i}{\theta}}}{y_i/Y}.$$

Then, we use this along with the first derivative of Υ in equation (E.8) to express market shares as in equation (9) in Lemma 1 and obtain

$$\sigma_i = \frac{\Upsilon'\left(\frac{y_i}{Y}\right)\frac{y_i}{Y}}{\sum_j \Upsilon'\left(\frac{y_j}{Y}\right)\frac{y_j}{Y}} = \frac{e^{-\frac{\epsilon_i}{\theta}} \left(\frac{\epsilon_i}{\theta} - \log\frac{y_i}{Y}\right)^{\frac{1}{\theta}} \frac{y_i}{Y}}{\sum_j e^{-\frac{\epsilon_j}{\theta}} \left(\frac{\epsilon_j}{\theta} - \log\frac{y_j}{Y}\right)^{\frac{1}{\theta}} \frac{y_j}{Y}} = \frac{e^{-\frac{\varepsilon_i}{\theta}} \left(\frac{\varepsilon_i}{\theta}\right)^{\frac{1}{\theta}}}{\sum_j e^{-\frac{\varepsilon_j}{\theta}} \left(\frac{\varepsilon_j}{\theta}\right)^{\frac{1}{\theta}}}. \tag{E.13}$$

□

Appendix F. Estimation Appendix

F.1. Recovering demand elasticities

Cournot Markups. Consider the reciprocal of the markup as expressed in equation (25) of Proposition 2 and the elasticity in (16) in Proposition 1.

$$\frac{1}{\mu_i^m} = 1 - \frac{1}{\bar{\varepsilon}_i^m} = 1 - \frac{1}{\gamma}\sigma_i^m - \left(\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right) + \frac{1}{\bar{\varepsilon}_{-i}^m}\sigma_i^m\right)\left(1-\sigma_i^m\right). \tag{F.1}$$

Rearranging,

$$\left(\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right) + \frac{1}{\bar{\varepsilon}_{-i}^m}\sigma_i^m\right)\left(1-\sigma_i^m\right) = 1 - \frac{1}{\gamma}\sigma_i^m - \frac{1}{\mu_i^m} \tag{F.2}$$

$$\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right) + \frac{1}{\bar{\varepsilon}_{-i}^m}\sigma_i^m = \frac{1-\frac{1}{\gamma}\sigma_i^m}{1-\sigma_i^m} - \frac{1}{\mu_i^m\left(1-\sigma_i^m\right)} \tag{F.3}$$

$$\frac{1}{\varepsilon_i^m}\left(1-\sigma_i^m\right) + \sum_{j\neq i}\frac{1}{\bar{\varepsilon}_j^m}\sigma_j^m\frac{\sigma_i^m}{1-\sigma_i^m} = \frac{1-\frac{1}{\gamma}\sigma_i^m}{1-\sigma_i^m} - \frac{1}{\mu_i^m\left(1-\sigma_i^m\right)}. \tag{F.4}$$

This is a linear system of equations on the reciprocal of the own-elasticities, $\left\{\varepsilon_i^m\right\}$. In matrix form,

$$\begin{bmatrix} 1-\sigma_1^m & \frac{\sigma_1^m}{1-\sigma_1^m}\sigma_2^m & \cdots & \frac{\sigma_1^m}{1-\sigma_1^m}\sigma_{N_m}^m \\ \frac{\sigma_2^m}{1-\sigma_2^m}\sigma_1^s & 1-\sigma_2^m & \cdots & \frac{\sigma_2^m}{1-\sigma_2^m}\sigma_{N_m}^m \\ \vdots & & \ddots & \\ \frac{\sigma_{N_m}^m}{1-\sigma_{N_m}^m}\sigma_1^m & \frac{\sigma_{N_m}^m}{1-\sigma_{N_m}^m}\sigma_2^m & \cdots & 1-\sigma_{N_m}^m \end{bmatrix} \begin{bmatrix} \frac{1}{\varepsilon_1^m} \\ \frac{1}{\varepsilon_2^m} \\ \vdots \\ \frac{1}{\varepsilon_{N_m}^m} \end{bmatrix} = \begin{bmatrix} \frac{1-\frac{1}{\gamma}\sigma_1^m}{1-\sigma_1^m} - \frac{1}{\mu_1^m(1-\sigma_1^m)} \\ \frac{1-\frac{1}{\gamma}\sigma_2^m}{1-\sigma_2^m} - \frac{1}{\mu_2^m(1-\sigma_2^m)} \\ \vdots \\ \frac{1-\frac{1}{\gamma}\sigma_{N_m}^m}{1-\sigma_{N_m}^m} - \frac{1}{\mu_{N_m}^m\left(1-\sigma_{N_m}^m\right)}. \end{bmatrix} \tag{F.5}$$

