Variable Markups, Demand Elasticity and Pass-through of Marginal Costs into Prices*

Anastasia Burya[†] and Shruti Mishra[‡]

Latest version

Abstract

Using a novel theoretical framework for a general oligopolistic market, we derive sufficient statistics to empirically estimate elasticities of demand and optimal pass-through of marginal costs into prices, using the ACNielsen Retail Scanner database. Our main findings are: 1) elasticities of demand are large for small firms, but decrease as the firm's market share increases; 2) there is a positive dependence of demand elasticities on relative prices (superelasticity), in line with Marshall's second law of demand; 3) an individual firm's pass-through decreases with the firm's market share; and 4) pass-through depends positively on the size of the marginal cost shock. This last finding means that the total effect of marginal cost shock on prices is non-linear and that firm prices are more responsive to marginal cost increases than to marginal cost decreases. For market leaders, the pass-through of a large negative marginal cost shock would be close to zero, while the pass-through of a large positive marginal cost shock would approach that of small firms.

*We are grateful to our advisors Hassan Afrouzi, Jennifer La'O and Michael Woodford for invaluable comments and support as well as Émilien Gouin-Bonenfant, Matthieu Gomez, Gautam Gowrisankaranm, Rui Mano, Stephanie Schmitt-Grohé, Jesse Schreger, Andrey Simonov, Pietro Tebaldi, Yannick Timmer, Martin Uribe, Anke Weber, David Weinstein and all the participants of the Economic Fluctuation Colloquium at Columbia University.
Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.
The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[†]Columbia University

[‡]Bank of America Securities

1 Introduction

When elasticity of demand is not constant with respect to the firm's price, optimal pass-through of marginal costs into prices is imperfect. This means that a one-unit marginal cost shock causes a less-than-one unit price change. Intuitively, this occurs due to strategic considerations in firms' price-setting behaviors. For instance, if firm expects that a larger price would negatively affect its market share and profits, it might optimally choose not to increase prices as much, even if its costs are growing.

Non-constant elasticity and imperfect pass-through are important parameters in many macroeconomic models that influence some of their main predictions. Their effects on price-setting go beyond the common price adjustment friction. For instance, they make the prices of some firms act as if they were more sticky than would be expected from the price-adjusting friction alone. This will cause a change in the desired price response and, potentially, alter optimal inflation and influence long-run price dynamics.

How can we assess the optimal imperfect pass-through? This outcome is the result of the variability of the desired markup, which in most models can be expressed as a function of demand elasticity:

$$\mu = \frac{\xi}{\xi - 1}$$

To complete this task, we need to look at the determinants of demand elasticity. However, empirical evidence on this subject remains limited for macroeconomic settings. In particular, little is known about the parameters of oligopolistic markets. The literature has mostly focused on monopolistic competition with atomistic firms, but the recent trends of rising superstar firms and increasing firm concentration may make it crucial to relax this assumption.

This paper fills the gap in evidence by studying the parameters of variable demand elasticity, applying them to assess the degree of imperfect optimal pass-through, and then going one step further to discuss the possible non-linear effects of the marginal costs on prices. Our empirical strategy differs from the approaches commonly found in the macroeconomic literature. Previous estimates of pass-through and related parameters, for instance those in Amiti et al. (2019), relied on the direct estimation of the optimal pricing equation. Our approach instead focuses on estimating the relevant parameters directly from the demand equation, which holds regardless of how sticky the prices are. We then reconstruct the pass-through from these structural parameters. This allows us to estimate the optimal pass-through separately from the impact of sticky prices.

Central to our approach is the estimation of the demand system. Within the oligopolistic setting, we develop a theoretical framework to account for the effect of a single firm's pricing on the market aggregates. We start from a nested structure, separating firm-level demand from more aggregated, market-level demand that unites products with similar characteristics. We continue to allow for any arbitrary demand aggregator at the firm level and Kimball-type demand at the market level. This framework includes many models used in the macroeconomic literature, such as Atkeson-Burstein oligopoly Atkeson and Burstein (2008), Kimball oligopoly Wang and Werning (2020), and the original Kimball monopolistic competition Kimball (1995), but ultimately goes beyond them. We then consider a semi-novel concept of "partial elasticity," the elasticity of demand given the market aggregates. We specifically show that the response of the market aggregates to the price changes of a single firm can be represented using partial elasticities and market shares. This means that partial elasticities are sufficient statistics for most demand-related parameters, which is why they are central to our empirical strategy.

The theoretical results will ensure the external validity properties of our estimation. Due to decomposition of the effect of a firm's price on the market aggregates and additional invariant results, under specific symmetry assumptions we can fully account for the market-level heterogeneity originating from the distribution of market shares. In this case, our estimates are not simply averages across many markets with different numbers of firms and different degrees of competitiveness, but consistent statistics that can be potentially applied in multiple settings.

As in any demand estimation, we need an instrument that shifts the supply curve. It is especially challenging in macroeconomic settings to find a unified instrument for different markets. We address this issue by using the rich structure of the Nielsen Retail Scanner dataset, which lets us construct an instrument by combining the ideas of the National Price instrument Hausman (1996) and the Granular IV approach Gabaix and Koijen (2020). This instrument bears similarities with Nevo (2001) and especially DellaVigna and Gentzkow (2019). For the baseline model, we start from an assumption that for firms that operate in multiple regions, prices are set nationally and demand shocks are more aggregated then supply shocks. While demand only varies for each individual firm, whether across all regions or for each individual region, supply shocks can occur at the firm-region level due to transport cost shocks or other supply-chain effects. This assumption would be violated if demand changed significantly for an individual firm in a particular region, such as due to a local advertising campaign. As a robustness check, we relax this last assumption and obtain similar results.

Our results allow to estimate pass-through and strategic complementarity on an unprecedentedly disaggregated level. We allow for heterogeneity across both firms and market structures. We estimate demand partial elasticities for various firm sizes, assess partial cross-elasticity, estimate the "elasticity of the elasticity" (often referred to as superelasticity), and then repeat everything at the market level. The resulting estimates produce sufficient statistics to calculate optimal pass-through of marginal costs into prices, as well as non-linear marginal cost effects.

Apart from being practical benchmarks for further analysis, our estimates offer several big-picture insights into the market structure. Our superelasticity estimates are positive, suggesting that elasticity increases with prices, which corroborates Marshall's second law. We also find that firm-level elasticity is higher than market-level elasticity, which would broadly suggest that elasticities are larger for smaller firms. Superelasticities at the firm-level are lower than superelasticities at the market level, contributing to the fact that the elasticity of demand changes more sharply with the price for the larger firms. Firm-level elasticities depend strongly on the market shares, suggesting a possible persistent heterogeneity across firms.

We use these results to assess the optimal pass-through and non-linear marginal cost effect. We find that pass-through is heterogeneous across firms with different market shares. Smaller firms with market shares below 1% would have a pass-through of approximately 0.8, while for the large firms with market shares over 30%, pass-through will be just 0.4. Next, we go one step further to study the previously unaccounted for non-linearity of marginal cost shock effects on the desired price.

We find that the elasticity of pass-through to a marginal cost shock is positive and significant. This creates a non-linearity in price responses to a marginal cost shock, and causes firm-level prices to be more responsive to marginal cost increases than to marginal cost decreases. This would mean, for instance, that even for large firms, the pass-through for a large positive marginal cost shock will be high and approach that of a small firm, while the pass-though for a negative marginal cost shock will be close to zero. In an economy where large firms play a significant role, we expect to see both missing deflation in the presence of negative cost shocks and strong inflation in the presence of large positive cost shocks. This feature of price setting offers an explanation for the recent trend in inflation and points out how the prominence of large firms in the economy can create both periods of relative price stability and large spikes in inflation, in response to an increase in marginal costs.

Our paper relates to several branches of the literature. First, it contributes to the empirical understanding of markups and pass-through, including at the aggregate level. Building on work by Nevo (2001), Nakamura and Zerom (2010), De Loecker and Warzynski (2012), De Loecker et al. (2020), Hottman et al. (2016), Amiti et al. (2019), and Foster et al. (2022), we offer unique estimates with detailed heterogeneity and discovers several novel facts. Secondly, by highlighting the asymmetry of price responses to marginal cost shocks, we enhance research on the impact of variable markups in the aggregate economy, including Rotemberg and Woodford (1992), Rotemberg and Woodford (1999), Kimball (1995), Atkeson and Burstein (2008), Gopinath and Itskhoki (2011), Amiti et al. (2014), Edmond et al. (2015), Edmond et al. (2018), Klenow and Willis (2016), Mongey (2021), Afrouzi (2020), Wang and Werning (2020), and Baqaee et al. (2021). Thirdly, we contribute to the broader literature on firm heterogeneity and price setting behaviour, e.g., Alvarez et al. (2020); Rubbo (2020); Afrouzi et al. (2022); La'O and Tahbaz-Salehi (2020). Lastly, we add new empirical evidence to the dynamic demand models, following Klemperer (1987), Ravn et al. (2006), Gowrisankaran and Rysman (2012), and Shcherbakov (2016).

The rest of the paper is organized as follows: Section 2 describes the theoretical framework, Section 3 discusses identification strategy and presents results of our estimation, Section 4 calculates the optimal pass-through, non-linear marginal cost effect, and other important parameters, and Section 5 concludes.

2 Theoretical Framework

In this section we discuss the general oligopolistic theoretical framework. We first derive sufficient statistics for measuring the firm elasticities and superelasticities that inform our empirical strategy. We then study the pass-through of marginal costs into prices under flexible pricing and derive the second-order asymmetric marginal cost effect.

2.1 Demand System

We start from a nested demand structure and separate two aggregation levels. At the most disaggregated level we consider individual firms producing firm-specific goods, indexed by ω. Individual firm goods are aggregated into similar-good categories that we call markets, indexed by j. We assume that there is a finite number of firms within each market, so individual firms are allowed to have a non-zero impact on the market aggregates. Production of individual markets is aggregated once again into the final consumption composite. We assume that there is an infinitely large number of markets in the economy, so that each market is infinitesimally small compared to the economy as a whole.

In the model, we equate one firm to one unique good, then equate the market to a unity of sufficiently similar goods. We therefore impose the assumption that either firms only produce goods for a single market or the prices for all the firm's markets are set independently from each other.

We assume that the substitution between markets is governed by the Kimball aggregator while substitution within the market is governed by an arbitrary finite-argument function F of individual firm-level production shares, which would implicitly define the aggregate market production. As in well-known examples of CES and Kimball demand systems, we also assume homogeneity, so that individual good quantities are proportional to the aggregate consumption. From here on out we consider all the consumer problems in terms of consumption shares $\frac{y_{\omega jt}}{Y_{jt}}$. We additionally require that the function F satisfies several assumptions common to the demand functions used in the macroeconomic literature, and impose several additional assumptions to ensure the invertibility and the global existence of the solution.

Definition 1 *Function $F\left(\frac{y_1}{Y}...\frac{y_\omega}{Y}...\frac{y_n}{Y}\right)$ governs the substitution between individual firm goods within the market. The function is twice continuously differentiable. It satisfies the non-satiation and decreasing marginal utility assumptions, so that the function is concave with $F'_\omega > 0, \forall\omega$, and a negative-definite second derivative matrix $\{F''_{\omega\omega}\} < 0$. The function follows Inada conditions, meaning that $\forall\omega$ $F'_\omega \to +\infty$ as $\frac{y_{\omega jt}}{Y_{jt} \to 0}$.*

Another assumption commonly imposed in the theoretical literature is symmetry, meaning that the value of the function remains the same for any permutation of its arguments. This assumption would be needed to ensure that function F does not create any a priori heterogeneity between firms. In our setting, this assumption is not imposed unless the opposite is specified.

Consumers minimize their costs across individual firms and markets. Due to the two-level nested structure, consumer problems within a market and between markets are separable and can be solved independently.

We first consider a within-market cost minimization problem.

$$P_j = \min_{\frac{y_{\omega j}}{Y_j}} \sum_{\omega=1}^{n_j} p_{\omega j} \frac{y_{\omega j}}{Y_j}$$
$$\text{s.t.} \quad F\left(\frac{\nu_{1j} y_{1j}}{Y_j}, .. \frac{\nu_{\omega j} y_{\omega j}}{Y_j}, .. \frac{\nu_{n_j j} y_{n_j j}}{Y_j}\right) = 1$$

Where $y_{\omega j}$ is the quantity of the good from firm ω operating in market j, $\nu_{\omega j}$ is the demand shock, Y_j is the aggregate quantity of the market j, $p_{\omega j}$ is the price of ω in market j, and P_j is the aggregate price of market j.

This problem is static so the time subscript t is dropped, when applicable. Note that this problem can be adjusted to become a minimization of a convex function on a convex set, which would guarantee that the first-order conditions give a point of minimum. Thus, the demands for individual firm goods are going to be the solution of the following system.

$$\forall \omega : \quad \left\{ F'_\omega \left(\frac{\nu_{1j} y_{1j}}{Y_j} .. \frac{\nu_{\omega j} y_{\omega j}}{Y_j} .. \frac{\nu_{n_j j} y_{n_j j}}{Y_j} \right) = \frac{p_{\omega j}}{P_j D_j \nu_{\omega j}} \right.$$
$$\text{where } D_j = \left(\sum_\omega F'_\omega \frac{\nu_{\omega j} y_{\omega j}}{Y_j} \right)^{-1}$$
$$P_j = \sum_\omega \frac{y_{\omega j}}{Y_j} p_{\omega j}$$
$$F\left(\frac{\nu_{1j} y_{1j}}{Y_j} .. \frac{\nu_{\omega j} y_{\omega j}}{Y_j} .. \frac{\nu_{n_j j} y_{n_j j}}{Y_j}\right) = 1$$

The solution of this system of equations gives the demand for each of the individual firm's goods as a fraction of total consumption in the market. Two important aggregates arise here. First, there is the aggregate market price, P_{jt}, defined using the natural pricing assumption. The second aggregate is the price dispersion D_{jt}. Price dispersion in our setting is similar to the version seen in Kimball-type demand aggregators. In the case when the aggregator F is a function that is homogeneous to degree 1, this price dispersion term would be permanently pegged at unity.

Proposition 1 *Given the definition of the function F, this system of non-linear equations has a unique solution and this solution is a minimum of the cost minimization problem.*

As discussed before, the second part of the proposition follows from the fact that the problem can be adjusted to minimize a convex function on a convex set. The first part of the proposition comes directly from the results in Gale and Nikaidô (1965). There would be a global inverse function if the Hessian matrix of the second derivatives of the function F were a negative semi-definite matrix. Even though in this paper we mostly rely on the restrictive assumption of negative-definiteness, this assumption can be relaxed to include P-matrices and their subclass of dominant diagonal matrices, as discussed in Gale and Nikaidô (1965) and Mckenzie (1960), and as shown even further by Berry et al. (2013).

From here on, we define the solutions of this system as:

$$y_\omega = \frac{1}{\nu_{\omega j}} \Upsilon^{\omega j} \left(\frac{p_{1j}}{P_j D_j \nu_{1j}} \cdots \frac{p_{n_j j}}{P_j D_j \nu_{n_j j}} \right) Y_j$$

We now turn to the market-level demand and solve the second half of the consumer problem. As noted, we assume that the market-level demand follows a Kimball aggregator.

$$Y_j = \Upsilon^j \left(\frac{P_j}{PD} \right) Y$$

Combining the structures for firm-level and market-level demand, we can obtain the total demand function that firms would face.

$$y_{\omega j} = \frac{1}{\nu_{\omega j}} \Upsilon^{\omega j} \left(\frac{p_{1j}}{P_j D_j \nu_{1j}} \cdots \frac{p_{n_j j}}{P_j D_j \nu_{n_j j}} \right) \Upsilon^j \left(\frac{P_j}{PD} \right) Y$$

Note that this demand structure is a straightforward generalization of the Atkeson-Burstein oligopoly. This system additionally covers the Kimball oligopoly in Wang and Werning (2020) and the original Kimball monopolistic competition in Kimball (1995). The demand system would turn into a Kimball oligopoly when there is a separability between $\frac{y_{\omega j}}{Y_j}$, meaning that the cross-derivatives $F''_{\omega k}, \omega \neq k$ are equal to zero. The demand would be similar to Atkeson-Burstein oligopoly if the F function were a sum of power functions: $F = \sum_\omega \left(\frac{y_{\omega j}}{Y_j} \right)^{\sigma - 1}$, similar to CES.

2.2 Decomposition

The next step is to find a decomposition for the firm-level demand parameters that are important for price-setting. We are going to start with the decomposition of the elasticity and proceed to talk about a similar decomposition for higher-order terms.

The general intuition regarding firm-level elasticity stems from the oligopolistic structure of the firm-level demand. For small firms with market shares close to zero, elasticity would be similar to elasticity of the firm-level demand, with the aggregates almost independent on their prices. For large firms with shares close to unity, the market price will be almost similar to their own price, so their demand elasticity would be similar to the elasticity of the market-level demand. This difference arises from the two types of firms' differential impact on the aggregates. We see a similar pattern when we write down the equation for the elasticity of demand: there are parameters that depend on the firm-level demand function and parameters that are determined by how much an individual firm impacts the aggregates.

$$\xi_{\omega\omega j} = \underbrace{-H^{-1}_{\omega\omega j} \frac{Y}{y_{\omega j} \nu_{\omega j}} F'_{y_{\omega j}}} + \sum_k \underbrace{H^{-1}_{\omega k j} \frac{Y_j}{y_{\omega j} \nu_{\omega j}} F'_{y_{kj}}} \frac{\partial P_j D_j}{\partial p_{\omega j}} \frac{p_{\omega j}}{P_j D_j} - \frac{\partial Y_j}{\partial p_{\omega j}} \frac{p_{\omega j}}{Y_j}$$

Where $\xi_{\omega\omega}$ is the elasticity of firm ω to its own price, H is the Hessian of the function F, and $H^{-1}_{\omega k}$ is the

ωk-th element of the inverse of the Hessian matrix of the aggregator function F.

Elements of the inverse Hessian are objects of the function F of the firm-market level substitution, as opposed to the partial derivatives of P, D and Y. Capturing this distinction would allow for decomposition of the elasticity and determination of the set of sufficient statistics, which would express the impact of the firm on the market aggregates in terms of objects of the function F. To do so, we introduce the concept of "partial elasticity."

Partial elasticity is the elasticity conditional on keeping the aggregate market parameters fixed. It would allow to separate the direct impact of the firm's price on its quantity and the impact of the firm's price on the market aggregates. Intuitively, it will only be an important concept for oligopolistic markets, where firms are large enough to impact the market aggregates. In the case of monopolistic competition, partial elasticity would coincide with elasticity.

Definition 2 *Let partial elasticity be the demand elasticity, given the aggregate values of the P_j, D_j, Y_j:*

$$\xi^r_{\omega kj} = -\frac{\partial y_{\omega j}}{\partial p_{kj}} \frac{p_{kj}}{y_{\omega j}}\bigg|_{Y_j, D_j, P_j} = -H^{-1}_{\omega kj} \frac{Y_j}{y_{\omega j}\nu_{\omega j}} \frac{p_{kj}}{P_j D_j \nu_{kj}} = -H^{-1}_{\omega kj} \frac{Y_j}{y_{\omega j}\nu_{\omega j}} F'_k$$

To illustrate the concept of partial elasticity, we consider the simpler case of an Atkeson-Burstein oligopoly. In this case, both the firm-level and market-level demand aggregators are CES and the demand function for an individual firm good is given by:

$$y_{\omega jt} = \left(\frac{p_{\omega jt}}{P_{jt}}\right)^{-\sigma} \left(\frac{P_{jt}}{P_t}\right)^{-\theta} Y_t$$

Note that the price dispersion term D would in this case always be equal to unity, so it is omitted from the demand equation. Partial elasticity in this case would be given by: $\xi^r_{\omega\omega} = \sigma$ and $\xi^r_{\omega k} = 0, \omega \neq k$.