The matrix in the left is a stochastic matrix and is generically invertible. However, the system can only be solved exactly if it holds that

$$1 > \frac{1}{\gamma}\sigma_i^m + \frac{1}{\mu_i^m} \tag{F.6}$$

for all firms in each market. This restriction comes from the implications of the model for concentration and markups. Larger firms hold more market power and command higher markups. This relationship has to be respected in the data. An alternative way of understanding this condition is that it is necessary for the the own-elasticities to be positive $\varepsilon_i^m > 0$. This can be seen by expressing the condition in terms of the firms' average elasticity

$$\frac{1}{\bar{\varepsilon}_i^m} > \frac{1}{\gamma}\sigma_i^m. \tag{F.7}$$

Bertrand Markups. We prove that the injectivity of the markup function

$$\mu_i = F_i(\vec{\varepsilon}) \equiv \frac{\sigma_i + \varepsilon_i\left[1 - \frac{\varepsilon_i\sigma_i}{\bar{\varepsilon}}\right]}{\sigma_i + \varepsilon_i\left[1 - \frac{\varepsilon_i\sigma_i}{\bar{\varepsilon}}\right] - 1},$$

where $\bar{\varepsilon} = \sum_{j=1}^{N} \varepsilon_j \sigma_j$. This implies that if, given market shares $\{\sigma_i\}$, there is a vector of elasticities $\{\varepsilon_i\}$ that generates the observed vector of markups $\{\mu_i\}$, then this vector of elasticities is unique. For notational simplicity we omit the sector script, s.

To prove that F is injective we evaluate its Jacobian and show that it is a P-matrix. The result then follows from Theorem 4 in Gale and Nikaido (1965).

We start by examining the first derivatives of F_i. It will be useful to define the auxiliary function

$$g_i(\varepsilon) = \sigma_i + \varepsilon_i\left[1 - \frac{\varepsilon_i\sigma_i}{\bar{\varepsilon}}\right]$$

so that:

$$F_i(\varepsilon) = \frac{g_i(\varepsilon)}{g_i(\varepsilon) - 1} \qquad \text{and} \qquad \frac{\partial F_i}{\partial \varepsilon_j} = \frac{-1}{\left(g_i(\varepsilon) - 1\right)^2}\frac{\partial g_i(\varepsilon)}{\partial \varepsilon_j}$$

We need to determine the value of the derivatives of g_i with respect to ε_i and ε_j:

$$\frac{\partial g_i(\varepsilon)}{\partial \varepsilon_i} = \left[1 - \frac{\varepsilon_i\sigma_i}{\bar{\varepsilon}}\right] - \varepsilon_i\sigma_i\left[\frac{\bar{\varepsilon} - \varepsilon_i\sigma_i}{\bar{\varepsilon}^2}\right] = \left[1 - \frac{\varepsilon_i\sigma_i}{\bar{\varepsilon}}\right]^2$$

$$\frac{\partial g_i(\varepsilon)}{\partial \varepsilon_j} = \frac{\varepsilon_i^2\sigma_i\sigma_j}{\bar{\varepsilon}^2}$$

We want to show that the Jacobian of F is a P-matrix, which is defined as a matrix having positive determinants for all its principal minors. The principal minor of the Jacobian excludes one of the rows and one of the columns with the same index. Instead of tackling the full Jacobian we note that it is proportional to:

$$\boldsymbol{J}^N \propto -\tilde{\boldsymbol{J}}^N \equiv -\begin{bmatrix} \left[\sum_{j\neq 1}\varepsilon_j\sigma_j\right]^2 & \cdots & \varepsilon_1^2\sigma_1\sigma_i & \cdots & \varepsilon_1^2\sigma_1\sigma_N \\ \vdots & \ddots & & & \vdots \\ \varepsilon_i^2\sigma_i\sigma_1 & & \left[\sum_{j\neq i}\varepsilon_j\sigma_j\right]^2 & & \varepsilon_i^2\sigma_i\sigma_N \\ \vdots & & & \ddots & \vdots \\ \varepsilon_N^2\sigma_N\sigma_1 & \cdots & \varepsilon_N^2\sigma_N\sigma_i & \cdots & \left[\sum_{j\neq N}\varepsilon_j\sigma_j\right]^2 \end{bmatrix}$$

So it is sufficient to prove that $\tilde{J}$ is a P-matrix. We do this by induction. It is immediate that for $N = 1$ and $N = 2$ all principal minors are (weakly) positive because the diagonal is non-negative. It is also easy to check the determinant directly when $N = 3$. Now we prove formally that $\tilde{\boldsymbol{J}}$ is a P-matrix for $N \geq 3$ under the induction hypothesis that it is a

P-matrix for $N-1$. Without loss of generality, the principal minor of $\tilde{\boldsymbol{J}}$ has the form:

$$\tilde{\boldsymbol{J}}^{N}_{PM} = \begin{bmatrix} \left[\sum_{j\neq 1}\varepsilon_j\sigma_j\right]^2 & \dots & \varepsilon_1^2\sigma_1\sigma_i & \dots & \varepsilon_1^2\sigma_1\sigma_{k-1} & \varepsilon_1^2\sigma_1\sigma_{k+1} & \dots & \varepsilon_1^2\sigma_1\sigma_N \\ \vdots & \ddots & & & & & & \vdots \\ \varepsilon_i^2\sigma_i\sigma_1 & & \left[\sum_{j\neq i}\varepsilon_j\sigma_j\right]^2 & & & & & \varepsilon_i^2\sigma_i\sigma_N \\ \vdots & & & \ddots & & & & \vdots \\ \varepsilon_{k-1}^2\sigma_{k-1}\sigma_1 & & & & \left[\sum_{j\neq k-1}\varepsilon_j\sigma_j\right]^2 & & & \varepsilon_{k-1}^2\sigma_{k-1}\sigma_N \\ \varepsilon_{k+1}^2\sigma_{k+1}\sigma_1 & & & & & \left[\sum_{j\neq k+1}\varepsilon_j\sigma_j\right]^2 & & \varepsilon_{k+1}^2\sigma_{k+1}\sigma_N \\ \vdots & & & & & & \ddots & \vdots \\ \varepsilon_N^2\sigma_N\sigma_1 & \dots & \varepsilon_N^2\sigma_N\sigma_i & \dots & \varepsilon_N^2\sigma_N\sigma_{k-1} & \varepsilon_N^2\sigma_N\sigma_{k+1} & \dots & \left[\sum_{j\neq N_s}\varepsilon_j\sigma_j\right]^2 \end{bmatrix}$$