To further note the distinction, consider the difference between partial elasticity $\xi^r_{\omega\omega} = \sigma$ and elasticity $\xi_{\omega\omega} = \sigma + (\theta - \sigma)\lambda_\omega$, with λ_ω being the revenue market share. Moreover, consider the distinction between partial cross-elasticity $\xi^r_{\omega k} = 0$ and cross-elasticity $\xi_{\omega k} = (\theta - \sigma)\lambda_k$. In all those cases, the part that depends on the market shares comes from the impact of the individual firms on the market aggregates. Similarly, in Kimball oligopoly models, the partial cross-elasticity would be equal to zero but the cross-elasticity itself would not equal zero due to the impact on the markets aggregates.

To finish the decomposition of the elasticity, we need to consider the last important parameter — market-level elasticity. Market elasticity is the elasticity of the market-level demand to the market-level price. It is determined by the parameters of the market-level aggregator.

Definition 3 *Let market elasticity be the market-level demand elasticity to the market-level price:*

$$\xi^m_j = -\frac{\partial Y_j}{\partial P_j} \frac{P_j}{Y_j}$$

The possibility of introducing such a definition is another consequence of the nested demand structure: individual firm prices only impact market demand through the market price. Effectively, we are treating the market demand separately from any composition effects arising from changing market share distributions across firms within the market.

Coming back to the example of the Atkeson-Burstein oligopoly, consider the market-level elasticity in this case: $\xi_j^m = \theta$.

Recall the elasticity formula in Atkeson-Burstein. Using our definition of partial elasticity and market elasticity, we obtain this well-known formula:

$$\xi_{\omega\omega j} = \xi_{\omega\omega j}^r + (\xi_j^m - \xi_{\omega\omega j}^r)\lambda_\omega$$

Note that if we now wished to calculate the elasticity for each firm in the market, we would only need to know two constants: $\xi_{\omega\omega j}^r$ and ξ_j^m. Those two parameters therefore form a sufficient statistic for firm-level elasticity.

Even though partial elasticities will no longer be constants, a similar decomposition for elasticity can be written in the general case, where partial elasticities and market-level elasticities are sufficient statistics to measure the elasticity of a firm.

Simplifying the equation for elasticity and substituting in the dependence of the market aggregates on the individual firm price, we obtain:

$$\xi_{\omega\omega} = \xi_{\omega\omega}^r + \left(\xi^m - \sum_k \xi_{\omega k}^r \frac{\sum_k \xi_{k\omega}^r \lambda_k}{\sum_k \sum_m \xi_{km}^r \lambda_k \lambda_\omega}\right)\lambda_\omega$$

Where $\xi_{\omega\omega}$ is the own price elasticity, $\xi_{\omega k}^r$ is the partial elasticity of demand for the good from firm ω with respect to the firm price k, ξ^m is the elasticity of the market demand for the market price, and λ_k is the revenue market share of the firm k. Apart from these elementary "building blocks," there are also several important aggregates. The first one is $\sum_k \xi_{\omega k}^r$. It reflects the total effect coming from the change of aggregate price and price dispersion. Since the change of market aggregates affects all the relative prices, it reflects data from the focal firm itself and all its competitors. The second aggregate is $\sum_k \xi_{k\omega}^r \lambda_k$, which reflects the market share-weighted effect of the own price on competitor demand. This is an important parameter for the dynamic of price dispersion. The last one is $\sum_k \sum_m \xi_{km}^r \lambda_k$, which can be interpreted as aggregate market elasticity. The fraction as a whole, $\theta_\omega = \frac{\sum_k \xi_{k\omega}^r \lambda_k}{\sum_k \sum_m \xi_{km}^r \lambda_k \lambda_\omega}\lambda_\omega$, corresponds to the response of the market aggregate price and price dispersion to the individual firm price: $\frac{\partial PD}{\partial p_\omega}\frac{p_\omega}{PD}$

The equation of elasticity that we derive in our model would be simplified into Atkeson-Burstein oligopoly demand elasticity if we make the corresponding assumptions. When the demand system is an Atkeson-Burstein oligopoly, all the partial cross-elasticities become zero and own price partial elasticity is constant across firms. Rewriting the equation for elasticity in accordance with these assumptions would get us back to the familiar elasticity formula.

Note that in some cases, the elasticity for a firm with close to zero market share would be different from the own price partial elasticity. This effect emerges from the non-zero partial cross-elasticity for the zero share. In this case, even the smallest firms have a significant effect on other firms' quantities. Intuitively, this would mean that the mere availability of the option on the market makes a difference for the demand of other firms. However, as we will see, this is not the most empirically relevant case. That title belongs to Kimball oligopoly, the case of zero partial cross-elasticity and heterogeneous elasticities across firms.

In this case, elasticity would be represented by:

$$\xi_{\omega\omega} = \xi^r_{\omega\omega} + \left(\xi^m - \xi^r_{\omega\omega}\frac{\xi^r_{\omega\omega}}{E^r}\right)\lambda_\omega$$
$$\text{where } E^r = \sum_k \xi^r_{kk}\lambda_k$$

where $\xi^r_{\omega\omega}$ is the own price partial elasticity, E^r is the aggregate market elasticity, ξ^m is the market-level elasticity, and λ_ω is the firm revenue share.

Once we have a decomposition for the elasticity, we can apply similar logic to obtain a decomposition for the higher-order terms, most importantly superelasticity. Superelasticities of the own price elasticity, with respect to the own price, are the main parameters that determine the variation of elasticity. This makes superelasticity an important parameter for the variability of markups and, consequently, for the pass-through of marginal costs into prices. By definition, superelasticity is the elasticity of the demand elasticity.

$$\eta_{\omega km} = \frac{\partial\xi_{\omega k}}{\partial p_m}\frac{p_m}{\xi_{\omega k}}$$

Where $\xi_{\omega k}$ is the cross-elasticity of demand for the good from firm ω with respect to the firm price k, p_m is the price of firm m, and $\eta_{\omega km}$ is the elasticity of the cross-elasticity of product ω to price k, with respect to the price of firm m. Note that we have a triple subscript because superelasticity is connected to the second derivative of demand for a particular good. When appropriate, such as in the case of own superelasticity, we omit the triple subscript for notation simplicity.

Similarly to elasticities, superelasticities can be decomposed into the direct effects of prices and the indirect effects from their impact on the market aggregates. To obtain such a decomposition, we first need to define a partial superelasticity. Similarly to partial elasticity, we define partial superelasticity as the elasticity of the partial elasticity given the market aggregates.

Definition 4 *Let partial superelasticity be the elasticity of the partial elasticity given the market aggregates.*

$$\eta^r_{\omega kmj} = \frac{\partial\xi^r_{\omega kj}}{\partial p_{mj}}\frac{p_{mj}}{\xi^r_{\omega kj}}\bigg|_{P_j,D_j,Y_j}$$

Proposition 2 *Superelasticity can be decomposed into direct and market aggregate effects, and this decomposition can be expressed in terms of partial elasticities, partial superelasticities, market elasticities, market* *superelasticities, and revenue shares.*

The general formula for the superelasticity decomposition can be found in the appendix. Here we include the decomposition for a Kimball oligopoly, which as noted above is the most empirically relevant case. Superelasticity decomposition would be given by:

$$\eta_{\omega\omega} = \frac{\xi^r_\omega}{\xi_\omega}\eta^r_\omega(1-\theta_\omega) + \frac{\xi^m\lambda_\omega}{\xi_{\omega\omega}}(\eta^m\lambda_\omega + \Lambda_{\omega\omega}) - \frac{(\xi^r_\omega)^2\lambda_\omega}{\xi_{\omega\omega}E^r}(2\eta^r_\omega(1-\theta_\omega) + \Lambda_{\omega\omega} - N^r_\omega)$$

$$\text{where } E^r = \sum_k \xi^r_k\lambda_k \qquad \theta_\omega = \frac{\xi^r_\omega\lambda_\omega}{E^r}$$

$$\Lambda_{k\omega} = \frac{\partial\lambda_k}{\partial p_\omega} = \mathbb{I}\{k=\omega\} - \xi_{k\omega} - (1-\xi^m)\lambda_\omega$$

$$N^r_\omega = \frac{\partial E^r}{\partial p_\omega}\frac{p_\omega}{E^r} = \sum_k \frac{\xi^r_k\lambda_k}{E^r}\Lambda_{k\omega}$$

Note several important aggregates. As before, we have aggregate elasticity E^r, as well as the impact of price on the aggregate price and price dispersion, θ_ω. New important parameters are the response of market share to the change of individual firm price, $\Lambda_{k\omega}$, and the superelasticity of the market level elasticity, N^r_ω

The most important takeaway from this subsection is that the set of partial elasticities and revenue shares is a sufficient statistic for calculating the total oligopolistic firm elasticity. This is an important building block for our empirical strategy, since we will be able to estimate a limited set of parameters and control well for the effects coming from changes in the competitors' prices, which is a common concern in demand estimations. It is especially important given the oligopolistic structure, where the impact of the firm of interest on other firms' demands, and an additional effect on its own demand via the market aggregates, cannot be ignored.

2.3 Optimal Flexible Prices

In this subsection we consider optimal price setting in a flexible-price economy. This is an important benchmark to discuss, since it will determine the dynamic of desired prices for firms that are constrained in their price setting. If the desired flexible price is changing rapidly, we might expect more response in a sticky price economy. The flexible-price economy benchmark would also determine the desired inflation benchmark, so it is a useful tool to start a conversation about inflationary pressures in the economy.

We consider a simple firm problem and derive the price setting equation. We prove that if we abstract from the possibility of collusion, under the assumptions that are true for many demand structures, this system of equations will have a unique solution.

Consider the usual partial-equilibrium price-setting problem. Given the aggregate wage W_t, quantities

Y_t, and prices P_t, the firm maximizes instantaneous profits.

$$\begin{aligned} &\max_{p_{\omega jt}} p_{\omega jt} y_{\omega jt} - W_t l_{\omega jt} \\ &\text{s.t. } y_{\omega jt} = \Upsilon^{\omega j}\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right)\Upsilon^{j}\left(\frac{P_{jt}}{P_t D_t}\right) Y_t \\ &\qquad y_{\omega jt} = a_{\omega j} l_{\omega jt} \end{aligned}$$

We have two sources of heterogeneity. First, we allow for the non-symmetric aggregator function F, interpreting this as a long-term difference in the appeal of the firm's good. Secondly, we include differential fixed productivity $a_{\omega j}$. Intuitively, this could act as a placeholder for differential levels of capital for large and small firms. The optimal flexible price, as in many other models, would be given by:

$$p^{*}_{\omega jt} = \frac{\xi_{\omega\omega jt}}{\xi_{\omega\omega jt} - 1} mc_{\omega jt}$$

This system of equations would determine flexible price partial equilibrium.

Proposition 3 *Flexible price equilibrium exists and is unique when*

1. *The demand aggregator satisfies the conditions of definition 1 and is homogeneous of degree 0 with respect to prices; own price elasticity is positive and other goods in the same market are substitutes, so that all the cross-superelasticities are negative.*

2. *The demand aggregator is Kimball.*

Note that this equilibrium uniqueness result does not exclude the possibility of extra collusive equilibrium, when firms can decide all the prices collectively and act as a cartel. In this case, they are able to internalize the externalities caused to each of them by competition from the others, and thus do not follow the same pricing equation.

The uniqueness of the flexible price equilibrium means that prices and market shares are uniquely determined by firm-level characteristics, such as fixed firm-level productivity or the exogenous asymmetries of the aggregation function F. This means that unless there is a change to those individual firm characteristics, the relative prices would return to the same equilibrium. For instance, if there is an aggregate shock, it will not impact individual firms' price-setting behavior in the flexible price case. Under sticky prices, this would mean that relative prices would converge back to the same levels.

The uniqueness of the flexible price equilibrium does not mean that this economy is perpetually locked in zero inflation. While relative prices stay fixed, prices would still change, even though their change does not impact real variables in such an economy. The degree of the price change will, however, be important if we keep in mind that this economy is a benchmark and the desired changes in prices in it still correspond to the desired changes in prices in a sticky price economy. For instance, if we find that prices in this flexible price economy increase sharply, this would mean that the prices in the more realistic sticky price economy are subject to significant inflationary pressures. This makes the flexible price economy a useful tool for understanding the dynamics of inflation, without, of course, saying anything about its possible real effects.

The result of the proposition also means that given the demand function, all the firm-level characteristics, including partial elasticities, can be uniquely determined from the firm-level shares. In the case with non-zero partial cross-elasticities, it would be important to know the shares of the firm itself and all its competitors; however, in the Kimball oligopoly case, it would be enough to only know the market share of the firm itself.

Proposition 4 *In the Kimball oligopoly case without a priori asymmetry, the partial elasticity of a firm with the same "adjusted" output share* $\frac{y_{\omega j}\nu_{\omega j}}{Y_j}$ *is the same in any market.*

This corollary means that in the Kimball oligopoly case with no asymmetry, it is sufficient to know the output share of a firm to calculate the partial elasticity, and that all the dependence on the share distribution is contained in the response of the market aggregates. In an empirical setting, this would mean that once we observe output shares, we can estimate partial elasticities independently of the market structure. This would make our estimates for the partial elasticity usable for any market, with any distribution of market shares across firms.

In reality, however, one more step is required, since "adjusted" output shares $\frac{y_{\omega j}}{Y_j}$ are difficult to observe. Usually we can only observe revenue market shares $\lambda_{\omega j} = \frac{p_{\omega j} y_{\omega j}}{P_j Y_j} = \frac{y_{\omega j}\nu_{\omega j}}{Y_j} F'_\omega\left(\frac{y_{\omega j}\nu_{\omega j}}{Y_j}\right) D_j$ that are dependent on the market aggregates. This means that in order to get estimates for partial elasticity that are independent of the market structure, we would also need to control for the market aggregates.

It is additionally worth noting that the no-asymmetry assumption can be somewhat relaxed in the applications while maintaining the corollary. We can allow for limited asymmetry across firms, while maintaining symmetry between markets. The only requirement in this case is that any a priori asymmetry across firms is observable and can be directly accounted for. For instance, in one of the robustness checks, we additionally allow elasticity to depend on the several lags of the prices, in the spirit of customer acquisition models, such as Ravn et al. (2006). As long as all the sources of apriori asymmetry are accounted for, the result of the corollary still holds.

The proposition and the corollary of this section are important for our empirical strategy. As long as there is no significant change in the distribution of the market shares across firms, and no collusion among firms, markets fluctuate around the same steady state with similar elasticities. Additionally, given market aggregates, partial elasticities for firms with similar revenue market shares can vary across markets, as they depend only on the market aggregates. This means that partial elasticity estimates that control for the revenue market share and the market aggregates are independent from the market structure, so they can be used to reconstruct firm-level elasticity in a market with any share distribution. We get this kind of partial elasticity estimates from a regression with fixed effects in the market and market share firm groups.

2.4 Pass-through and Asymmetric Effect

In this subsection we are going to derive the important demand parameters that influence the dynamics of the desired markup and hence the desired price level. We will start with the linear approximation and discuss the linear pass-through of marginal costs into prices, and then consider the non-linear effect on marginal cost and derive the elasticity of pass-through to a marginal cost shock.

Consider the log-linearization of the equation for the optimal flexible price. This will effectively give the linear approximation of the optimal response function for a given firm's own marginal costs, prices of the competitors, and the economy aggregates.

$$\tilde{p}^*_{\omega jt} = \underbrace{\frac{\xi_{\omega\omega jt} - 1}{\xi_{\omega\omega jt} - 1 + \eta_{\omega\omega\omega jt}}}_{\text{pass-through}} \tilde{mc}_{\omega jt} + \sum_{k \neq \omega} \underbrace{\frac{-\eta_{\omega\omega kjt}}{\xi_{\omega\omega jt} - 1 + \eta_{\omega\omega\omega jt}}}_{\text{strategic complementarity}} \tilde{p}_k + \underbrace{\frac{\eta_{\omega\omega mjt}}{\xi_{\omega\omega jt} - 1 + \eta_{\omega\omega\omega jt}}}_{\text{market strategic complementarity}} (\tilde{P}_t + \tilde{D}_t)$$

Where $\xi_{\omega\omega jt}$ is the own price elasticity, $\eta_{\omega\omega\omega jt}$ is the own price superelasticity, $\eta_{\omega\omega kjt}$ is the superelasticity with respect to the price of the competitor k, and $\eta_{\omega\omega mjt}$ is the superelasticity with respect to the economy aggregate price and price dispersion.

By definition, superelasticity is the elasticity of the elasticity. It is an important parameter, since it determines how elasticity changes with prices and, consequently, how markups change with prices. Own price elasticity and own price superelasticity are the only two numbers that we need to know in order to calculate the pass-through of marginal costs into prices. Note that if superelasticity is zero, there is a perfect pass-through and firms respond one-to-one to the changes in their marginal costs. This would be a case of CES demand.

There are two types of strategic complementarity. The first one comes from the effects of competitor prices from the same market. The degree of this strategic complementarity depends on the responsiveness of the elasticity for the firm of interest to the prices of a particular competitor. It can be heterogeneous across firms; for instance, one would expect larger firms to have a larger impact than small firms. The second type comes from the effects of the economy aggregate price. This term arises because of the assumption of the Kimball aggregator on the market level. The impact of the aggregate price would depend on the superelasticity of the market-level demand and on the size of the firm of interest.

Note additionally that in homogeneous demand cases, such as the one considered here, the coefficients in the first-order decomposition sum up to one. Intuitively, this means that if the firm's marginal cost, all the competitors' costs, and the total economy price adjust similarly, it will be optimal for the firm to also adjust fully.

Pass-through will be different for different firms and for different equilibria. The pass-through is smaller for larger firms, so that they are less responsive to changes in marginal costs. This means that when the role of large firms in the economy is larger, the aggregate pass-through will decrease. In turn, lower pass-through would mean smaller adjustment of the desired price and, if we abstract from the non-linearities, a smaller response of inflation. This would explain some puzzles such as missing deflation and be consistent with other commonly cited facts, such as the flattening of the Phillips curve studied by Baqaee et al. (2021).

Our rich structure allows us to hint at another effect. If the share of large firms increases, it might not only impact the large firms themselves, but also cause not-so-large firms in the same market to have a lower pass-through, a result appearing in oligopolistic settings with non-constant partial elasticity. This means that not only will aggregate pass-through be lower, but each firm's individual pass-through as well. This will further amplify the effects.

Proposition 5 *When elasticity is a convex function of prices, pass-through is smaller for larger firms.*

Studying linear approximation is useful, but it could lead us to miss important non-linearities. For instance, it makes the effect of marginal cost on prices seem symmetric and uniform across different sizes of marginal cost shocks, which might not be the case. To study this non-linearity, we take a closer look at the elasticity of pass-though with respect to a marginal cost shock. This will reveal how the pass-through changes depending on the sign and magnitude of the marginal cost shock.

$$\frac{\partial \log [\text{pass-through}]}{\partial \log mc_\omega} = \left(\frac{\eta_\omega}{\xi_\omega - 1}\right)\left(\eta_\omega \frac{\xi_\omega}{\xi_\omega - 1} - \psi_\omega\right)\left(\frac{\xi_\omega - 1}{\xi_\omega - 1 + \eta_\omega}\right)^2$$

Where ξ_ω is elasticity, η_ω is superelasticity, and $\psi_\omega = \frac{\partial \eta_\omega}{\partial p_\omega}\frac{p_\omega}{\eta_\omega}$ is the elasticity of the superelasticity.

In many models, this second-order coefficient will be positive, such as in Kimball monopolistic competition with constant superelasticity or in an Atkeson-Burstein oligopoly. Note additionally that in the flexible price equilibrium, if we consider one effect of the common marginal cost shock, all the second-order terms will sum up to zero, ensuring a total pass-through of unity.

Proposition 6 *When elasticity is a convex function of prices, the pass-through of marginal costs into prices is asymmetric and is larger for marginal cost increases.*

This result suggests that positive marginal cost shocks have a larger pass-through when superelasticity depends negatively on prices. A positive second-order effect means that the response of prices to the marginal cost shock is asymmetric, such that prices are more responsive to a marginal cost increase than to a marginal cost decrease. Moreover, the non-linear effect is more pronounced for larger cost increases. This means that even though the linear pass-through for a more concentrated aggregate economy might be lower, the positive non-linear effect would mean that a large positive marginal cost shock can still have a high pass-through and create significant inflationary pressures.

3 Empirical Strategy and Results

In this section we discuss the data used, the empirical strategy, the construction of our instrument, and the results of the estimation.

3.1 Data

Our estimation requires firm-level data on prices and quantities, which makes the ACNielsen Retail Scanner Database a good option. This database collects weekly prices, sales, and barcodes of participating retail stores across all US markets. The data contains barcode-level product prices and quantities, recorded weekly from about 35,000 participating grocery, drug, mass merchandise, convenience, and liquor stores. It covers more than half of the total sales of US grocery and drug stores and more than 30% of all US mass merchandise. The total size of the dataset is over 1300 GB.

Our theoretical framework first requires defining a "firm" and a "market." First, we define the market as a Nielsen product category. Our dataset has approximately 1100 product categories. A single product category unites similar goods, so "canned fruit-grapefruit" and "canned fruit-oranges" are two different product categories. Likewise, "toaster and toaster oven appliance" and "microwave appliance" are also two separate product categories.

Defining a firm in this context presents a bit more of a technical challenge. Unlike in our theoretical framework, in the data firms might produce multiple different goods or even operate in different markets. We define a firm as a single producer, that, according to the barcode assignment rule, should be coded by the first several numbers of the code. We then artificially construct a firm composite product for each of the markets. We are able to do this by invoking the homogeneity assumption. In the case of homogeneous demand, aggregate price changes would be given by a weighted average of the individual changes, with the weights equal to the market revenue shares of individual products. We compute a quantity composite as the difference between the aggregate revenue and the aggregate price.

Due to the size of the dataset, working with it also presents several technical obstacles. The raw version of the dataset reports weekly sales across individual stores. For the baseline estimation, we collapse the dataset to the firm-market-time level. This aggregation procedure is likewise based on the assumption of homogeneity of demand. We additionally limit ourselves to the years 2007 to 2015. For the collapsed firm-market-time and market-time level datasets, we use all the data available.

The instruments that we construct use regional variation. For the instrument to be constructed properly, a firm must operate in multiple regions. We define a single region as a FIPS county, and we limit ourselves to the firms that operate in over 10 regions. To maintain comparability and since we are primarily interested in oligopolistic firms, we additionally exclude all firms with revenue shares below 0.1% of the corresponding market. When we later turn to analyzing the aggregate market dynamics, inclusion of those firms does not influence the results.

3.2 Granular Instrument

The estimation of demand presents a classic example of simultaneous equations, meaning we need to employ an instrumental variable approach.

Our baseline instrument builds on the ideas of the National Price Instrument Hausman (1996) sometimes used in the IO literature, for instance in Nevo (2001) and DellaVigna and Gentzkow (2019), and on the Granular Instrument Approach by Gabaix and Koijen (2020). We make use of the fact that firms set their prices at the national level and are usually unable to pick prices for each specific region. At the same time, individual regions are large enough that regional firm-level shocks are significant enough to influence national firm-level prices.

The first stage for constructing the granular instrument is to separate local supply shocks. The main identifying assumption for our baseline instrument is that demand shocks are more aggregated than supply shocks. We assume that demand is determined at the firm-market-time level or the region-market-time level, and that there are no specific region-firm-market-time shocks. This assumption seems plausible in our case, since all the goods that we consider are nationally traded and hence we might expect all the product characteristics and advertising campaigns to be decided nationally. At the same time, there are region-firm-market-time supply shocks. While the production of goods in our sample is conducted at the national level, there is still a production-like step involved in getting the goods from the factory to the final consumer. There are different cost shocks coming from changing transport costs or changing local wages, which affect (for instance) drivers and retail workers. Another example of a local shock that would be included in the instrument is a change in the price at a particular supermarket chain due to a temporary sale. As discussed in DellaVigna and Gentzkow (2019), such variation, after controlling for seasonality, is unlikely to be related to the local demand. If both assumptions are fulfilled, we can extract supply shocks by taking the residual from the following regression:

$$\tilde{y}_{r\omega jt} = \gamma_{\omega jt} + \gamma_{rjt} + \epsilon_{r\omega jt}$$

Where $\tilde{y}_{r\omega jt}$ is the change in the quantity sold by firm ω in market j in the region r, $\gamma_{\omega jt}$ is the fixed effect at the firm-market-time level, and γ_{rjt} is the fixed effect at the region-market-time level. Note that those two sets of fixed effects effectively take care of all the firm-level demand shocks and all the region-market-level demand shocks. Firm-market-time-level effects capture, for instance, increased national appeal of the firm, such as one resulting from a national advertising campaign, or an increase in a good's quality that makes it more valuable to the consumer. Regional-market-time-level effects capture, for instance, increases in aggregate demand within the region, such as the presence of a booming local industry, or a temporary increase in the appeal of a particular type of good in a particular region, such as an increased demand for ice cream in California during summer. These effects could also arise from the closure of an important local firm in a corresponding market, such as a large California-only ice cream producer going bust.

Due to these differences, all constant-over-time regional differences in demand for different goods or even different firm goods would be canceled out, such as constant higher demand for Ben&Jerry's ice cream in California. The identification threat for this strategy is the presence of large and important, non-constant over time, firm-region-market-level demand shocks, like a local advertising campaign for Ben&Jerry's ice cream in California.

The second stage is to use the residuals to construct a granular instrument at the firm-market-time level by creating a weighted sum of all the regions, except for the region of interest. The weights are lagged revenue shares.

$$z_{\omega jt} = \sum_{r} \lambda_{r\omega jt-1} \epsilon_{r\omega jt}$$

Where r is the region, ω is the firm, j is the market, t is time, $z_{\omega jt}$ is the constructed firm-level granular instrument, $\lambda_{r\omega jt-1}$ is the revenue share of region r within the firm ω, and $\epsilon_{r\omega jt}$ is the residual from the fixed effects regression discussed above.

The main identifying assumption on regional firm-market-time shocks can be relaxed if we construct a granular instrument at the regional level. This would require working with a subsample of the data and will map less clearly onto our theoretical framework, so we address it as a robustness check. The new assumption for the regional-level shock is that region-firm-market-level demand shocks do not spillover across regions—this would mean, for instance, that there is no demand increase so great that consumers are willing to travel to another region to buy the product. We perform this robustness check and find similar results.

For a better illustration of what kinds of shocks enter the granular instrument and why it makes sense to formulate it in this way, consider a simple extension of the model. Say a firm makes all pricing and production decisions on the aggregate level, but then needs to supply the goods locally. We would consider two types of shocks here: iceberg transport shocks and demand shocks.

Iceberg transport costs are commonly found in the trade literature and follow a classic paper by Samuelson (1952). In this case, it is assumed that for each unit of goods that leaves the factory, only a fraction is able to reach the destination and be paid for. This shock can be interpreted in three ways. First, it can be seen as a share of product lost on the way, for instance due to lost or damaged packages or goods expiring before they are sold. Second, it can be interpreted as supply chain disruptions and delays, e.g., from transport companies being overwhelmed or going out of business. The delays are especially important in our setting, since Nielsen mostly contains information on short-lived consumer goods that are produced and supplied to stores in a continuous flow. Delays would cause products to be stuck in the warehouses for longer, which incurs losses for the company since the goods are produced but not paid for. Lastly, iceberg cost shocks can be related to decisions made by local retailers, such as reducing store hours due to high labor costs in the region. The important property of the iceberg cost shock is that it influences quantities in different regions directly.

Demand shocks are similar to the ones commonly found in the literature. They capture an increase in the appeal of a particular good, which makes consumers more willing to buy it. We specifically consider three types of demand shocks, which mirror the fixed effects of the regressions for the granular instrument construction discussed above. The first one is a common regional shock that affects all the firms operating in this region. The second one is a firm-level demand shock, reflecting the general appeal of the firm across regions. The last one represents a constant-over-time consumer preference for a certain firm in a particular region.

In this case, the firm solves:

$$\begin{aligned} max_{p_{\omega jt}} \quad & \sum_r p_{r\omega jt} y_{r\omega jt} - W_t L_{\omega jt} \\ \text{s.t.} \quad & y_{r\omega jt} = \delta_{r\omega jt} \nu_{r\omega} \nu_{rjt} \nu_{\omega jt} \left(\frac{p_{r\omega jt}}{p_\omega} \right)^{-\sigma} y_{\omega jt} \\ & p_{r\omega jt} = p_{\omega jt} \\ & V_{\omega jt} y_{\omega jt} = A_{\omega jt} L_{\omega jt} \\ & V_{\omega jt} y_{\omega jt} = \sum_r \frac{p_{r\omega t}}{p_{\omega jt}} \frac{y_{r\omega jt}}{\delta_{r\omega jt}} \end{aligned}$$

Where $\delta_{r\omega jt}$ is the region-firm-market-time iceberg cost shock, $\nu_{r\omega}$ is the constant region-firm taste, ν_{rjt} is the time-varying region-market-level common demand shock, and $\nu_{\omega jt}$ is the time-varying firm-market-level demand shock. There is an additional term, $V_{\omega jt} = \sum_r \nu_{r\omega} \nu_{rjt} \nu_{\omega jt}$, since in the case of granularity, local demand shocks do not sum up to a unity.

Note that we assume a lack of specific firm-level demand shocks within the region. As discussed above, the presence of such shocks would be a threat to identification. Note moreover that, as we will explain below, the possible threats in region-firm-market-level demand shocks do not include transmission through the national price, since up to the first order of approximation, it does not depend on the demand shocks. Due to national pricing, this will remain true even if we relax the CES demand assumption on the regional level.

Note particularly the last condition, which is similar to the common natural pricing assumption but has an unusual term: $\frac{1}{\delta_{r\omega jt}}$. This term reflects the fact that a firm is only able to charge for the goods that do get delivered.

The solution for the optimal national price is given by:

$$p^*_{\omega jt} = \mu_{\omega jt} \frac{1}{\sum_r \frac{\nu_{rjt} \nu_{r\omega} \nu_{\omega jt}}{\sum_r \nu_{rjt} \nu_{\omega jt} \nu_{r\omega}} \delta_{r\omega jt}} mc_{\omega jt}$$

Where $\mu_{\omega jt}$ is the firm's markup and $mc_{\omega jt}$ is the marginal cost. After log-linearization, the equation for the price becomes:

$$\tilde{p}^*_{\omega jt} = \tilde{\mu}_{\omega jt} + \tilde{mc}_{\omega jt} - \sum_r \lambda_{r\omega jt} \tilde{\delta}_{r\omega jt}$$

Note that none of the demand shocks have a first-order effect on the prices, a common feature in such frameworks. For our case, this would mean that there is no spillover of demand shocks through the national price.

The last step is to construct the granular instrument in this model. Given that prices are the same across regions, the log-linear version of local demand is given by:

$$\tilde{y}_{r\omega jt} = \tilde{\delta}_{r\omega jt} + \tilde{\nu}_{rjt} + \tilde{\nu}_{\omega jt} + \tilde{y}_{\omega jt}$$

This corresponds exactly to the regression with fixed effects mentioned earlier, with the residuals of that regression identifying $\nu_{r\omega jt}$.

The aggregate granular instrument would be given by:

$$z_{\omega jt} = \sum_{k} \lambda_{k\omega jt} \tilde{\nu}_{r\omega jt}$$

This exactly corresponds to how the iceberg cost shocks enter the equation for the optimal price. Moreover, it means that the coefficient in front of the granular instrument $z_{\omega jt}$ from the first-stage regression would allow us to get an approximate estimate for the pass-through parameter, similarly to Amiti et al. (2019). However, if we choose to see our estimate as an estimate for the pass-through, it will be biased towards zero if our granular shock correlates with other productivity shocks. Additionally, due to higher frequency of the data, it will be more exposed to impacts from sticky prices.

The first-stage regression is similar to the log-linear optimal pricing equation. The result of the first-stage regression is given in Table 1 and confirms the relevance of the suggested instrument. We see a strong negative relation between the shock and the price, which is exactly what would be expected for a supply shock. The negative sign is intuitive and suggests that positive supply shocks, such as an increase in productivity, will make it optimal to decrease the prices.

3.3 Regression Design

The main objective of our empirical estimation is to estimate the partial elasticities discussed in the theoretical section. We set out to estimate four main parameters: own price partial elasticity, own price superelasticity, competitor price partial elasticity, and market-level elasticity.

As follows from the theoretical section, partial elasticities are elasticities with respect to individual firm prices, given the market aggregates. Partial elasticities are important in an oligopolistic setting since they allow us to separate the effects of the firm's own price from its effects on the market-level price. This same fact creates an empirical challenge since any instrument for the price will not satisfy the excludability constraint, by affecting not only the firm's own price but also the market-level price, and potentially even the prices of other firms in the market. Thus, caution is required when choosing controls.

First note that market-level prices and quantities are, by definition, market-level variables and might be accounted for by market-time fixed effects. However, since we cannot rule out heterogeneous responses to

	Price $\tilde{p}_{\omega jt}$			
	(1)	(2)	(3)	(4)
Granular Instrument $\tilde{z}_{\omega jt}$	-0.052***	-0.109***	-0.101***	-0.115***
	(0.003)	(0.005)	(0.004)	(0.005)
Market Share $\lambda_{\omega jt}$			0.019***	0.036***
			(0.001)	(0.004)
Obs.	13,485,909	7,098,861	8,104,721	7,098,218
F	259	506	298	297
Group × Market × Time		✓		✓
Producer			✓	✓

Table 1: First-Stage Regression *The results of the first-stage regression* $\tilde{p}_{\omega jt} = \alpha\tilde{z}_{\omega jt} + P^c_{-\omega jt} + \gamma_{gjt} + \varepsilon_{\omega jt}$ *where* α *is the first-stage coefficient, the parameter of interest. The granular instrument is given as a weighted sum of the regional regression residuals:* $z_{\omega jt} = \sum_k \lambda_{k\omega jt}\tilde{\nu}_{r\omega jt}$. *Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt}y_{\omega jt}}{P_{jt}Y_{jt}}$. *Regression is performed for firm group-market-time fixed effects, and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

the market aggregates by firms of different sizes, it is also crucial to establish control on an even finer level and allow for different constants for firms of different sizes. Thus, we introduce firm-size group-market-time fixed effects into the regressions.

Another concern involves the direct effect of competitor prices, as not all the effect will be captured by the market aggregate controls. This would not be the case in Kimball oligopoly frameworks and only presents an issue if the partial cross-elasticity is non-zero. In this case, the estimate of the partial elasticity from the regression with fixed effects will be biased. Instead of estimating the partial elasticity, the estimate will be:

$$\hat{\xi}^r_{\omega\omega} \approx \xi^r_{\omega\omega} + \{\text{Strategic Complementarity}\} \times \xi^r_{\omega k}$$

Due to the detailed fixed effect strategy, it is impossible to control for the prices of the competitors directly. We will instead consider an alternative procedure, using the simultaneous estimation of two equations with and without fixed effects to test whether $\xi^r_{\omega k}$ is equal to zero. Please refer to the subsection on competitor prices for details. We find that partial cross-elasticity is close to zero and insignificant. Due to this result, it would be safe to assume that partial cross-elasticity is equal to zero, all the competitor price effect is contained in the market aggregates, and the results of the regression with fixed effects are unbiased estimates of partial elasticities.

The baseline specification for the regression is determined by the demand equation, and is given by:

$$\tilde{y}_{\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{\omega jt} + \gamma_{gjt} + \varepsilon_{\omega jt}$$

Where $\xi^r_{\omega\omega}$ is the own price partial elasticity, the parameter of interest, and $\gamma_{g_\omega jt}$ is the firm-size group-market-time fixed effect, with subscript g_ω reflecting the group to which firm ω belongs.

This equation corresponds to the theoretical model and the extension we discussed above. To demonstrate this, we consider the demand equation which will arise at the firm-market-time aggregate level in this model.

$$y_{\omega jt} = \sum_r \nu_{r\omega}\nu_{rjt}\nu_{\omega jt}\Upsilon^{\omega j}\left(\frac{p_{1jt}}{P_{jt}D_{jt}}\dots\frac{p_{n_jjt}}{P_{jt}D_{jt}}\right)Y_{jt}$$

Note that there is an aggregate of regional demand shocks included in this equation, and moreover that iceberg cost shocks do not impact the demand equation. Log-linearization of this equation yields:

$$\tilde{y}_{\omega jt} = \underbrace{\widetilde{(\textstyle\sum_r \nu_{r\omega}\nu_{rjt}\nu_{\omega jt})}}_{\text{demand shock}} + \underbrace{(-\xi^r_{\omega\omega j})\tilde{p}_{\omega jt}}_{\text{own price}} + \underbrace{\sum_{k\neq\omega}(-\xi^r_{\omega kj})\tilde{p}_{kjt}}_{\text{competitors}} + \underbrace{\sum_k(\xi^r_{\omega kj})[\tilde{P}_{jt}+\tilde{D}_{jt}]+\tilde{Y}_{jt}}_{\text{market aggregates}}$$

This equation corresponds to the empirical specification. As discussed, the only thing missing is the direct effect of the competitor prices. The estimation of partial elasticity will be unbiased when the partial cross-elasticities are zero. We test for this in a separate section and find that they are indeed close to zero and insignificant. Note that our constructed aggregate instrument $z_{\omega jt}$ in this case satisfies both excludability and exogeneity requirements, since it does not correlate with the aggregate demand shocks.

Other regressions, including the regression for the regional level, are constructed in a similar fashion. Details will be covered in corresponding sections.

3.4 Partial Elasticity Estimation

In this section we discuss the partial elasticity estimation and its results. As discussed above, we estimate partial elasticities from the following regression:

$$\tilde{y}_{\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{\omega jt} + \gamma_{gjt} + \varepsilon_{\omega jt}$$

Where $\xi^r_{\omega\omega}$ is the own price partial elasticity, the parameter of interest, and $\gamma_{g_\omega jt}$ is the firm-size group-market-time fixed effect, with subscript g_ω reflecting the group to which firm ω belongs.

We perform this regression for two sets of fixed effects, starting with the firm group-market-time fixed effects. This is the common market-time fixed effect, interacted with the firm group by the market share. The latter is added to capture possible heterogeneity in response to aggregates for firms with different market shares. The second set of fixed effects additionally includes firm fixed effects. We double cluster standard errors at the firm and market levels.

The result of the estimation is the partial elasticity for an average firm. The results of the instrumented regression are drastically different from the OLS regression, suggesting that the instrument is powerful.