The matrix $\tilde{\boldsymbol{J}}^{N}_{PM}$ can be written as:

$$\tilde{\boldsymbol{J}}^{N}_{PM} = \tilde{\boldsymbol{J}}^{N-1}_{PM} + \text{diag}\left[\varepsilon_k\sigma_k\left[2\sum_{j\neq i,k}\varepsilon_j\sigma_j + \varepsilon_k\sigma_k\right]\right]$$

where diag $\left[a_i\right]$ is a diagonal matrix with diagonal entries $\{a_1, \dots, a_{N-1}\}$. The result follows because

$$\left|\tilde{\boldsymbol{J}}^{N}_{PM}\right| \geq \left|\tilde{\boldsymbol{J}}^{N-1}_{PM}\right| + \varepsilon_k\sigma_k\prod_{i\neq k}\left[2\sum_{j\neq i,k}\varepsilon_j\sigma_j + \varepsilon_k\sigma_k\right] \geq \left|\tilde{\boldsymbol{J}}^{N-1}_{PM}\right| \geq 0,$$

where the first inequality follows from the properties of determinants, the second inequality from the fact $\varepsilon_i, \sigma_i \geq 0$ for all i, and the last inequality comes from the induction hypothesis. This completes the proof.

F.2. Identification of demand parameters

We construct the result in Proposition 4 in three steps. We first relate changes in market shares to changes in firm output. Then, we relate changes in the elasticity of demand to changes in firm output. Finally, we bring these two results together to complete the proof of Proposition 4. In what follows we omit the sector script, s.

Market shares and output changes. We derive an equation relating the change in the market share of a firm σ_i when there is an (exogenous) change in firm output y_i, while taking into account the effects on aggregate quantity (Y) and price (P).

We start from the definition of $\sigma_i \equiv p_i y_i / PY$ and group the effect of the change in output (y_i) on relative price (p_i/P) and quantity (y_i/Y). The effect is obtained from the total differential of σ_i.

$$\sigma_i = \left(\frac{p_i}{P}\right)\left(\frac{y_i}{Y}\right)$$

$$d\sigma_i = \left(\frac{1}{P}\frac{\partial p_i}{\partial y_i}dy_i - \frac{p_i}{P^2}\frac{\partial P}{\partial p_i}\frac{\partial p_i}{\partial y_i}dy_i\right)\left(\frac{y_i}{Y}\right) + \left(\frac{p_i}{P}\right)\left(\frac{1}{Y}dy_i - \frac{y_i}{Y^2}\frac{\partial Y}{\partial y_i}dy_i\right) \qquad \text{Total Differential}$$

$$d\sigma_i = \left(\frac{1}{P} - \frac{p_i}{P^2}\frac{\partial P}{\partial p_i}\right)\left(\frac{y_i}{Y}\right)\frac{\partial p_i}{\partial y_i}dy_i + \left(\frac{p_i}{P}\right)\left(\frac{1}{Y} - \frac{y_i}{Y^2}\frac{\partial Y}{\partial y_i}\right)dy_i \qquad \text{Group } dy_i \text{ terms}$$

$$d\sigma_i = \left(\frac{1}{P} - \frac{p_i}{P^2}\frac{y_i}{Y}\right)\left(\frac{y_i}{Y}\right)\frac{\partial p_i}{\partial y_i}dy_i + \left(\frac{p_i}{P}\right)\left(\frac{1}{Y} - \frac{y_i}{Y^2}\frac{p_i}{P}\right)dy_i \qquad \frac{\partial P}{\partial p_i} = \frac{y_i}{Y}, \frac{\partial Y}{\partial y_i} = \frac{p_i}{P}$$

$$d\sigma_i = \frac{1}{P}(1-\sigma_i)\left(\frac{y_i}{Y}\right)\frac{\partial p_i}{\partial y_i}dy_i + \frac{1}{Y}\left(\frac{p_i}{P}\right)(1-\sigma_i)\,dy_i \qquad \text{Express ratios as } \sigma_i$$

$$d\sigma_i = \left(\frac{1}{p_i}\frac{\partial p_i}{\partial y_i} + \frac{1}{y_i}\right)\frac{y_i}{Y}\frac{p_i}{P}(1-\sigma_i)\,dy_i \qquad \text{Factor terms}$$

$$d\sigma_i = \left(\frac{1}{p_i}\frac{\partial p_i}{\partial y_i} + \frac{1}{y_i}\right)\sigma_i(1-\sigma_i)\,dy_i \qquad \text{Express ratios as } \sigma_i$$