	Quantity $\tilde{y}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Price $\tilde{p}_{\omega jt}$	-1.519***	-2.074***	-18.153***	-9.003***	-9.756***	-8.269***
	(0.075)	(0.086)	(1.114)	(0.383)	(0.400)	(0.311)
	OLS	OLS	GIV	GIV	GIV	GIV
Obs.	13,518,751	7,105,937	13,470,836	7,098,677	8,104,421	7,098,037
F	409	580	265	554	328	396
Group × Market × Time		✓		✓		✓
Producer					✓	✓

Table 2: Partial Elasticity *The results of the elasticity estimation from the regression $\tilde{y}_{\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{\omega jt} + P^c_{-\omega jt} + \gamma_{gjt} + \varepsilon_{\omega jt}$ where $\xi^r_{\omega\omega}$ is the own price partial elasticity, the parameter of interest. Market share is calculated as a revenue share $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$. Regression is performed for the firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: * - 90%, ** - 95%, *** - 99%*

Moreover, the result of the OLS is closer to zero. This would be expected for the uninstrumented demand regression since apart from supply shocks, which would change prices and quantities in opposite directions and identify demand, the variation also comes from demand shocks that would change prices and quantities in the same direction.

The results for our firm-level partial elasticity are higher then the ones commonly found for elasticity in the macro literature. This distinction should mainly be attributed to the heterogeneity between firms. Remember, the parameter we are estimating is what elasticity would be for a firm with zero market share. For larger firms, it will be closer to the elasticity of the aggregate market and therefore will have a smaller value. Additionally, we need to account for the fact that partial elasticity of a firm is non-constant and there might be a non-zero partial superelasticity. We include this fact in the following subsection, right after addressing a more pressing issue of possible bias coming from the direct effect of the competitors' prices.

As discussed above, one issue that we might encounter with our elasticity estimation is the presence of non-zero partial cross-elasticity, which would bias the results. In the next section we are going to test whether partial cross-elasticity is zero.

3.5 Partial Cross-elasticity Test

In this section we discuss the procedure for assessing partial cross-elasticity. Without additional assumptions we cannot present unbiased estimates for this number, but we can test whether it is equal to zero.

The test procedure is two-fold. First, we consider the regression without fixed effects but with competitor price controls.

$$\tilde{y}_{\omega jt} = (-\xi_{\omega\omega})\tilde{p}_{\omega jt} + \sum_{k\neq\omega, k\in g_\omega} (-\xi_{\omega k})\tilde{p}_{kjt} + \sum_{k\notin g_\omega} (-\xi_{\omega k})\tilde{p}_{kjt} + \varepsilon$$

From this regression without the fixed effects, we can estimate own price elasticity and cross elasticity. Note that this time we are talking about the "total" elasticity estimates, including the effect from the market aggregates, and not the partial elasticity as in the fixed effects regressions in the previous section.

Going back to the model, we note that the estimates of the regressions without the fixed effects would be given by these formulas:

$$\xi_{\omega\omega} = \xi^r_{\omega\omega} + (\xi^m - \sum_k \xi^r_{\omega k} \frac{\sum_k \xi^r_{k\omega}\lambda_k}{\sum_m \sum_k \xi^r_{km}\lambda_k\lambda_\omega})\lambda_\omega$$
$$\xi_{\omega c} = \xi^r_{\omega c} + (\xi^m - \sum_k \xi^r_{\omega k} \frac{\sum_k \xi^r_{kc}\lambda_k}{\sum_m \sum_k \xi^r_{km}\lambda_k\lambda_c})\lambda_c$$

Where $\xi_{\omega\omega}$ is the own price elasticity and $\xi_{\omega c}$ is the cross-elasticity with respect to the price of the competitor c. $\xi^r_{\omega\omega}$ represents the partial elasticity, $\xi^r_{\omega c}$ represents the partial cross-elasticity, λ_{omega} is the market share of firm ω, and ξ^m is the market-level elasticity.

Note that for firms with the same market share, the component coming from the effect on market aggregates is going to be the same. Hence, if we take the difference of the estimated own price total elasticity and total cross elasticity, we get the difference between partial own price elasticity and partial cross-elasticity.

$$c \in g_\omega : \;\; \hat{\beta} = \xi_{\omega\omega} - \xi_{\omega c} = \xi^r_{\omega\omega} - \xi^r_{\omega c}$$

The idea of the next step is to compare this result with estimates of the fixed effects regression to show that the two are close, meaning that partial cross-elasticity needs to be zero.

Recall our discussion of the possible bias with the fixed effect regression. In the case when partial cross-elasticity is non-zero, we estimate a combination of partial cross-elasticity, strategic complementarity, and partial cross-elasticity.

$$\hat{\xi}^r_{\omega\omega} = \xi^r_{\omega\omega} + \{\text{Strategic Complementarity}\}\xi^r_{\omega c}$$

When we take the difference between the estimates of the regression with and without the fixed effects, we obtain:

$$(1 + \{\text{Strategic Complementarity}\})\xi^r_{\omega c}$$

We can then test whether this statistic is equal to zero. Note that because strategic complementarity component is positive or at the very least larger then -1, this equation is only equal to zero when partial cross-elasticity is equal to zero. In the case when strategic complementarity component is negligible, for instance due to sticky prices, the resulting statistic would be a correct estimate for the partial cross-elasticity. In the case of non-zero partial cross-elasticity, the estimation result would be an average partial cross-elasticity

	Quantity $\tilde{y}_{\omega jt}$			
	(1)	(2)	(3)	(4)
Partial Cross-elasticity $\xi^r_{\omega k}$	-0.875***	0.047	-0.740***	0.035
	(0.166)	(0.284)	(0.192)	(0.182)
Obs.	13,470,836	7,098,677	8,104,421	7,098,037
Group × Market × Time		✓		✓
Producer			✓	✓
$\lambda_{\omega jt}$			✓	✓

Table 3: Partial Cross-Elasticity Test *The results of the partial cross-elasticity test. The calculated statistic is equal to the partial cross-elasticity when strategic complementarity is equal to zero; otherwise, the statistic is* $(1+[\textit{Strategic Complementarity}])\xi^r_{\omega k}$. *Due to the fact that strategic complementarity is larger than* -1 *in both cases, the statistic equal to zero indicates zero partial cross-elasticity. The estimation is conducted using GMM. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$. *Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are clustered at the market level. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

across different markets. Under the assumption that goods inside the market are substitutes, partial cross-elasticity should always be positive, so equating average cross-elasticity to zero would still be equivalent to saying that the elasticity in each particular market is close to zero, and the test would still be valid.

Due to our need to estimate two equations simultaneously and obtain a correct variance-covariance matrix, we use a two-step GMM procedure. The structure of the table is similar to the one in the previous section. We preform the regression for two sets of fixed effects: firm group-market-time, and again adding firm fixed effects. When including group-market-time in the "fixed-effect" regression, we include market fixed effects in the "non-fixed-effect" regression. Standard errors are clustered at the market level and are calculated using a delta method.

The resulting statistic, standard errors, and significance levels are given in Table 3. In the incorrect specification, without inclusion of the fixed effects, partial cross-elasticity seems to have a counterintuitive sign. With controls the estimate becomes positive, close to zero, and insignificant, suggesting zero partial cross-elasticity.

The result of this section confirms that the elasticity estimates obtained from the regression for elasticity with firm group-time-market fixed effects are unbiased. In the next section, we proceed to estimate more parameters of demand, starting with partial superelasticity.

3.6 Partial Superelasticity Estimation

In this section we discuss the partial superelasticity estimation and its results. We estimate partial superelasticity from the following regression:

$$\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega j})\tilde{p}_{r\omega jt} + [(-\xi^r_{\omega\omega j}) \times (\eta^r_{\omega\omega\omega})]\tilde{p}^2_{r\omega jt} + \gamma + \delta_{r\omega jt}$$

	Quantity $\tilde{y}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Price $\tilde{p}_{\omega jt}$	-1.594***	-2.170***	-15.751***	-7.602***	-8.223***	-6.948***
	(0.073)	(0.072)	(0.781)	(0.240)	(0.266)	(0.192)
$\tilde{p}^2_{\omega jt}$	-0.389***	-1.364***	-12.415***	-11.243***	-11.571***	-12.059***
	(0.074)	(0.142)	(2.295)	(1.638)	(1.674)	(1.592)
Partial Superelasticity η^r_ω	0.244***	0.628***	0.788***	1.479***	1.407***	1.736***
	(0.002)	(0.004)	(0.019)	(0.042)	(0.040)	(0.048)
	OLS	OLS	GIV	GIV	GIV	GIV
Obs.	13,518,751	7,105,937	13,470,836	7,098,677	8,104,421	7,098,037
F	236	471	204	509	343	459
Group × Market × Time		✓		✓		✓
Producer					✓	✓

Table 4: Partial Superelasticity *The results of the partial superelasticity estimation from the regression* $\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega j})\tilde{p}_{r\omega jt} + [(-\xi^r_{\omega\omega j}) \times (\eta^r_{\omega\omega\omega})]\tilde{p}^2_{r\omega jt} + \gamma + \delta_{r\omega jt}$ *where* $\xi^r_{\omega\omega}$ *is the own price partial elasticity,* $\eta^r_{\omega\omega\omega}$ *is the own price superelasticity, the parameter of interest. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$. *Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market level. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

Where $\xi^r_{\omega\omega}$ is the own price partial elasticity, $\eta^r_{\omega\omega}$ is the own price superelasticity, the parameter of interest, and $\gamma_{g_\omega jt}$ is the firm-size group-market- time fixed effect, with subscript g_ω reflecting the group to which firm ω belongs.

The structure of the table is the same as before. We perform this regression for two sets of fixed effects. We double cluster standard errors at the firm and market levels. The result of this estimation is the partial elasticity and partial superelasticity for an average firm. Note that to obtain the estimate for the superelasticity, as shown in the regression equation, we need to divide the two coefficients from the regression. The final estimate for superelasticity is given in a separate line. Standard errors are calculated using the delta method.

The estimate for partial elasticity is 6.95. This is the elasticity for the zero-share firm. Other firms' elasticities will be affected by their impact on the market aggregates and will be different. Moreover, the actual estimates for the non-zero share firm might not only be affected by its own market share, but by other characteristics of the market, such as concentration.

The estimate for the partial superelasticity is 1.74. Again, this will not reflect the superelasticity of all the firms, especially the largest firms, since there are going to be additional effects from their impact on the market aggregates. The partial superelasticity is positive, which is in line with Marshall's second law. The superelasticity of 1.74 would mean that for a 1% increase of price, the demand becomes 1.74% more elastic. In the next section we are going to address the fact that different firms are at the different points of their respective demand functions in the equilibrium.

	Quantity $\tilde{y}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Price $\tilde{p}_{\omega jt}$	-1.830***	-2.010***	-12.887***	-9.525***	-10.347***	-8.682***
	(0.104)	(0.094)	(0.730)	(0.445)	(0.462)	(0.358)
Price $\tilde{p}_{\omega jt}$ × Market Share $\lambda_{\omega jt}$	-1.858***	-3.049***	23.269***	23.912***	15.709***	18.528***
	(0.471)	(0.623)	(3.062)	(3.329)	(2.071)	(2.636)
	OLS	OLS	GIV	GIV	GIV	GIV
Obs.	8,121,239	7,105,937	8,104,856	7,098,677	8,104,421	7,098,037
F	221	511	144	282	225	334
Group × Market × Time		✓		✓		✓
Producer					✓	✓

Table 5: Partial Elasticity Heterogeneity *The results of the across-share heterogeneity estimation from the regression* $\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{r\omega jt} + \alpha\tilde{p}_{r\omega jt} \times \lambda + \gamma + \delta_{r\omega jt}$ *where* $\xi^r_{\omega\omega}$ *is the own price partial elasticity,* α *is the degree of heterogeneity of elasticity with respect to firm shares, the parameter of interest. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$*. Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

3.7 Heterogeneity

In this section, we discuss how partial elasticity depends on the market share of the firm. We estimate this dependence from the following regression:

$$\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{r\omega jt} + \alpha\tilde{p}_{r\omega jt} \times \lambda + \gamma + \delta_{r\omega jt}$$

Where $\xi^r_{\omega\omega}$ is the own price partial elasticity, α is the degree of homogeneity of partial elasticity with respect to firm shares, the parameter of interest, $\lambda_{\omega jt}$ is market share, and $\gamma_{g_\omega jt}$ is the firm-size group-market-time fixed effect, with subscript g_ω reflecting the group to which firm ω belongs. The result of this estimation is shown in Table 5. The structure of the table is the same as before; we start from a simple OLS and then proceed to do the GIV estimation with different sets of fixed effects. As discussed in the theoretical section, the presence of this type of asymmetry will not affect the external validity properties if there are no other asymmetries that are unaccounted for.

Our result indicates a strong dependence of partial elasticity on market shares. The estimate for the degree of this dependence is 18.53, meaning that if a zero-share firm has a partial elasticity of 8.68, then a firm with 10% market share would have a partial elasticity of 6.82. This heterogeneity is not just important for the correct estimation of elasticity, but also for estimating the additional effect of market concentration on elasticity for all firms.

	Market Quantity $\tilde{Y}_{jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Market Price $\tilde{P}_{jt}$	-2.167***	-2.156***	-8.971***	-6.411***	-7.046***	-3.955***
	(0.276)	(0.261)	(0.699)	(0.571)	(0.663)	(0.494)
$\tilde{P}^2_{jt}$					-24.596***	-34.305***
					(5.114)	(4.662)
Market Superelasticity η^m					3.491***	8.674***
					(0.744)	(3.111)
	OLS	OLS	GIV	GIV	GIV	GIV
Obs.	506,798	506,798	506,798	506,798	506,798	506,798
Time		✓		✓		✓
Market		✓		✓		✓

Table 6: Market-level Elasticity *The results of the market elasticity and market cross-elasticity estimations from the regression* $\tilde{Y}_{jt} = (-\xi^m)\tilde{P}_{jt} + (-\xi^m)(\eta^m)\tilde{P}^2_{jt}\gamma + \delta_{jt}$ *where* ξ^m *is market elasticity and* η^m *is market superelasticity, the parameters of interest. The regression is performed for time and market fixed effects. Standard errors are clustered at the market level. Asterisks mark significance levels: * - 90%, ** - 95%, *** - 99%*

3.8 Market Elasticity

In this section we discuss the market elasticity and superelasticity estimations, along with their results. We estimate market elasticity and superelasticity from the following regression:

$$\tilde{Y}_{jt} = (-\xi^m)\tilde{P}_{jt} + (-\xi^m)(\eta^m)\tilde{P}^2_{jt}\gamma + \delta_{jt}$$

Where ξ^m is the market-level elasticity and η^m is the market-level superelasticity, both parameters of interest, and γ_j and γ_t are the time and market fixed effects.

The aggregation to the market level, as before, relies on the homogeneity assumption with the aggregate market price and quantity equal to revenue-share weighted prices and quantities, respectively. The granular instrument for the market level is also a collapsed firm-level granular instrument, given by:

$$z^m_{jt} = \sum_{\omega} \lambda_{\omega jt-1} z_{\omega jt}$$

As before, this instrument is a supply shock, so the sign of the coefficient in front of the granular instrument in the first-stage regression should be negative. This is indeed what we get in the first stage, with the coefficient equal to -0.14. Please see the table in the appendix. The results of the market elasticity and superelasticity estimations are given in Table 6. Note that the market-level elasticity is much lower then the firm-level partial elasticity. This is an intuitive result since when we go to a higher level of aggregation, we should expect elasticity to decrease. Moreover, for our oligopolistic framework this means that elasticity of the firms with larger revenue-based market shares would be lower than the elasticity of the lower- share firms.

In our estimations, market-level elasticity is 4.00 and market-level superelasticity is 8.54. The superelasticity is positive, once again confirming Marshall's second law. Moreover, product-level demand is large, meaning that constant elasticity would not be a good approximation and we need to use a richer Kimball demand aggregator.

3.9 Robustness

In this subsection we will perform a robustness check with the regional granular shock. This procedure will allow us to relax the assumption of no firm-region-market-time shocks and replace it with the assumption of no spillover of firm-region-market-time shocks across regions. For this section, we are going to work with a 5% sample of the data collapsed to the firm-region-market-time level, instead of the whole dataset collapsed to the firm-market-time level as in the previous sections. The sample is balanced across regions and time. The granular shock is formulated as a sum of all the residuals, except for the ones coming from the region of interest.

$$z_{r\omega jt} = \sum_{k \neq r} \lambda_{k\omega jt-1} \tilde{\nu}_{k\omega jt}$$

Where $\lambda_{k\omega jt-1}$ is the market share of firm ω in market j at time t in region k, and $\tilde{\nu}_{k\omega jt}$ is the shock coming from region k. Similarly to the baseline regression, we estimate partial elasticity from the following regression:

$$\tilde{y}_{r\omega jt} = (-\xi^{r}_{\omega\omega}) \tilde{p}_{r\omega jt} + \gamma_{gjt} + \varepsilon_{r\omega jt}$$

Where $\tilde{y}_{r\omega jt}$ is the change in quantity, $\tilde{p}_{r\omega jt}$ is the change in prices, $\xi^{r}_{\omega\omega}$ is partial elasticity, the parameter of interest, and γ_{gjt} is the fixed effect. This regression is performed for various sets of fixed effects. Standard errors are triple clustered on the firm, market, and region levels. To account for the fact that regions have different sizes, we weight the observations according to the revenue from a particular region for the firm. The results of this estimation are given in Table 7. Note that the estimates for elasticity are similar to the estimates we got in the baseline regression. This suggests that the assumption of non-significant firm-region-market-time demand shocks is reasonable.

3.10 Empirical Result Conclusion

We have estimated partial elasticity and partial superelasticity, found that elasticity is persistently heterogeneous with respect to shares, connected it to possible dynamic effects of prices, and finally estimated elasticity and superelasticity for the market level.

Our partial elasticity estimates are larger then the estimates of the elasticity commonly found in the macroeconomic literature. This is due to the fact that partial elasticities only coincide with the elasticities of low-share firms; for larger firms, elasticities would be lower. This is also highlighted by the fact that the market-level elasticities are lower then the firm-level elasticities. This is the first hint that larger firms have smaller elasticity.

	Quantity $\tilde{y}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Price $\tilde{p}_{r\omega jt}$	-1.365***	-1.356***	-10.167***	-6.873***	-7.555***	-6.237***
	(0.063)	(0.059)	(0.627)	(0.527)	(0.329)	(0.251)
Market Share $\lambda_{r\omega jt-1}$					0.179***	0.321***
					(0.010)	(0.023)
	OLS	OLS	GIV	GIV	GIV	GIV
Obs.	320,998,850	320,998,850	320,998,850	320,998,850	320,998,850	320,998,850
F	463	530	262	432	309	372
Group × Market × Time		✓		✓		✓
Producer					✓	✓

Table 7: Regional Estimation of Partial Elasticity *The results of the partial elasticity estimation from the regression $\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{r\omega jt} + \gamma_{gjt} + \varepsilon_{r\omega jt}$ where $\xi^r_{\omega\omega}$ is the own price partial elasticity, the parameter of interest. Observations are weighted by the lagged share of the revenue from a particular region for the firm. Market share is calculated as a revenue share $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$. Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are clustered at the market level. Asterisks mark significance levels: * - 90%, ** - 95%, *** - 99%*

Partial superelasticity and market superelasticity are both positive, confirming Marshall's second law of demand. We find that partial superelasticity at the firm level is low, possibly making it negligible for some practical applications. At the same time, market-level elasticity is large, making it important to use a Kimball demand aggregator instead of CES.

Even though partial superelasticity is close to zero, there is a significant heterogeneity of partial elasticity across firms, with larger firms enjoying much smaller partial elasticity. As suggested by our results for the dynamic effects of prices, this heterogeneity might be due to the ability of firms to form their customer base over time.

In the next section we are going to calculate total elasticity, markup, pass-through, superelasticity, strategic complementarity, and second-order marginal cost effects, using the formulas derived in the theoretical section.

4 Elasticity, Pass-through, and Second-Order Marginal Cost Effects

In this section, we will calculate elasticities that include the impact on the market aggregates. We will use the following estimate as our benchmark:

$$\tilde{y}_{\omega jt} = \underbrace{-\underset{(0.23)}{7.39}}_{\xi_0^r} \left[\tilde{p}_{\omega jt}\right] + \underbrace{\underset{(2.55)}{18.71}}_{\alpha_\xi} \left[\tilde{p}_{\omega jt} \times \lambda_{\omega jt-1}\right] - \underset{(1.57)}{11.83} \left[\tilde{p}^2_{\omega jt}\right] + \underset{(0.22)}{0.19} \left[\lambda_{\omega jt-1}\right] + \gamma_{gjt} + \varepsilon$$

We additionally considered an interaction between market shares and the second-order price term, but found that it is insignificant and hence omitted it from this regression and other calculations.

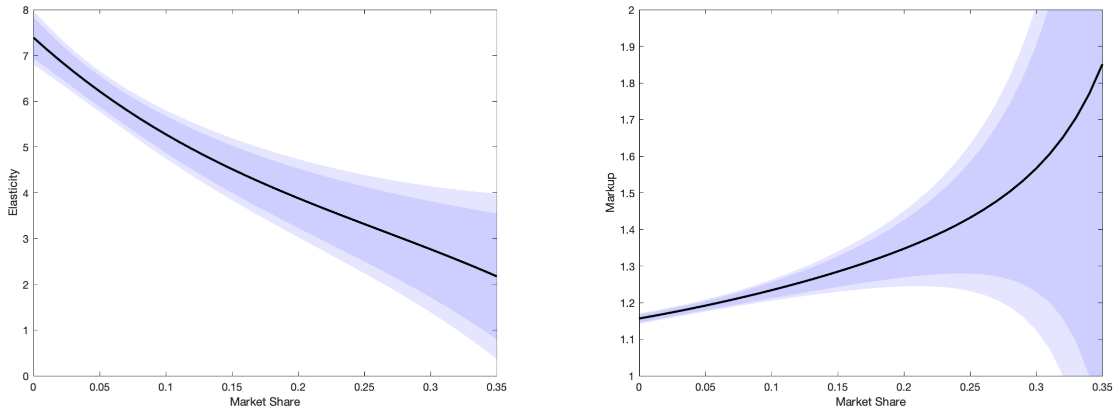

Figure 1: Dependence of Elasticity and Markup on Market Share
This figure shows the result of the elasticity and markup estimation in terms of dependence of market shares. The left panel shows elasticity while the right panel shows elasticity. Market shares are calculated as revenue shares $\frac{p_{\omega j} y_{\omega j}}{P_j Y_j}$. *Standard errors are double clustered on the market and producer levels and are calculated using the delta method. Shaded regions represent 95% and 99% confidence intervals.*

The values of elasticities and other relevant parameters depend on the market-share distribution in the market. For most of the following analyses, we have fixed HHI at 0.14, which corresponds to the median HHI level in our sample. Some of the calculations also require other market share aggregates; for instance, superelasticity also requires the sum of cubes of shares. These aggregates were taken as averages for the markets in the corresponding HHI bracket. For the baseline HHI of 0.14, the H^3 is equal to 0.03. More detailed descriptions of the assumptions can be found in the appendix.

For most of the following work, we plot the estimated parameters for the market shares from 0 to 35%, which corresponds to the ??th percentile in our sample. For cases where we need to compare large and small firms, we take a firm with 1% market share as a small firm and a firm with 30% market share as a large firm. The firm with 30% market share can be treated as a "typical" large firm, as 65% of the markets have a firm with this market share or larger.

4.1 Elasticity

We calculate elasticities and markups, then demonstrate how both depend on the market share. The results are shown in Figure 1. Elasticity decreases with shares, while markup increases. Elasticity for the smallest firms exceeds 7, while elasticity of the firms controlling a third of their corresponding market is below 3. Markups also increase sharply, from below 1.2 for the smallest firms to potentially over 1.5 for the largest firms.

In the case of heterogeneous partial elasticity, it is no longer the case that only the firm's own market share matters. Instead, the characteristics of the market, in particular market concentration, influence the elasticities of individual firms. In our simplified case, where elasticity is a linear function of the market share, concentration would be the only market share distribution aggregate that is needed to calculate the

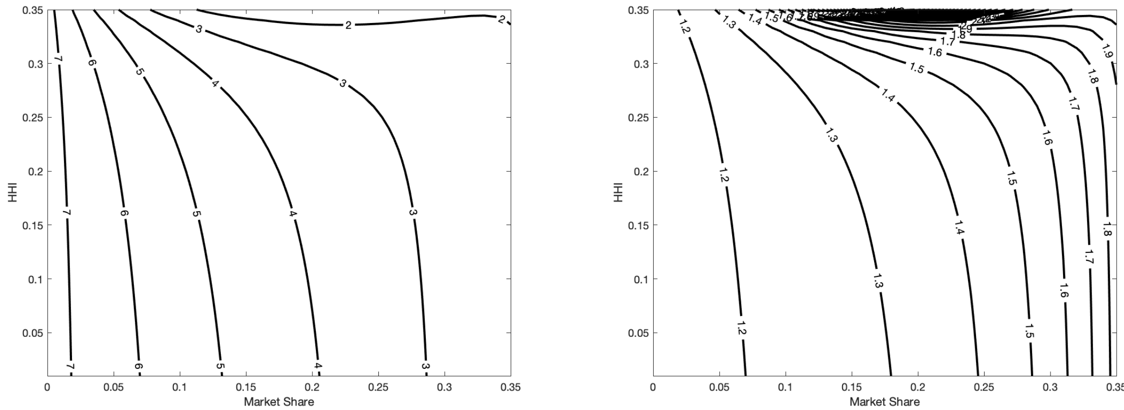

Figure 2: Dependence of Elasticity and Markup on the Market Share and Concentration
This figure shows the result of the elasticity estimation, including the impact of the firms on the market share, for different levels of market concentration. The left panel shows elasticity and the right panel shows markup. Market share is revenue share $\lambda_\omega = \frac{p_\omega y_\omega}{PY}$. *Levels of concentration are calculated by HHI =* $\sum_\omega \lambda_\omega^2$ *and are written as a share of 1.*

elasticity. In fact, it will be given by:

$$\xi_{\omega\omega} = (\xi_0^r + \alpha_\xi \lambda_\omega) + \xi^m \lambda_\omega - \frac{(\xi_0^r + \alpha_\xi \lambda_\omega)^2}{\xi_0^r + \alpha_\xi \text{HHI}} \lambda_\omega$$

$$\text{HHI} = \sum_\omega \lambda_\omega^2$$

Note that the elasticity of any given firm would decrease as HHI increases. The effect of HHI is however fairly mild, especially for the smaller firms, so we are only able to present suggestive evidence of the magnitude of this effect in Figure 2. Not all the differences presented in this graph are significant. First note the already discussed effect of decreasing elasticity with respect to market share. This can be seen as we propagate along the x axis of the map graph. The additional effect of HHI can be seen as we propagate along the y axis of the map graph. We can note that as the HHI increases, elasticity decreases, even if the firm's own market share remains the same. For instance, for a firm that controls 20% of the market with an HHI of 0.14, elasticity will be over 4, while a firm with the same share in a more concentrated market can have an elasticity of just 3. Similar patterns can be observed for markups: not only will larger firms will have larger markups, but also firms in more concentrated markets.

4.2 Pass-through and Superelasticity

In this subsection we are going to take a look at the linear terms of price adjustment: pass-through and superelasticity. As discussed in the theoretical section, pass-through would be given by:

$$\text{Pass-through} = \frac{\xi_\omega - 1}{\xi_\omega - 1 + \eta_\omega}$$

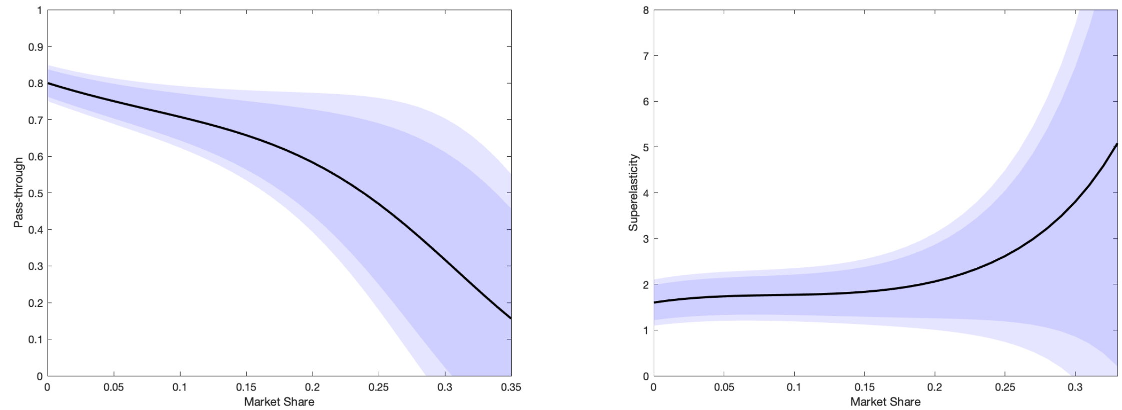

Figure 3: Dependence of the Optimal Pass-through and Superelasticity on the Market Share
This figure shows the result of the optimal pass-through and superelasticity estimation in terms of dependence on market shares. The left panel shows pass-through and the right panel shows superelasticity. Market shares are calculated as revenue shares $\frac{p_{\omega j} y_{\omega j}}{P_j Y_j}$. *Standard errors are double clustered on the market and producer levels and are calculated using the delta method. Shaded regions represent 95% and 99% confidence intervals.*

The two important parameters to calculate the pass-through are elasticity and superelasticity. Note that pass-through will always be positive and smaller than one whenever Marshall's second law is fulfilled and superelasticity is positive. As we have seen, this holds empirically, so we can get intuitive estimates for the pass-through in an interval between zero and one without imposing it as a constraint. The results for pass-through and superelasticity are given in Figure 3. To analyse pass-through and superelasticity even in our simple case, we would need to know more about the distribution of shares in the market. One more parameter must be added to the HHI index: the sum of cubes of shares.

$$\mathrm{H}^3 = \sum_{\omega} \lambda_{\omega}^3$$

Since HHI and H^3 are closely connected, we choose one H^3 for each HHI, the average H^3 found in the market with that particular HHI. The resulting dependence can be found in the appendix.

The resulting pass-through and elasticity plots are given in Figure 3. Note that, as before, pass-through depends on the market share of the firm. A larger firm would have smaller pass-through. The degree of pass-through differs dramatically between firms, with the smallest firms exhibiting a pass-through close to 0.8 and firms controlling a third of their market going as low as 0.2.

Similarly to our results for elasticity, we also find a dependence of pass-through on market concentration. These effects again seem mild and we are only able to show some of the suggestive evidence. The results are shown in the Figure 4. A firm that controls 20% of the market would have a pass-through of over 0.5 in an non-concentrated market, while the same firm's pass-through can get as low as 0.3 if the market concentration is high. Note most importantly that not only do firms' own shares matter for their pass-through, but also the concentration of the market.

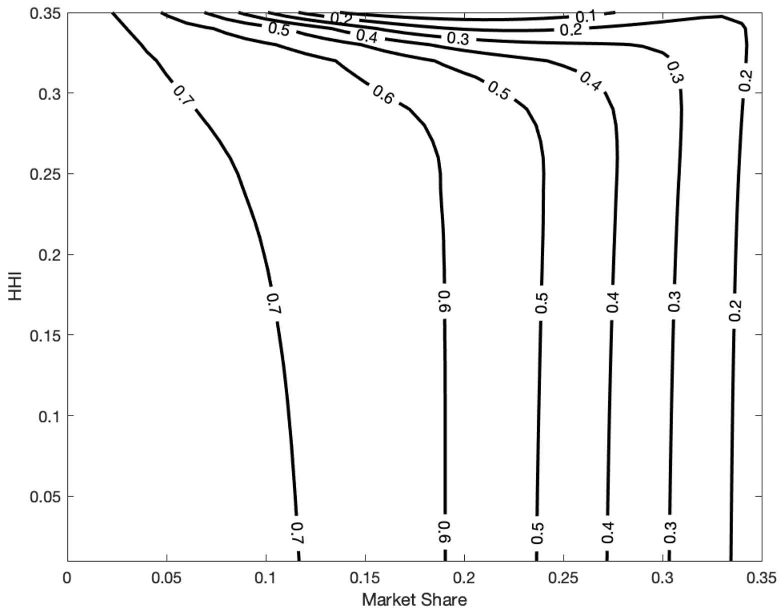

Figure 4: Dependence of the Optimal Pass-through on the Market Share and Concentration
This figure shows the result of the superelasticity estimation for different firms and markets with different concentration. Market share is a revenue share $\lambda_\omega = \frac{p_\omega y_\omega}{PY}$. *Relevant concentration parameters are: HHI* $= \sum_\omega \lambda_\omega^2$, $H^3 = \sum_\omega \lambda_\omega^3$. H^3 *is chosen uniquely for each* HHI, *and the assumption details can be found in the appendix.*

The direct effect of the firm's market share and the more mild effect of the market concentration mean that a larger firm in a more concentrated market has a lower pass-though. In other words, given the same change in marginal costs, larger firms would experience smaller changes in their desired prices and, as a consequence, adjust their prices by less. In the aggregate economy this would mean that the inflation response is dampened and we would not observe large fluctuations in final goods prices, even if there is a significant change to the marginal costs of the firms. This, however, is the result from a linear approximation that does not take into account the non-linear effect of marginal cost shocks. We consider this non-linear effect in the next subsection.

4.3 Second Order Marginal Cost Effect

For the larger marginal cost shocks, a linear approximation of the change in the desired price might not be enough. To address this issue we consider the elasticity of pass-through with respect to the marginal cost shock. As we derived in the theoretical section, the elasticity of the pass-through with respect to the marginal cost shock is given by:

$$\frac{\partial \log \text{Pass-through}}{\partial \log mc} = \left(\frac{\eta_\omega}{\xi_\omega - 1}\right)\left(\eta_\omega \frac{\xi_\omega}{\xi_\omega - 1} - \psi_\omega\right)\left(\frac{\xi_\omega - 1}{\xi_\omega - 1 + \eta_\omega}\right)^2$$

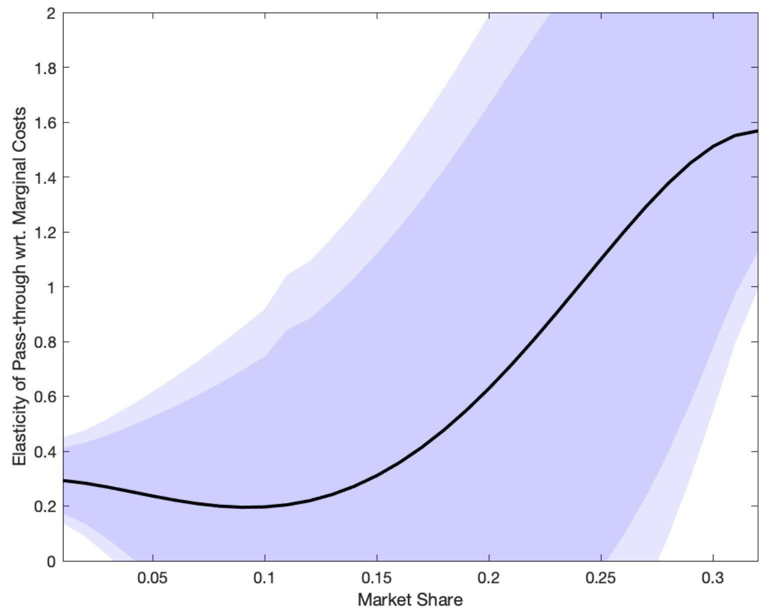

Figure 5: Dependence of Pass-through Elasticity on Market Share
This figure shows optimal pass-through elasticity with respect to the marginal cost shock in terms of dependence of market shares. Market shares are calculated as revenue shares $\frac{p_{\omega j} y_{\omega j}}{P_j Y_j}$. *Standard errors are double clustered on the market and producer levels and are calculated using the delta method. Shaded regions represent 95% and 99% confidence intervals.*

Note that for the elasticity of pass-through with respect to the marginal cost shock, we additionally need to know the elasticity of the superelasticity ψ. For this calculation we are going to assume that partial superelasticity is a constant, so the partial elasticity of the superelasticity is zero. Then ψ would only depend on the effect of the firm on the market aggregates, allowing us to calculate it using a decomposition formula, similar to the ones we employed for elasticity and superelasticity.

The estimation result for the elasticity of the pass-through with respect to the marginal cost shock is given in Figure 5. Note that the elasticity of the pass-through is positive for firms with any market share and for the largest and smallest firms it is significant at the 99% level. Moreover, the elasticity of the pass-through is significantly larger for large firms, meaning that for larger firms non-linear effects are more important.

The positive elasticity of pass-through with respect to the marginal cost shock means that there is asymmetry in the responses of prices to marginal cost shocks; moreover, the response to positive marginal cost shocks is larger than the response to negative marginal cost shocks. The total effect of the marginal cost shocks on prices can be seen in Figure 6. The straight line represents the linear approximation of the price change, while the curve represents the total price change, including the second-order marginal cost effect. The left and right panels depict a small firm with 1% market share and a large firm with 30% market share, respectively. Note that due to the positive marginal cost effect on the pass-through in both cases, there

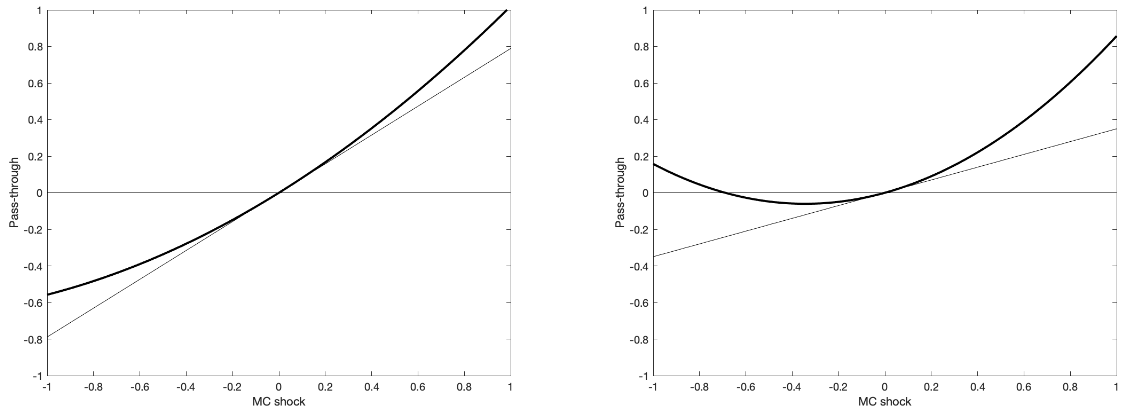

Figure 6: Total Response of the Optimal Price to the Marginal Cost Shock
This figure shows the total response of the optimal price to a marginal cost shock for large and small firms. The left panel shows the total response of a small firm with a market share of 1%, while the right panel shows the total response of a large firm with a market share of 30%. The tangent straight line represents the linear approximation.

is an asymmetric effect of the marginal costs on prices, with positive marginal cost shocks having larger pass-through than the negative marginal cost shocks. This effect is especially prominent for the larger firm: large negative marginal cost shocks can have a pass-through close to zero, while large positive marginal cost shocks can have a pass-through of over 0.8.