$$d\sigma_i = \left(\frac{y_i}{p_i}\frac{\partial p_i}{\partial y_i} + 1\right)\sigma_i(1-\sigma_i)\frac{dy_i}{y_i} \qquad \text{Factor out } \frac{1}{y_i}$$

$$d\sigma_i = \left(\frac{1}{\varepsilon_i} + 1\right)\sigma_i(1-\sigma_i)\,d\log y_i \qquad \text{Express as } \varepsilon, d\log y_i$$

The final step gives us the equilibrium relationship we wanted, so to a first order approximation we know that the market share of firm i evolves as:

$$\sigma_i = \overline{\sigma}_i + \left(\frac{1}{\varepsilon_i} + 1\right)\sigma_i(1-\sigma_i)\,d\log y_i \tag{F.8}$$

Alternatively we can write the relationship in terms of the change in logs:

$$d\log\sigma_i = \left(\frac{1}{\varepsilon_i} + 1\right)(1-\sigma_i)\,d\log y_i. \tag{F.9}$$

Demand elasticities and output changes. We derive an equation relating the change in the elasticity of a firm ε_i when there is an (exogenous) change in firm output y_i, while taking into account the effects on aggregate quantity (Y) and price (P).

We start from the definition of the elasticity,

$$\varepsilon_i \equiv \left(-\frac{y_i}{p_i}\frac{\partial p_i}{\partial y_i}\right)^{-1} = -\frac{\Upsilon'\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon''\left(\frac{y_i}{Y}\right)}, \tag{F.10}$$

and use the first order condition of the aggregator firm to express it in terms of relative output (y_i/Y). The effect is obtained from the total differential of ε_i.

$$\begin{aligned}
\varepsilon_i &= -\frac{\Upsilon'\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon''\left(\frac{y_i}{Y}\right)} \\
d\varepsilon_i &= -\left(\frac{\Upsilon''\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon''\left(\frac{y_i}{Y}\right)}d\left(\frac{y_i}{Y}\right) - \frac{\Upsilon'\left(\frac{y_i}{Y}\right)}{\left(\frac{y_i}{Y}\right)^2\Upsilon''\left(\frac{y_i}{Y}\right)}d\left(\frac{y_i}{Y}\right) - \frac{\Upsilon'\left(\frac{y_i}{Y}\right)\Upsilon'''\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\left(\Upsilon''\left(\frac{y_i}{Y}\right)\right)^2}d\left(\frac{y_i}{Y}\right)\right) \\
d\varepsilon_i &= -\left(\left(\frac{y_i}{Y}\right)^{-1} + \varepsilon_i\left(\frac{y_i}{Y}\right)^{-1} + \varepsilon_i\frac{\Upsilon'''\left(\frac{y_i}{Y}\right)}{\Upsilon''\left(\frac{y_i}{Y}\right)}\right)d\left(\frac{y_i}{Y}\right) \\
\varepsilon_i &= -\left(\left(\frac{y_i}{Y}\right)^{-1} + \varepsilon_i\left(\frac{y_i}{Y}\right)^{-1} + \varepsilon_i\frac{\Upsilon'''\left(\frac{y_i}{Y}\right)}{\Upsilon''\left(\frac{y_i}{Y}\right)}\right)\left(\frac{1}{Y} - \frac{y_i}{Y^2}\frac{\partial Y}{\partial y_i}\right)dy_i \\
d\varepsilon_i &= -\left(\left(\frac{y_i}{Y}\right)^{-1} + \varepsilon_i\left(\frac{y_i}{Y}\right)^{-1} + \varepsilon_i\frac{\Upsilon'''\left(\frac{y_i}{Y}\right)}{\Upsilon''\left(\frac{y_i}{Y}\right)}\right)\frac{y_i}{Y}\left(1 - \frac{y_i}{Y}\frac{p_i}{Y}\right)d\log y_i \\
d\log\varepsilon_i &= \frac{-1}{\varepsilon_i}\left(1 + \varepsilon_i + \varepsilon_i\frac{\frac{y_i}{Y}\Upsilon'''\left(\frac{y_i}{Y}\right)}{\Upsilon''\left(\frac{y_i}{Y}\right)}\right)\left(1-\sigma_i\right)d\log y_i
\end{aligned}$$

We can simplify this expression by introducing the super-elasticity of demand (see, Klenow and Willis 2016):

$$\xi_i \equiv -\frac{p_i}{P}\frac{\partial\log\varepsilon_i}{\partial\left(\frac{p_i}{P}\right)} = 1 - \frac{\Upsilon'\left(\frac{y_i}{Y}\right)}{\frac{y_i}{Y}\Upsilon''\left(\frac{y_i}{Y}\right)} - \frac{\Upsilon'\left(\frac{y_i}{Y}\right)\Upsilon'''\left(\frac{y_i}{Y}\right)}{\left(\Upsilon''\left(\frac{y_i}{Y}\right)\right)^2} = 1 + \varepsilon_i + \varepsilon_i\frac{\frac{y_i}{Y}\Upsilon'''\left(\frac{y_i}{Y}\right)}{\Upsilon''\left(\frac{y_i}{Y}\right)}. \tag{F.11}$$

Replacing gives the desired expression:

$$d\log\varepsilon_i = -\frac{\xi_i}{\varepsilon_i}\left(1-\sigma_i\right)d\log y_i. \tag{F.12}$$

Demand elasticities and market share changes. We derive a an equation relating the change in the elasticity of a firm ε_i to the change in its market share σ_i.

From the first set of results we get:

$$d\log y_i = \frac{d\log \sigma_i}{\left(\frac{1}{\varepsilon_i}+1\right)(1-\sigma_i)}.$$

Replacing in the second equation gives us the desired result:

$$d\log \varepsilon_i = -\left(\frac{\xi_i}{\varepsilon_i}\right)\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right) d\log \sigma_i. \tag{F.13}$$

This expression provides us with an equation that we can take to the data, exploiting the demand elasticities $\{\varepsilon_i\}$ that we recover from data on markups and market shares $\{\mu_i, \sigma_i\}$.

Regression parameters. We estimate the equation in Proposition 4 to obtain

$$\Delta\log \varepsilon_i^{st} = \hat{\beta}\cdot \Delta\log \sigma_i^{st}. \tag{F.14}$$

The population value of the coefficient β is a weighted average of the super-elasticity to elasticity ratio:

$$\begin{aligned}
\beta &= \frac{\sum (\Delta\log \varepsilon_i)(\Delta\log \sigma_i)}{\sum (\Delta\log \sigma_i)^2} \\
\beta &= \frac{\sum \left(-\left(\frac{\xi_i}{\varepsilon_i}\right)\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\Delta\log \sigma_i\right)(\Delta\log \sigma_i)}{\sum (\Delta\log \sigma_i)^2} \\
\beta &= -\sum \left(\frac{\xi_i}{\varepsilon_i}\right)\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\frac{(\Delta\log \sigma_i)^2}{\sum (\Delta\log \sigma_i)^2}
\end{aligned}$$

We can further express this coefficient in terms of the underlying function Υ_i by replacing the super-elasticity

$$\beta = -\sum \left(1+\frac{\varepsilon_i}{1+\varepsilon_i}\frac{\frac{y_i}{Y}\Upsilon'''\left(\frac{y_i}{Y}\right)}{\Upsilon''\left(\frac{y_i}{Y}\right)}\right)\frac{(\Delta\log \sigma_i)^2}{\sum (\Delta\log \sigma_i)^2} \tag{F.15}$$

However, this does not prove to be useful, as the ratio ξ/ε proves to be more tractable for the leading functional for Υ_i in the literature.