The asymmetric effect of the marginal costs has an important implication for calculating desired prices and desired inflation. As we saw from the section on pass-through, in an economy where large firms are more prominent, the aggregate pass-through will be lower– this would create the potential for limited propagation of marginal costs into prices and a lack of final goods price response. On the other hand, there will be a pronounced asymmetry in the reaction of prices to marginal costs. Negative marginal cost shocks will have a close-to-zero pass-through, while positive marginal cost shocks might have a pass-though that is comparable to that of an economy without large firms.

5 Conclusion

In this paper we examined how variable markups affect pass-through. We estimated heterogeneous pass-through for firms controlling different shares of the market, then assessed the elasticity of the pass-through to marginal shock effects. We found that the response of desired prices to marginal cost shocks is asymmetric, with negative marginal cost shocks having smaller pass-through than positive marginal cost shocks. We also found that this asymmetric effect is more pronounced for large firms and, hence, in economies where larger shares of the market are controlled by large firms. In such economies, the asymmetric effect can be such that negative marginal costs shocks have pass-through close to zero, while pass-through of the large positive marginal cost shocks approaches that of competitive economies without a significant share of large firms.

This result helps to explain the recent trend in inflation. Consider a concentrated economy with many firms controlling large shares of their respective markets. First, in such an economy the response to small marginal cost shocks is going to be staggered due to the low pass-through of marginal costs into desired prices. This would mean that in general, the inflationary pressures would be kept at low levels. Second, if such an economy faced a large negative marginal cost shock, similar to one that could occur during a crisis caused by a negative aggregate demand shock, the expected deflationary response would be virtually non-existent, creating a missing deflation puzzle. But if such an economy encountered a large positive marginal cost shock, the pass-through would suddenly resemble that of a competitive economy without large firms, and the resulting inflationary pressures would be significant. In such an economy, we would observe long periods of low inflation that would rapidly turn into an inflation spike when there is a large positive marginal cost shock.

References

Afrouzi, Hassan (2020) "Strategic inattention, inflation dynamics, and the non-neutrality of money",Technical report, Working Paper. 4

Afrouzi, Hassan, Andres Drenik, and Ryan Kim (2022) "Growing by the masses",Technical report, Working Paper. 4

Alvarez, Fernando E., Francesco Lippi, and Aleksei Oskolkov (2020) "The macroeconomics of sticky prices with generalized hazard functions",Technical report, National Bureau of Economic Research. 4

Amiti, Mary, Oleg Itskhoki, and Jozef Konings (2014) "Importers, exporters, and exchange rate disconnect.", *American Economic Review*, 104 (7), pp. 1942–1978. 4

Amiti, Mary, Oleg Itskhoki, and Jozef Konings (2019) "International shocks, variable markups and domestic prices", *The Review of Economic Studies*, 86 (6), p. 2356–2402. 2, 4, 20

Atkeson, Andrew and Ariel Burstein (2008) "Pricing-to-market, trade costs, and international relative prices", *American Economic Review*, 98 (5), pp. 1998–2031. 2, 4

Baqaee, David, Emmanuel Farhi, and Kunal Sangani (2021) "The supply-side effects of monetary policy (no. w28345). national bureau of economic research",Technical report, National Bureau of Economic Research. 4, 15

Berry, Steven, Amit Gandhi, and Philip Haile (2013) "Connected substitutes and invertibility of demand", *Econometrica*, 81 (5), p. 2087–2111. 6

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), pp. 561–644. 4

De Loecker, Jan and Frederic Warzynski (2012) "Markups and firm-level export status", *The Quarterly Journal of Economics*, 102 (6), pp. 2437–2471. 4

DellaVigna, Stefano and Matthew Gentzkow (2019) "Measuring market power in the ready-to-eat cereal industry", *The Quarterly Journal of Economics*, 134 (4), p. 2011–2084. 3, 17

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu (2015) "Competition, markups, and the gains from international trade", *American Economic Review*, 105 (10), pp. 3183–3221. 4

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu (2018) "How costly are markups?",Technical report, National Bureau of Economic Research. 4

Foster, Lucia S., John C. Haltiwanger, and Cody Tuttle (2022) "Rising markups or changing technology?",Technical report, National Bureau of Economic Research. 4

Gabaix, Xavier and Ralph S.J. Koijen (2020) "Granular instrumental variables",Technical report, National Bureau of Economic Research. 3, 17

Gale, David and Hukukane Nikaidô (1965) "The jacobian matrix and global univalence of mappings", *Mathematische Annalen*, 159, pp. 81–93. 6, 41, 44

Golosov, Mikhail and Robert E. Lucas Jr. (2007) "Menu costs and phillips curves", *Journal of Political Economy*, 115 (2), pp. 171–199.

Gopinath, Gita and Oleg Itskhoki (2010) "Currency choice and exchange rate pass-through", *American Economic Review*, 100 (1), pp. 304–336.

Gopinath, Gita and Oleg Itskhoki (2011) "In search of real rigidities", *NBER Macroeconomics Annual*, 25 (1), pp. 261–310. 4

Gowrisankaran, Gautam and Marc Rysman (2012) "Dynamics of consumer demand for new durable goods", *Journal of political Economy*, 120 (6), pp. 1173–1219. 4

Hausman, Jerry A. (1996) "Valuation of new goods under perfect and imperfect competition.", *The economics of new goods*, pp. 207–248. 3, 17

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein (2016) "Quantifying the sources of firm heterogeneity", *The Quarterly Journal of Economics*, 131 (3), pp. 1291–1364. 4

Jordá, Óscar (2005) "Estimation and Inference of Impulse Responses by Local Projections", *American Economic Review*, 95 (1), pp. 161–182.

Kimball, Miles S. (1995) "The quantitative analytics of the basic neomonetarist model", *Journal of Money, Credit and Banking*, 27 (4), pp. 1241–1277. 3, 4, 7

Klemperer, Paul (1987) "Markets with consumer switching costs", *The Quarterly Journal of Economics*, 102 (2), pp. 375–394. 4

Klenow, Peter J. and Jonathan L. Willis (2016) "The rise of market power and the macroeconomic implications", *Economica*, 83 (331), pp. 443–472. 4

La'O, Jennifer and Alireza Tahbaz-Salehi (2020) "Optimal monetary policy in production networks",Technical report, National Bureau of Economic Research. 4

Leibtag, Ephraim, Alice Nakamura, Emi Nakamura, and Dawit Zerom (2007) "Cost pass-through and sticky prices in the ground coffee industry", *US Department of Agriculture Online Journal*, pp. 2013–297.

Mckenzie, Lionel (1960) *Matrices With Dominant Diagonals and Economic Theory*. 6

Mongey, Simon (2021) "Market structure and monetary non-neutrality",Technical report, National Bureau of Economic Research. 4

Nakamura, Emi and Dawit Zerom (2010) "Accounting for incomplete pass-through", *The Review of Economic Studies*, 77 (3), pp. 1192–1230. 4

Nevo, Aviv (2001) "Measuring market power in the ready-to-eat cereal industry", *Econometrica*, 69 (2), pp. 307–342. 3, 4, 17

Ravn, Morten, Stephanie Schmitt-Grohé, and Martín Uribe (2006) "Deep habits", *The Review of Economic Studies*, 73 (1), pp. 195–218. 4, 13, 45

Rotemberg, Julio J. and Michael Woodford (1990) "Cyclical markups: Theories and evidence.",Technical report.

Rotemberg, Julio J. and Michael Woodford (1992) "Oligopolistic pricing and the effects of aggregate demand on economic activity", *Journal of Political Economy*, 100 (6), pp. 1153–1207. 4

Rotemberg, Julio J. and Michael Woodford (1999) "The cyclical behavior of prices and costs", *Handbook of Macroeconomics*, 1, pp. 1051–1135. 4

Rubbo, Elisa (2020) "Networks, phillips curves and monetary policy",Technical report, Working Paper. 4

Samuelson, Paul A. (1952) "The transfer problem and transport costs: The terms of trade when impediments are absent", *The Economic Journal*, 62 (246), p. 278–304. 18

Shcherbakov, Oleksandr (2016) "Measuring consumer switching costs in the television industry", *The RAND Journal of Economics*, 47 (2), pp. 366–393. 4

Wang, Olivier and Iván Werning (2020) "Dynamic oligopoly and price stickiness",Technical report, National Bureau of Economic Research. 3, 4, 7

A Proposition 1: Demand Invertibility

Proposition 1 *Given the definition of the function F, this system of non-linear equations has a unique solution and this solution is a minimum of the cost minimization problem.*

Consider the problem of choosing $\frac{y_{\omega jt}}{Y_{jt}}$ given $p_{\omega jt}$.

$$\begin{aligned} P_j &= \min_{\frac{y_{\omega j}}{Y_j}} \sum_{\omega=1}^{n_j} p_{\omega j} \frac{y_{\omega j}}{Y_j} \\ \text{s.t.} \quad & F\left(\frac{y_{1j}}{Y_j}, .. \frac{y_{\omega j}}{Y_j}, .. \frac{y_{n_j j}}{Y_j}\right) = 1 \end{aligned}$$

Where $y_{\omega j}$ is the quantity of the good of the firm ω operating at the market j, Y_j is the aggregate quantity of the market j, $p_{\omega j}$ is the price of the good of the firm ω operating at the market j, P_j is the aggregate price of the market j. The first-order conditions can be written as:

$$\begin{aligned} \forall \omega : \quad & \left\{ F'_\omega \left(\tfrac{y_{1j}}{Y_j} .. \tfrac{y_{\omega j}}{Y_j} .. \tfrac{y_{n_j}}{Y_j}\right) = \tfrac{p_{\omega j}}{P_j D_j} \right. \\ \text{where } D_j &= \left(\sum_\omega F'_\omega \frac{y_{\omega j}}{Y_j}\right)^{-1} \\ P_j &= \sum_\omega \frac{y_{\omega j}}{Y_j} p_{\omega j} \\ F & \left(\frac{y_{1j}}{Y_j} .. \frac{y_{\omega j}}{Y_j} .. \frac{y_{n_j j}}{Y_j}\right) = 1 \end{aligned}$$

By assumption, the function F is such that the first derivatives are all positive and the Hessian matrix of the second derivatives is negative definite. Consider the following system of equations that defines each of the $\frac{y_{\omega jt}}{Y_{jt}}$ in terms of the relative price vector $\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right)$.

$$\forall \omega \left\{ F_\omega \left(\tfrac{y_{1jt}}{Y_{jt}} \dots \tfrac{y_{n_j jt}}{Y_{jt}}\right) = \tfrac{p_{\omega jt}}{P_{jt} D_{jt}} \right.$$

The Jacobian matrix of this system coincides with the Hessian matrix of the function F and hence it is a negative definite matrix. Hence the Jacobian for the $(-1) \times F_\omega$ is a positive definite matrix. Using the fact that a positive definite matrix is a P-matrix and the result of Gale and Nikaidô (1965), we conclude that both $(-1) \times F_\omega$ and F_ω are injective maps. This means that this system has a unique solution and there is a smooth function:

$$\frac{y_{\omega jt}}{Y_{jt}} = D^\omega \left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right)$$

that determines the solution for each of the points $\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right)$.
It is left to prove that there are unique values for the market aggregates P_{jt}, D_{jt}. For this, we would first

need to check that there is a unique solution for the product PD. For this, we substitute the newly obtained function D^ω into the constraint $F\left(\frac{y_{1jt}}{Y_{jt}} \dots \frac{y_{n_j jt}}{Y_{jt}}\right)$. We then obtain:

$$F\left(D^1\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right) \dots D^{n_j}\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right)\right) = 1$$

Note that all the $p_{\omega jt}$ are given parameters, so this is just a constraint in terms of the univariate function of $P_{jt}D_{jt}$. To determine that this function is an injection and the equation has a unique solution, consider the first derivative in terms of $P_{jt}D_{jt}$.

$$-\sum_\omega \sum_k D_k'^{\omega} \frac{p_{\omega jt}}{P_{jt}D_{jt}} \frac{p_{kjt}}{P_{jt}D_{jt}} \frac{1}{P_{jt}D_{jt}} = -\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right) H'^{-1}\left(\frac{p_{1jt}}{P_{jt}D_{jt}} \dots \frac{p_{n_j jt}}{P_{jt}D_{jt}}\right) \frac{1}{P_{jt}D_{jt}} > 0$$

The first derivative of this function would always be positive due to the fact that the Hamiltonian of the function F is a negative definite matrix and $P_{jt}, D_{jt} > 0$. Hence there is a unique solution for the product $P_{jt}D_{jt}$.

Now we can conclude that for the given $(p_1 \dots p_{n_j})$ there is a unique solution of the product $P_{jt}D_{jt}$ and hence $\left(\frac{y_{1jt}}{Y_{jt}} \dots \frac{y_{n_j jt}}{Y_{jt}}\right)$. It is now trivial to determine P_{jt} and D_{jt} individually using the formulas:

$$P_{jt} = \sum_\omega \frac{y_{\omega jt}}{Y_{jt}} p_\omega$$
$$D_{jt} = \left(\sum_\omega \frac{y_{\omega jt}}{Y_{jt}} F'_\omega\right)^{-1}$$

This concludes the proof of the uniqueness of the solution.

The solution would also be a minimum of the cost-minimization problem since the problem is a minimization of a convex function on a convex set.

B Proposition 2: Decomposition for Superelasticity

Proposition 2 *Superelasticity can be decomposed into direct and market aggregate effect and this decomposition can be expressed in terms of partial elasticities, partial superelasticities, market elasticities, market superelasticities and revenue shares.*

We have used the demand function and introduced a definition of partial elasticity to obtain a decomposition for partial elasticity. Using a definition for partial superelasticity, we can obtain a similar decomposition for the higher-order demand parameters.

For instance, the decomposition for the superelasticity of a general oligopolistic firm would be given by:

$$\eta_{\omega\omega\omega jt} = \frac{\xi^r_{\omega\omega jt}}{\xi_{\omega\omega jt}}\left(\eta^r_{\omega\omega\omega jt} - \sum_k \eta^r_{\omega\omega kjt}\theta_{\omega jt}\right) + \frac{\xi^m_{jt}\lambda_{\omega jt}}{\xi_{\omega\omega jt}}\left(\eta^m_{jt}\lambda_{\omega jt} + (1 - \xi_{\omega\omega jt} - (1-\xi^m_{jt})\lambda_{\omega jt})\right) +$$
$$+ \frac{\theta_{\omega jt}}{\xi_{\omega\omega jt}}\frac{\sum_k \xi^r_{\omega k}}{\theta_{\omega jt}}\sum_k\left[\eta^r_{\omega k\omega} + \sum_m \eta^r_{\omega km}\right] + \frac{\sum_k \xi^r_{k\omega}\lambda_k}{\xi_{\omega\omega jt}}\sum_k \frac{\xi^r_{k\omega}\lambda_k}{\sum_k \xi^r_{k\omega}\lambda_k}\left[\eta^r_{k\omega\omega} - \sum_m \eta^r_{k\omega m}\theta_\omega + \mathbb{I}\{k=\omega\} - \xi_{k\omega} - (1-\xi^m)\lambda_\omega\right] +$$
$$+ \frac{\sum_m\sum_k \xi^r_{km}\lambda_k}{\xi_{\omega\omega jt}}\sum_m\sum_k \frac{\xi^r_{km}\lambda_k}{\sum_m\sum_k \xi^r_{km}\lambda_k}\left[\eta^r_{km\omega} - \sum_l \eta^r_{kml}\theta_\omega + \mathbb{I}\{k=\omega\} - \xi_{k\omega} - (1-\xi^m)\lambda_\omega\right]$$

Note that, again, we only need to estimate partial elasticities, partial superelasticities, market elasticities, and market superelasticities in order to calculate the firm-level superelasticities.

As noted in the main text of the paper, in the Kimball case, the decomposition of the superelasticity would be much simplified.

C Proposition 3: Flexible Price Equilibrium

Proposition 3 *Flexible price equilibrium exists and is unique when*

1. *Demand aggregator satisfies the conditions of the definition 1 and is homogeneous of degree 0 with respect to prices; own price elasticity is positive and other goods at the same market are substitutes, so that all the cross-superelasticities are negative.*

2. *Demand aggregator is Kimball.*

C.1 Case 1

From proposition 1 we know that there is a unique solution for the $\frac{y_{\omega jt}}{Y_{jt}}$ given the individual prices. For the uniqueness of the flexible price equilibrium it is enough to prove that there is a unique optimal price, as determined by the implicit equation:

$$p^*_{\omega jt} = \frac{\xi_{\omega\omega jt}}{\xi_{\omega\omega jt} - 1} mc_{\omega jt}$$

As a first step of the proof, we take logs, obtaining:

$$\log p_{\omega jt} - \log \mu_{\omega jt} = \log mc_{\omega jt}$$

Note that now the left-hand side depends on the prices, while the right-hand side only depends on the marginal costs. We thus have a system of equations that determines the vector of prices $\log \bar{p}_{jt}$ in terms of the vector of marginal costs $\log \bar{mc}_{jt}$. We will be dealing with determining whether there is a solution to this system by employing a strategy similar to the one used in proposition 1. We will consider the matrix of the first derivatives, and determine that under the assumptions that we have made about the aggregator function F it will be a diagonally dominant matrix allowing us to use the general result of Gale and Nikaidô (1965).

The matrix of the first derivatives would be given by:

$$\text{Mat}\left\{\frac{\partial p_{\omega jt}}{\partial p_{kjt}}\right\} = I + D \times A = I + \text{Diag}\left\{\frac{1}{\xi_{\omega\omega jt} - 1}\right\} \times \text{Mat}\left\{\frac{\partial \log \xi_{\omega\omega jt}}{\partial \log p_{kjt}}\right\}$$
$$\text{Mat}\left\{\frac{\partial p_{\omega jt}}{\partial p_{kjt}}\right\} = I + \text{Diag}\left\{\frac{1}{\xi_{\omega\omega jt} - 1}\right\} \times \text{Mat}\left\{\eta_{\omega\omega\omega jt}\right\}$$

Note that the matrix A is a matrix of superelasticities. Due to the homogeneity of the demand structure, elasticity would be homogeneous of degree zero, which would require that the sum of superelasities is equal to zero. This means that the sum of the elements of the $D \times A$ matrix should be equal to zero. Assuming additionally that superelasticities with respect to the prices of the competitors are negative, we obtain the following equation:

$$1 + d_\omega a_{\omega\omega} = 1 - \sum_k d_\omega a_{\omega k} = 1 + \sum_k |d_\omega a_{\omega k}| > \sum_k |d_\omega a_{\omega k}|$$

where d_ω is the ω-th diagonal element of the matrix D and $a_{\omega k}$ is the ωk-th element of the matrix A.
This equation means that the matrix of the first derivatives has a dominant diagonal.
We can now use the result of Gale and Nikaidô (1965) to say that the system of equations has a unique global solution.
This would mean that the optimal pricing equation determines prices uniquely in terms of marginal costs, which ensures the uniqueness of flexible price equilibrium.

C.2 Case 2

In the case of the Kimball Oligopoly, the problem of finding the flexible price equilibrium is much simplified. Elasticity would only depend on the firm's own relative price.

$$p_{\omega j} = \frac{\xi_{\omega j}}{\xi_{\omega j} - 1} mc_{\omega j}$$
$$\xi_{\omega j} = \xi\left(\frac{p_{\omega j}}{P_j D_j}\right)$$

Hence the first derivative matrix of the elasticity with respect to the relative price is going to be diagonal. This means that relative price can be expressed as a function of the fraction of marginal costs and market aggregates:

$$p_{\omega j} = p\left(\frac{mc_{\omega j}}{P_j D_j}\right) P_j D_j$$

We can then substitute relative prices into the aggregator function F to obtain the unique values for the product $P_j D_j$.

$$F\left(\dots F_{\omega}'^{-1}\left(p\left(\frac{mc_{\omega j}}{P_j D_j}\right)\right)\dots\right) = 1$$

The solution would be unique since the derivative of the left-hand side is always positive. Hence, there would be a unique solution for the prices $p_{\omega j}$.

Note additionally that since the derivative of the left-hand is always positive, if the marginal costs of each of the firms change by the same multiplier, it will have to be the case that $P_j D_j$ change by the same multiplier, meaning that there is no effect on the relative prices.

D Proposition 4: External Validity

Proposition 4 *In the Kimball oligopoly case without apriori asymmetry, partial elasticity of a firm with the same output share* $\frac{y_\omega}{Y}$ *is the same at any market*

Since the flexible price equilibrium is unique, we would have the unique set of $\frac{y_{\omega j}}{Y_j}$ for each of the markets. Due to demand invertibility, output shares will uniquely determine relative prices. Due to the Kimball structure, partial elasticity for each of the firms would only be determined by its own relative price.

$$\xi_{\omega}^{r} = \xi\left(\frac{p_\omega}{PD}\right)$$

In the case of no apriori asymmetry between markets or firms, the function determining the partial elasticity would be the same for different markets. This would mean that it is enough to fix the output share for the partial elasticities to be the same.

We can allow for limited asymmetry on the firm level. In this case, the elasticity function would become:

$$\xi_{\omega}^{r} = \xi\left(A_\omega, \frac{p_\omega}{PD}\right)$$

Note that apriori asymmetry between firms is indexed by the parameter A. In this case, in order for our estimation to still hold, we need to additionally control for anything that influences the variation in the asymmetry parameter A_ω.

As an example of the asymmetry that can be accounted for in the applications, consider a variation of a customer acquisition model based on Ravn et al. (2006). We perform this exercise as one of the robustness checks.

In this case, the aggregate demand function would be given by:

$$y_{\omega jt} = \hat{y}_{\omega jt} + \kappa y_{\omega jt-1}$$
$$\hat{y}_{\omega jt} = \Upsilon^{\omega j}\left(\frac{p_{1j}}{P_j D_j} \cdots \frac{p_{1n}}{P_j D_j}\right)\Upsilon^{j}\left(\frac{P_{jt}}{P_t D_t}\right) Y_t$$

Note that here, we take into account the possibility of the firm developing a customer base over time - a share κ of consumers is going to be locked and will not be able to stop buying from a firm even if the price is no longer satisfactory. As a simple "patch" we assume that new consumers follow the demand function that we discussed in the main section.

In this case, partial elasticity can be expressed as:

$$\xi_{\omega}^{r} = \hat{\xi}_{\omega}^{r}\left(1 - \frac{y_{\omega jt-1}}{y_{\omega jt}}\right) = \hat{\xi}_{\omega}^{r}\left(1 - \frac{\nu_{\omega jt-1} y_{\omega jt-1}}{Y_{jt-1}} \frac{Y_{jt}}{\nu_{\omega jt} y_{\omega jt}}\right)$$

Note that the equation of the partial elasticity, apart from the term similar to the one coming from the previous section, includes an additional one that depends on the previous period's market share. When the market share was larger in the previous period, the firm's elasticity is smaller today. In other words, if the price was smaller in the previous period, the elasticity is smaller today.

In this case, the additional control is needed to account for the dynamic effect. The one that needs to be included is the previous period's prices and market aggregates. When these controls are added, we would once again get the correct partial elasticity estimates.

E Proposition 5 and 6: Pass-through and pass-through elasticity

Proposition 5: *When elasticity is a convex function of prices, pass-through is smaller for larger firms.*
Proposition 6: *When elasticity is a convex function of prices, pass-through of marginal costs into prices is asymmetric and is larger for the marginal cost increases.*

Note that in our static model, larger firms would always have lower prices. Then consider the derivative of pass-through with respect to price.

$$\frac{\partial[\text{Pass-through}]}{\partial p_{\omega}} = \left(\frac{1}{p_{\omega}}\right)\left(\frac{\eta_{\omega}(\xi_{\omega} - 1)}{(\xi_{\omega} - 1 + \eta_{\omega})^2}\right)\left(\eta_{\omega}\frac{\xi_{\omega}}{\xi_{\omega} - 1} - \psi_{\omega}\right)$$

This derivative is positive as long as ψ is negative, which would be the case for the case when elasticity is a convex function of prices.

The pass-through of positive marginal cost shocks will be larger when there is a positive elasticity of pass-through to the marginal costs. Recall the formula for the elasticity of pass-through with respect to marginal costs:

$$\frac{\partial \log[\text{Pass-through}]}{\partial \log mc_\omega} = \left(\frac{\xi_\omega - 1}{\xi_\omega - 1 + \eta_\omega}\right)^2 \left(\frac{\eta_\omega}{\xi_\omega - 1}\right)\left(\eta_\omega \frac{\xi_\omega}{\xi_\omega - 1} - \psi_\omega\right)$$

It will be positive whenever ψ is negative. So it is once again enough to require that elasticity is the convex function of prices.

F Formulas

In this section, we list all the formulas used for the computation of the final results. Note that the market-level and time indexes are omitted for simplicity. All the formulas are given for the case where both firm and market-level demands are Kimball since this is the most empirically relevant case.
Primitives:

$$\xi^r_\omega = -\frac{\partial y_\omega}{\partial p_\omega}\frac{p_\omega}{y_\omega}\bigg|_{P,Y,D} \qquad \eta^r_\omega = \frac{\partial \xi^r_\omega}{\partial p_\omega}\frac{p_\omega}{\xi^r_\omega}\bigg|_{P,Y,D} \qquad \psi^r_\omega = \frac{\partial \eta^r_\omega}{\partial p_\omega}\frac{p_\omega}{\eta^r_\omega}\bigg|_{P,Y,D}$$
$$\xi^m = -\frac{\partial Y}{\partial P}\frac{P}{Y} \qquad \eta^m = \frac{\partial \xi^m}{\partial P}\frac{P}{\xi^m}$$

Elasticity:

$$E^r = \sum_k \xi^r_k \lambda_k \qquad \theta_\omega = \frac{\xi^r_\omega \lambda_\omega}{E^r}$$
$$\xi_\omega = \xi^r_\omega + \xi^m \lambda_\omega - \xi^r_\omega \theta_\omega$$
$$\xi_{k\omega} = \xi^m \lambda_\omega - \xi^r_k \theta_\omega, \;\; k \neq \omega$$

Superelasticity:

$$\Lambda_\omega = 1 - \xi^r_\omega - \lambda_\omega + \xi^r_\omega \theta_\omega$$
$$\Lambda_{k\omega} = -\lambda_\omega + \xi^r_k \theta_\omega, \;\; k \neq \omega$$
$$N^r_\omega = \sum_k \theta_k(-\eta^r_k \theta_\omega - \lambda_\omega + \xi^r_k \theta_\omega) + \theta_\omega(\eta^r_\omega + 1 - \xi^r_\omega)$$
$$\bar{\theta}_\omega = \eta^r_\omega(1 - \theta_\omega) + \Lambda_\omega - N^r_\omega$$
$$\bar{\theta}_{k\omega} = -\eta^r_k \theta_\omega + \Lambda_{k\omega} - N^r_\omega$$
$$\eta_\omega = \frac{\xi^r_\omega}{\xi_\omega}\eta^r_\omega(1 - \theta_\omega) + \frac{\xi^m \lambda_\omega}{\xi_\omega}(\eta^m \lambda_\omega + \Lambda_\omega) - \frac{\xi^r_\omega \theta_\omega}{\xi_\omega}(\eta^r_\omega(1 - \theta) + \bar{\theta}_\omega)$$
$$\eta_{kk\omega} = -\frac{\xi^r_k}{\xi_k}\eta^r_k \theta_\omega + \frac{\xi^m \lambda_k}{\xi_k}(\eta^m \lambda_\omega + \Lambda_{k\omega}) - \frac{\xi^r_k \theta_k}{\xi_k}(-\eta^r_k \theta_\omega + \bar{\theta}_{k\omega})$$
$$\eta_{\omega m} = \frac{\xi^m \lambda_\omega}{\xi_\omega}\eta^m$$

Second-order effect:

$$\bar{\Lambda}_\omega = -\frac{\xi^r_\omega}{\Lambda_\omega}\eta^r_\omega(1-\theta_\omega) - \lambda_\omega + \frac{\xi^r_\omega\theta_\omega}{\Lambda_\omega}(\eta^r_\omega(1-\theta_\omega)+\bar{\theta}_\omega)$$

$$\bar{N}^r_\omega = \sum_k \frac{\theta_k(-\eta^r_k\theta_\omega - \lambda_\omega + \xi^r_k\theta_\omega)}{N^r_\omega}\bar{\theta}_{k\omega} + \frac{\theta_\omega(-\eta^r_\omega\theta_\omega - \lambda_\omega + \xi^r_\omega\theta_\omega)}{N^r_\omega}(\eta^r_\omega + 1 - \xi^r_\omega) +$$
$$\sum_k \frac{\theta_k}{N^r_\omega}(-\eta^r_k\theta_\omega(-\psi^r_k\theta_\omega + \bar{\theta}_\omega) - \lambda_\omega\bar{\Lambda}_\omega + \xi^r_k\theta_\omega(-\eta^r_k\theta_\omega + \bar{\theta}_\omega)) + \frac{\theta_\omega}{N^r_\omega}(-\eta^r_\omega\theta_\omega\psi^r_\omega + \xi^r_\omega\theta_\omega\eta^r_\omega) +$$
$$\frac{\theta_\omega(\eta^r_\omega + 1 - \xi^r_\omega)}{N^r_\omega}\bar{\theta}_\omega + \frac{\theta_\omega}{N^r_\omega}(\eta^r_\omega\psi^r_\omega(1-\theta_\omega) - \xi^r_\omega\eta^r_\omega(1-\theta_\omega))$$

$$\bar{\bar{\theta}}_\omega = \frac{\eta^r(1-\theta_\omega)}{\bar{\theta}_\omega}\psi^r_\omega(1-\theta_\omega) - \eta^r\theta_\omega + \frac{\Lambda_\omega}{\bar{\theta}_\omega}\bar{\Lambda}_\omega - \frac{N^r_\omega}{\bar{\theta}_\omega}\bar{N}^r_\omega$$

$$\psi_\omega = \frac{\xi^r_\omega\eta^r_\omega(1-\theta_\omega)}{\xi_\omega\eta_\omega}\left(\eta^r(1-\theta_\omega) + \psi^r_\omega(1-\theta_\omega) - \frac{\theta_\omega}{1-\theta_\omega}\bar{\theta}_\omega - \eta_\omega\right) +$$
$$\frac{\xi^m\lambda_\omega(\eta^m\lambda_\omega + \Lambda_\omega)}{\xi_\omega\eta_\omega}(\eta^m\lambda_\omega + \Lambda_\omega - \eta_\omega) + \frac{\xi^m\lambda_\omega}{\xi_\omega\eta_\omega}(\eta^m\lambda_\omega(\psi^m\lambda_\omega + \Lambda_\omega) + \Lambda_\omega\bar{\Lambda}_\omega) -$$
$$\frac{\xi^r_\omega\theta_\omega(\eta^r_\omega(1-\theta_\omega+\bar{\theta}_\omega))}{\xi_\omega\eta_\omega}(\eta^r_\omega(1-\theta_\omega) + \theta_\omega - \eta_\omega) - \frac{\xi^r_\omega\theta_\omega}{\xi_\omega}(\eta^r_\omega\psi^r_\omega(1-\theta_\omega)^2 - \eta^r_\omega\theta_\omega\bar{\theta}_\omega + \bar{\theta}_\omega\bar{\bar{\theta}}_\omega)$$

G Market-level First Stage

	Price $\tilde{P}_{jt}$	
	(1)	(2)
$\tilde{z}_{jt}$	-0.110***	-0.147***
	(0.010)	(0.011)
Obs.	506,798	506,798
Time		✓
Market		✓

Table 8: Market-level First Stage *This table shows the result of the first-level regression on the market level:* $\tilde{P}_{jt} = \alpha\tilde{z}_{jt} + \gamma + \delta_{jt}$*, where* α *is the first-stage coefficient, the parameter of interest. The granular instrument is given by:* $\tilde{z}_{jt} = \sum_\omega \lambda_{\omega jt-1}\tilde{z}_{\omega jt}$*, the market price is given by:* $\tilde{P}_{jt} = \sum_\omega \lambda_{\omega jt-1}\tilde{p}_{\omega jt}$ *due to the homogeneity of demand assumption. The regression is performed without fixed effects and then again with market and time fixed effects. Standard errors are clustered at the market level. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

H Assumption on the market share distribution aggregates

For each of the considered HHI levels we calculate other relevant market share distribution characteristics as an average of that characteristic across the markets with this HHI level. For instance, the resulting $H^3 = \sum\lambda^3_\omega$ and $H^4 = \sum\lambda^4_\omega$ are given in the Figure 7.

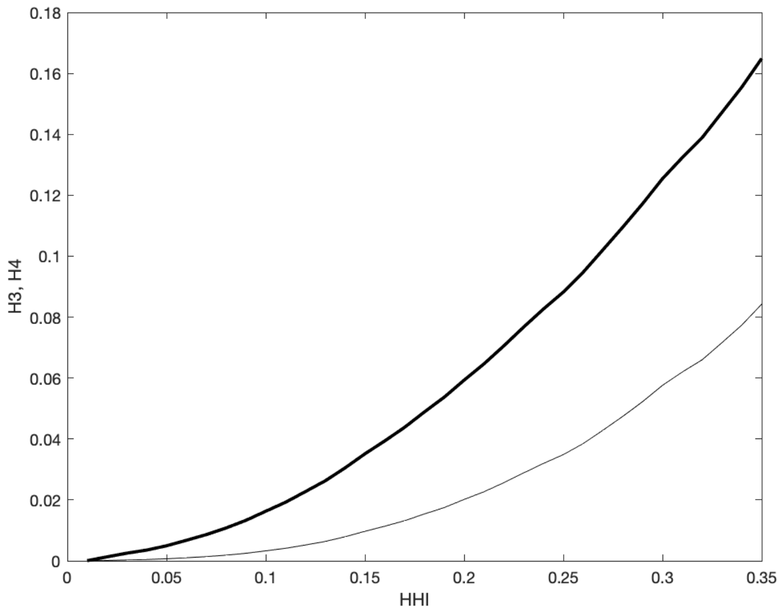

Figure 7: Assumptions about H^3 and H^4 *This figure shows the assumptions about H^3 and H^4. The x axis represents the concentration index $HHI = \sum_\omega \lambda_\omega^2$ with λ_ω being the revenue market share of a firm ω. The bold line shows $H^3 = \sum_\omega \lambda_\omega^3$ and the thin line shows $H^4 = \sum_\omega \lambda_\omega^4$.*

H.1 Dynamic Demand

In this section, we discuss the possibility of persistent effects of prices on demand elasticity. We estimate this dependence from the following Jorda-style regression:

$$\tilde{y}_{r\omega jt+k} = (-\xi^r_{\omega\omega})\tilde{p}_{r\omega jt+k} + \alpha_1 \tilde{p}_{r\omega jt} + (-\xi^r_{\omega\omega})(\alpha_2)\tilde{p}_{r\omega jt+k} \times \tilde{p}_{r\omega jt} + \gamma + \delta_{r\omega jt}$$

Where $\xi^r_{\omega\omega}$ is the own price partial elasticity, α_1 is the persistent effect of previous period prices, α_2 is the degree of dependence of current elasticity on the previous period elasticity, both parameters of interest, and $\gamma_{g_\omega jt}$ is the firm-size group-market-time fixed effect, with subscript g_ω reflecting the group to which firm ω belongs.

The result of the estimation is shown in Table 9. First, note that the past-period price has a significant effect on quantities, even after two months with 1% higher prices today, decreasing quantity one period ahead by 1.25% in the first period and by 0.75% two months ahead. The more important effect, though, is the effect on elasticity, obtained from the interaction coefficient between past and contemporaneous prices. We find that a 1% higher price today increases elasticity by 1.6% one period ahead and by 0.80% two months ahead. The effect is persistent and is both noticeable and significant even after two months.

This result suggests that there might be a persistent effect of current prices on future elasticity. This notes a potential source of bias in our baseline estimation of the elasticity. To address this issue we consider

	Quantity $\tilde{y}_{\omega jt+k}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\tilde{p}_{\omega jt+k}$		-9.071***	-8.025***	-7.684***	-7.655***	-7.645***	-7.579***	-7.541***	-7.506***
		(0.420)	(0.305)	(0.269)	(0.269)	(0.266)	(0.256)	(0.254)	(0.254)
$\tilde{p}_{\omega jt}$	-8.269***	-1.250***	-1.027***	-0.801***	-0.881***	-0.861***	-0.731***	-0.668***	-0.749***
	(0.311)	(0.092)	(0.074)	(0.064)	(0.064)	(0.064)	(0.056)	(0.055)	(0.057)
$\tilde{p}_{\omega jt+k} \times \tilde{p}_{\omega jt}$		-14.638***	-9.471***	-8.161***	-5.861***	-7.330***	-7.608***	-7.236***	-6.019***
		(3.054)	(2.865)	(2.567)	(1.924)	(2.097)	(2.043)	(2.155)	(1.574)
α_2		1.614***	1.180***	1.062***	0.766***	0.959***	1.004***	0.960***	0.802***
		(0.080)	(0.111)	(0.101)	(0.059)	(0.068)	(0.067)	(0.075)	(0.041)
Obs.	7,098,037	6,825,157	6,765,561	6,721,725	6,681,816	6,641,057	6,598,780	6,561,938	6,530,723
F	396	271	279	274	267	300	287	307	292
Group × Market × Time	✓	✓	✓	✓	✓	✓	✓	✓	✓
Producer		✓	✓	✓	✓	✓	✓	✓	✓
Period	0	1	2	3	4	5	6	7	8

Table 9: Dynamic Effect of the own Price *The results of the dynamic effects of the own price estimation from the regression* $\tilde{y}_{r\omega jt+k} = (-\xi^r_{\omega\omega})\tilde{p}_{r\omega jt+k} + \alpha_1 \tilde{p}_{r\omega jt} + (-\xi^r_{\omega\omega})(\alpha_2)\tilde{p}_{r\omega jt+k} \times \tilde{p}_{r\omega jt} + \gamma + \delta_{r\omega jt}$ *where* $\xi^r_{\omega\omega}$ *is the own price partial elasticity,* α_1 *is the persistent effect of previous period price, and* α_2 *is the degree of dependence of current elasticity on the previous period elasticity, both parameters of interest. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$. *Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

an alternative version of the main regressions enriched with four lags of prices that are instrumented with lagged granular instruments. The structure of regressions and tables follows the ones in the main text of the paper. The results are similar to the ones that we get in the baseline regression, suggesting that the dynamic demand is small and does not cause significant bias.

	Price $\tilde{p}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Granular Instrument $\tilde{z}_{\omega jt}$	-0.100***	-0.107***	-0.122***	-0.115***	-0.120***	-0.125***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
Market Share $\lambda_{\omega jt-1}$				0.011***	0.016***	0.030***
				(0.001)	(0.001)	(0.003)
Obs.	9,436,021	9,433,876	6,222,711	7,131,643	7,121,839	6,222,296
F	143	139	139	129	135	121
Market × Time		✓			✓	
Group × Market × Time			✓			✓
Producer				✓	✓	✓

Table 10: First-Stage Regression with More Controls *The results of the first-stage regression* $\tilde{p}_{\omega jt} = \sum_{h=0}^{3} \alpha_h \tilde{z}_{\omega jt-h} + \gamma_{gjt} + \varepsilon_{\omega jt}$ *where* α_0 *is the first stage coefficient, the parameter of interest. Granular Instrument is given as a weighted sum of the regional regression residuals:* $z_{\omega jt} = \sum_k \lambda_{k\omega jt} \tilde{\nu}_{r\omega jt}$. *Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$. *Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

	Quantity $\tilde{y}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Partial Cross-Elasticity $\xi^r_{\omega k}$	17.254***	1.654***	7.425	-0.662***	-1.469*	0.789
	(3.620)	(0.618)	(14.467)	(0.128)	(0.813)	(1.616)
Obs.	9,364,820	9,362,724	3,474,305	4,318,561	4,310,833	3,474,252
Market × Time		✓			✓	
Group × Market × Time			✓			✓
Producer				✓	✓	✓
$\lambda_{\omega jt}$				✓	✓	✓

Table 11: Partial Cross-Elasticity Test with More Controls *The results of the partial cross-elasticity test with additional controls. The calculated statistic is equal to the partial cross-elasticity when the strategic complementarity is equal to zero, otherwise, the statistic is* $(1 + [\text{Strategic Complementarity}])\xi^r_{\omega k}$*. Due to the fact that strategic complementarity is larger then* -1*, in both cases the statistic equal to zero indicates zero partial cross elasticity. The estimation is conducted using GMM. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$*. The controls include four lags of prices. Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are clustered at the market level. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

	Quantity $\tilde{y}_{\omega jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Price $\tilde{p}_{\omega jt}$	-1.706***	-1.952***	-201.931	-9.773***	-12.313***	-8.663***
	(0.079)	(0.078)	(492.315)	(0.502)	(0.926)	(0.347)
Market Share $\lambda_{\omega jt-1}$					0.116***	0.287***
					(0.014)	(0.024)
	OLS	OLS	GIV	GIV	GIV	GIV
Obs.	9,400,778	6,223,587	9,368,713	6,214,983	7,122,336	6,214,579
F	532	749	2	179	129	176
Group × Market × Time		✓		✓		✓
Producer					✓	✓

Table 12: Partial Elasticity with More Controls *The results of the elasticity estimation from the regression* $\tilde{y}_{\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{\omega jt} + \sum_{h=1}^{3} \alpha_h \tilde{p}_{\omega jt-h} + \gamma_{gjt} + \varepsilon_{\omega jt}$ *where* $\xi^r_{\omega\omega}$ *is the own price partial elasticity, the parameter of interest. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$*. Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

	Quantity $\tilde{y}_{\omega jt}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Price $\tilde{p}_{\omega jt}$	-1.824***	-1.914***	-2.039***	-220.527	-13.638***	-8.682***	-10.634***	-8.719***	-7.668***
	(0.071)	(0.061)	(0.069)	(578.452)	(0.967)	(0.372)	(0.644)	(0.336)	(0.255)
Price^2 $\tilde{p}^2_{\omega jt}$	-0.813***	-0.717***	-0.937***	7.855	-15.688***	-10.322***	-11.776***	-11.697***	-11.121***
	(0.098)	(0.080)	(0.134)	(49.689)	(3.114)	(1.673)	(2.045)	(1.664)	(1.532)
Market Share $\lambda_{\omega jt-1}$							0.095***	0.160***	0.233***
							(0.011)	(0.011)	(0.026)
$\eta_{\omega\omega\omega}$	0.446	0.375	0.459	-0.036	1.150	1.189	1.107	1.342	1.450
	0.003	0.002	0.004	0.041	0.035	0.028	0.028	0.028	0.032
	OLS	OLS	OLS	GIV	GIV	GIV	GIV	GIV	GIV
Obs.	9,400,778	9,389,107	6,223,587	9,368,713	9,366,397	6,214,983	7,122,336	7,112,355	6,214,579
F	419	580	601	1	125	195	134	182	204
Market × Time		✓			✓			✓	
Group × Market × Time			✓			✓			✓
Producer							✓	✓	✓

Table 13: Partial Superelasticity with More Controls *The results of the partial superelasticity estimation from the regression* $\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega j})\tilde{p}_{r\omega jt} + [(-\xi^r_{\omega\omega j}) \times (\eta^r_{\omega\omega\omega})]\tilde{p}^2_{r\omega jt} + \sum_{h=1}^{3} \alpha_h \tilde{p}_{\omega jt-h} + \gamma + \delta_{r\omega jt}$ *where* $\xi^r_{\omega\omega}$ *is the own price partial elasticity,* $\eta^r_{\omega\omega\omega}$ *is the own price superelasticity, the parameter of interest. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$*. Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

	Quantity $\tilde{y}_{\omega jt}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Price $\tilde{p}_{\omega jt}$	-1.717***	-1.853***	-1.893***	-28.427***	-12.390***	-10.210***	-12.945***	-10.393***	-8.983***
	(0.095)	(0.076)	(0.082)	(7.672)	(0.840)	(0.557)	(1.014)	(0.515)	(0.387)
Price $\tilde{p}_{\omega jt}$ × Market Share $\lambda_{\omega jt-1}$	-1.611***	-2.008***	-2.556***	37.005***	17.797***	18.380***	13.197***	11.463***	12.674***
	(0.439)	(0.319)	(0.507)	(12.119)	(2.702)	(2.939)	(2.291)	(1.815)	(2.204)
Market Share $\lambda_{\omega jt-1}$	0.098***	0.179***	0.209***	0.265***	0.268***	0.323***	0.101***	0.181***	0.268***
	(0.007)	(0.005)	(0.013)	(0.079)	(0.017)	(0.027)	(0.014)	(0.011)	(0.023)
	OLS	OLS	OLS	GIV	GIV	GIV	GIV	GIV	GIV
Obs.	7,140,370	7,124,508	6,223,587	7,122,589	7,112,610	6,214,983	7,122,336	7,112,355	6,214,579
F	395	704	622	34	123	143	108	158	165
Market × Time		✓			✓			✓	
Group × Market × Time			✓			✓			✓
Producer							✓	✓	✓

Table 14: Partial Elasticity Heterogeneity with More Controls *The results of the across-share heterogeneity estimation from the regression* $\tilde{y}_{r\omega jt} = (-\xi^r_{\omega\omega})\tilde{p}_{r\omega jt} + \alpha\tilde{p}_{r\omega jt} \times \lambda + \sum_{h=1}^{3} \alpha_h \tilde{p}_{\omega jt-h} + \gamma + \delta_{r\omega jt}$ *where* $\xi^r_{\omega\omega}$ *is the own price partial elasticity,* α *is the degree of heterogeneity of elasticity with respect to firm shares, the parameter of interest. Market share is calculated as a revenue share* $\lambda_{\omega jt} = \frac{p_{\omega jt} y_{\omega jt}}{P_{jt} Y_{jt}}$*. Regression is performed for firm group-market-time fixed effects and with additional firm fixed effects. Standard errors are double clustered at the firm and market levels. Asterisks mark significance levels: *** p<0.01, ** p<0.05, * p<0.1.*

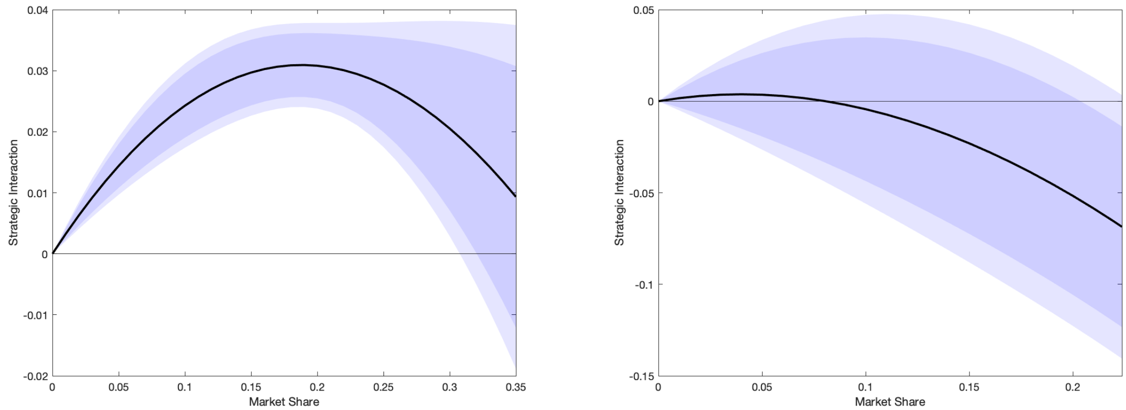

Figure 8: Strategic Complementarity *This figure shows the result of the strategic complementarity estimation in dependence of market shares of the competitor. The left panel considers the strategic complementarity of a firm with 1% market share, the right panel considers the strategic complementarity of a firm with 30% market share. HHI is fixed at 0.14, so the competitor shares on the right panel are only going up to the maximum shares allowed under this condition. Standard errors are calculated according to the delta method. Shaded regions represent 95% and 99% confidence intervals.*

I Strategic Complementarity

In this section, we discuss strategic complementarity. There are two important details to note. First, in our case, strategic complementarity would be different for different pairs of firms. The degree of strategic complementarity between a small and a large firm, between two small firms, between two large firms, or between large and small firms would all be different. Additionally, the level of competitiveness of the market will significantly alter the results, especially for the interaction with larger firms or for the larger firms themselves.

In general, in an oligopolistic framework, strategic complementarity would be a combination of the two effects. The first one is the one that is usually cited in the literature - firms would want to increase prices when their competitors are increasing prices, in order not to lose their market share. This would imply a positive strategic complementarity. This effect would drive the response of small and middle-sized firms to the prices of their competitors. Response to the smallest competitor's prices would be smaller than the response to the middle-sized competitors.

On the other hand, in the oligopolistic framework, there could be an alternative consideration - if all the large firms in the market increase the prices, the price of the market increases as well, which causes the whole market to shrink as compared to the economy as a whole. This would mean that when a large firm increases its price, it might sometimes be preferable for its large competitor to not increase its own price. This last consideration would dominate only when we consider the interaction between two large firms in the same market.

J Group-wise Partial Elasticity

Keeping in mind that dependence of elasticity on the market share might not be linear, we adopt a different strategy and instead of estimating a linear dependence of partial elasticity on the market share, estimate partial elasticity for different firm groups:

$$\tilde{y}_{\omega jt} = (-\xi^0_{\omega\omega})\tilde{p}_{\omega jt} + \sum_{g=1}^{4} \alpha_g \tilde{p}_{\omega jt} + \gamma + \delta\omega jt$$

We consider four groups. The lowest group includes market shares up to the 50th percentile of the market share distribution and with market shares up to approximately 1%. The second group includes shares up to the 90th percentile of the market share distribution and market shares up to approximately 9%. The third group includes shares up to the 99th percentile of the market share distribution and market shares up to approximately 32%. The last group includes top firms up from 99% of the market share distribution. The results remain similar if we include additional in-between groups and have groups corresponding to 10%, 25%, 50%, 75%, 90%, 95%, and 99% of the market share distribution. We decrease the number of groups to gain in the estimation precision. For the same reason, we keep the linear model as our baseline.

$$\tilde{y}_{\omega jt} = \underset{(0.495)}{-9.390}\,\tilde{p}_{\omega jt} + \underset{(0.454)}{2.858}\tilde{p}_{\omega jt} \times \mathbb{I}\{\omega \in G_2\} + \underset{(0.625)}{4.259}\tilde{p}_{\omega jt} \times \mathbb{I}\{\omega \in G_3\} + \underset{(0.699)}{4.674}\tilde{p}_{\omega jt} \times \mathbb{I}\{\omega \in G_4\} + \gamma + \delta\omega jt$$

In this section, we repeat all the similar calculations for elasticity, markup, superelasticity, pass-through, and elasticity of pass-through with respect to marginal costs.

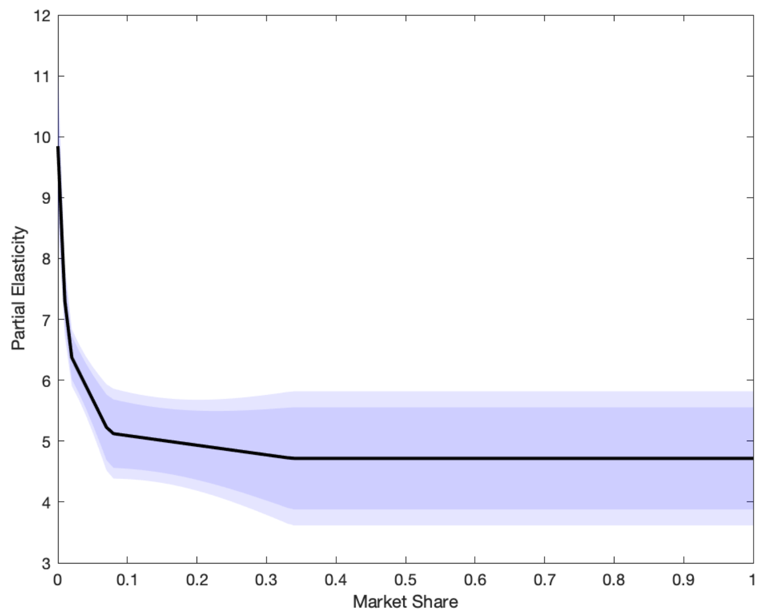

Figure 9: Group-wise Partial Elasticity *This figure shows the results for the estimation of the partial elasticity in dependence of market shares for the group-wise regression. Values for the regions between the estimation points are filled in according to the linear interpolation. Standard errors are calculated according to the delta method. Shaded regions represent the 95% and 99% confidence intervals.*

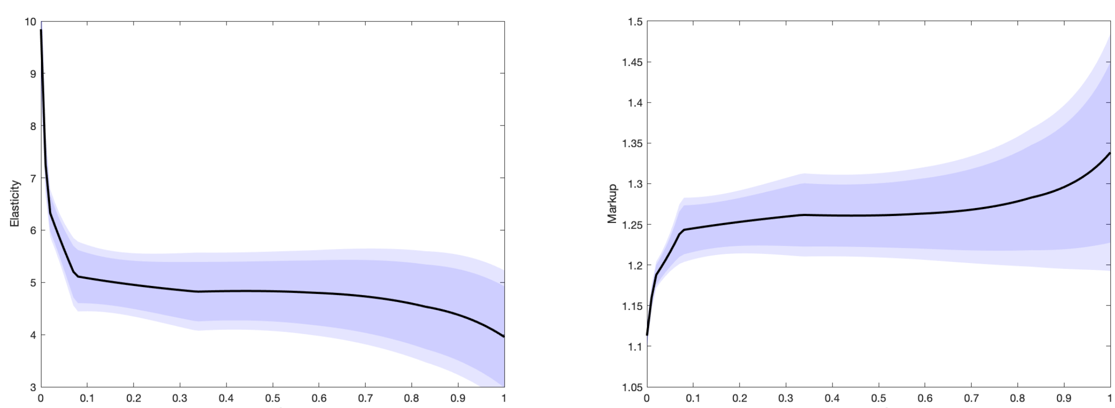

Figure 10: Group-wise Elasticity and Markup *This figure shows the results for the estimation of the elasticity and markup in dependence of market shares for the group-wise regression. The left panel shows elasticity, the right panel shows markup. Standard errors are calculated according to the delta method. Shaded regions represent the 95% and 99% confidence intervals.*

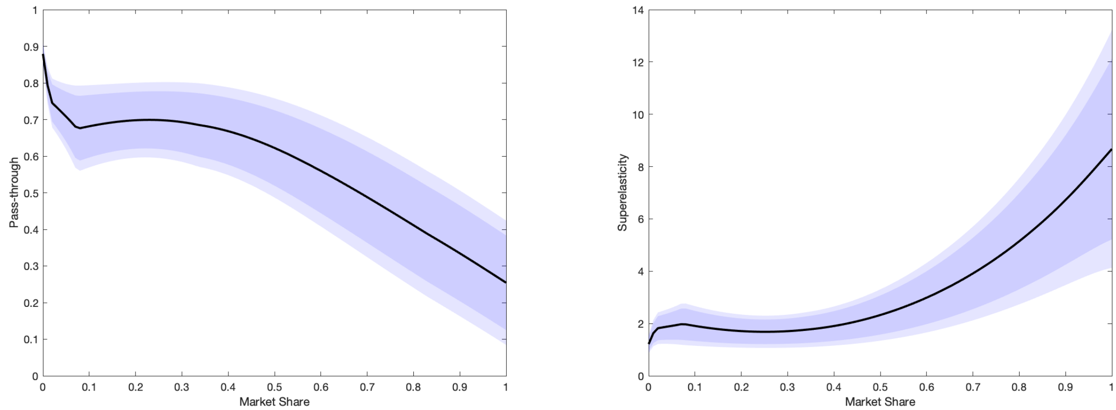

Figure 11: Group-wise Pass-through and Superelasticity *This figure shows the results for the estimation of the pass-through and superelasticity in dependence of market shares for the group-wise regression. The left panel shows pass-through, the right panel shows superelasticity. Standard errors are calculated according to the delta method. Shaded regions represent the 95% and 99% confidence intervals.*

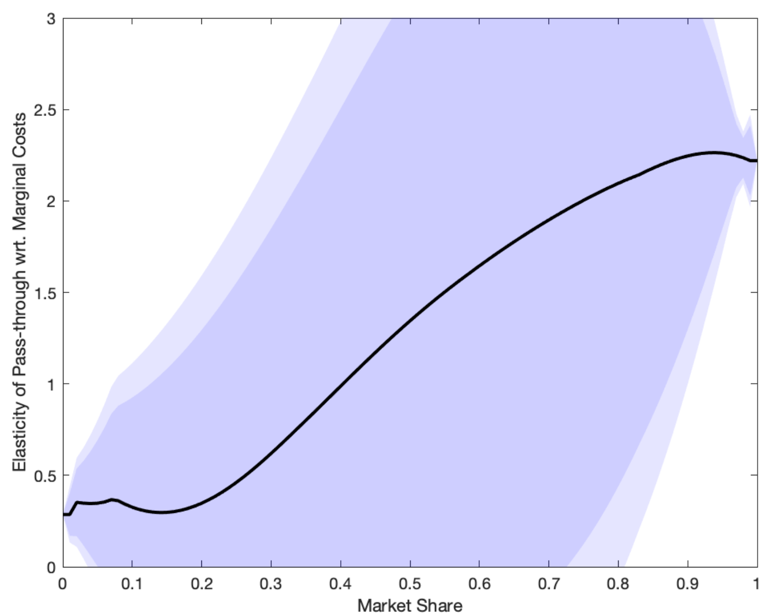

Figure 12: Group-wise Pass-through Elasticity *This figure shows the results for the elasticity of pass-through with respect to marginal cost shock in dependence of market shares for the group-wise regression. Standard errors are calculated according to the delta method. Shaded regions represent the 95% and 99% confidence intervals.*