F.3. Estimation for the Klenow & Willis aggregator

We first restate the elasticity and super-elasticity for the Klenow and Willis (2016)aggregator:

$$\varepsilon_i = \epsilon_i \left(\frac{y_i}{Y}\right)^{-\frac{\theta}{\epsilon_i}} \qquad \text{Elasticity}$$

$$\xi_i = \theta \left(\frac{y_i}{Y}\right)^{-\frac{\theta}{\epsilon_i}} \qquad \text{Superelasticity}$$

$$\frac{\xi_i}{\varepsilon_i} = \frac{\theta}{\epsilon_i} \qquad \text{Ratio}$$

With this, the regression coefficient becomes

$$\beta^{\text{KW}}\left(\theta, \{\epsilon_i, \sigma_i\}\right) = \frac{-\sum \left(\frac{\theta}{\epsilon_i}\right)\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\left(\Delta \log \sigma_i\right)^2}{\sum \left(\Delta \log \sigma_i\right)^2} = \frac{-\sum \left(\frac{1}{\epsilon_i}\right)\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\left(\Delta \log \sigma_i\right)^2}{\sum \left(\Delta \log \sigma_i\right)^2}\theta.$$

Algorithm.

(a) Observe $\{\mu_i, \sigma_i\}$ for a cross section of N firms in a given year (t) and sector (s)

(b) Use markup equation to back out elasticities $\{\varepsilon_i\}$. This is independent of Kimball parameters and depends only on the form of competition.

(c) Run regression in data, save $\hat{\beta} = \frac{\sum \Delta \log \varepsilon_i \cdot \Delta \log \sigma_i}{\sum (\Delta \log \sigma_i)^2}$

(d) Guess $\theta \geq 0$ that is common across firms. **Note:** $\theta = 0$ implies $\beta^{\text{KW}}(0, \cdot) = 0$.

(e) Solve for parameters $\{\epsilon_i\}$ so as to match market shares $\{\sigma_i\}$.

(f) Evaluate $\left|\hat{\beta} - \beta\left(\theta, \{\epsilon_i, \sigma_i\}\right)\right|$, note that all is observable given θ.

(i) The objective satisfies:

$$\left|\hat{\beta} - \beta\left(\theta, \{\epsilon_i, \sigma_i\}\right)\right| \propto \left|\sum \left(\Delta \log \varepsilon_i - \left(\frac{\theta}{\epsilon_i}\right)\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\Delta \log \sigma_i\right) \cdot \Delta \log \sigma_i\right|$$

which can be useful if we need to rescale since $\sum \left(\Delta \log \sigma_i\right)^2$ might be numerically hard to work with. The expression also has a nice interpretation as a weighted difference with weights given by the change in market shares. So we are trying to match closer the change in elasticity of the firms that changed their market shares the most.

(g) Repeat (4)-(6) to minimize distance

F.4. Estimation for the Dotsey & King aggregator

The process is the same, only changing the functional form of the elasticity ans super-elasticity in steps (e) and (f).

F.5. Recovering model quantities and prices

Relative quantities and prices.

(a) Record the sales share of each firm for a given sector s.

(b) Solve the system of non-linear equations defined by

$$\sigma_i^s = \frac{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)\frac{y_i^s}{Y_s}}{\sum_j \Upsilon'\left(\frac{y_j^s}{Y_s}\right)\frac{y_j^s}{Y_s}},$$

(i) Change variables to $x_i^s \equiv \Upsilon'\left(\frac{y_i^s}{Y_s}\right)\frac{y_i^s}{Y_s}$. The system is now:

$$\sigma_i^s \sum_j x_j^s = x_j^s,$$

or in matrix form:

$$\Sigma_s \vec{x}_s = \vec{x}_s,$$

where

$$\Sigma_s = \vec{\sigma}_s \otimes 1_{1\times N} = \begin{bmatrix} \sigma_1^s & \cdots & \sigma_1^s & \cdots & \sigma_1^s \\ \vdots & \ddots & & & \\ \sigma_n^s & \cdots & \sigma_n^s & \cdots & \sigma_n^s \\ \vdots & & & \ddots & \\ \sigma_N^s & \cdots & \sigma_N^s & \cdots & \sigma_N^s \end{bmatrix} \qquad \text{and} \qquad \vec{x}_s = \begin{bmatrix} x_1^s \\ \vdots \\ x_n^s \\ \vdots \\ x_N^s \end{bmatrix}.$$

(ii) Solve the linear problem. The solution is given by the eigenvector of Σ_s associated to its unit eigenvalue.

- The existence of a unit eigenvalue is guaranteed because Σ_s is a stochastic matrix. By the Perron–Frobenius theorem the unit eigenvector is also the largest eigenvalue, and the eigenvector can be chosen to be positive.
- Note that eigenvectors are only unique up to a scalar multiple.

(iii) Transform back to relative output units (y_i^s/Y_s) by solving the following system:

$$\alpha_s x_i^s = \Upsilon'\left(\frac{y_i^s}{Y_s}\right)\frac{y_i^s}{Y_s} \quad \forall_i \qquad \text{and} \qquad 1 = \sum_i \Upsilon\left(\frac{y_i^s}{Y_s}\right),$$

the scale $\alpha_s > 0$ is chosen to ensure that the aggregate output equation is satisfied. The solution of the system depends on the functional form of the Kimball aggregator. We consider two cases:

i. **D&K (2005):**

$$\Upsilon\left(\frac{y_i^s}{Y_s}\right) = \frac{1}{(1+\theta)\frac{\epsilon-1}{\epsilon}}\left[(1+\theta)\frac{y_i^s}{Y_s} - \theta\right]^{\frac{\epsilon-1}{\epsilon}} - \left[1 + \frac{1}{(1+\theta)\frac{\epsilon-1}{\epsilon}}\right]$$

$$\Upsilon'\left(\frac{y_i^s}{Y_s}\right) = \left((1+\theta)\,\frac{y_i^s}{Y_s} - \theta\right)^{\frac{-1}{\epsilon}}$$

ii. **K&W (2016):**

$$\Upsilon\left(\frac{y_i^s}{Y_s}\right) = 1 + (\epsilon - 1)\exp\left(\frac{1}{\theta}\right)\theta^{\frac{\epsilon}{\theta}-1}\left[\Gamma\left(\frac{\epsilon}{\theta}, \frac{1}{\theta}\right) - \Gamma\left(\frac{\epsilon}{\theta}, \frac{(y_i^s/Y)^{\frac{\theta}{\epsilon}}}{\theta}\right)\right]$$

$$\Upsilon'\left(\frac{y_i^s}{Y_s}\right) = \frac{\epsilon - 1}{\epsilon}\exp\left(\frac{1 - \frac{y_i^s}{Y_s}^{\frac{\theta}{\epsilon}}}{\theta}\right)$$

(c) Obtain relative prices as

$$\frac{p_i^s}{P_s} = \frac{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)}{\sum_j \Upsilon'\left(\frac{y_j^s}{Y_s}\right)\frac{y_j^s}{Y_s}} = \frac{\sigma_i^s}{y_i^s/Y_s}$$

(d) Compute elasticities of demand by firm:

$$\varepsilon\left(\frac{y_i^s}{Y_s}\right) = -\left(\frac{y_i^s}{Y_s}\frac{\Upsilon''\left(\frac{y_i^s}{Y_s}\right)}{\Upsilon'\left(\frac{y_i^s}{Y_s}\right)}\right)^{-1} = \begin{cases} \epsilon\left(1 - \frac{\theta}{1+\theta}\left(\frac{y_i^s}{Y_s}\right)^{-1}\right) & If \quad D\&K(2005) \\ \epsilon\left(\frac{y_i^s}{Y_s}\right)^{-\frac{\theta}{\epsilon}} & If \quad K\&W(2016) \end{cases}$$

(e) Compute markups according to type of competition: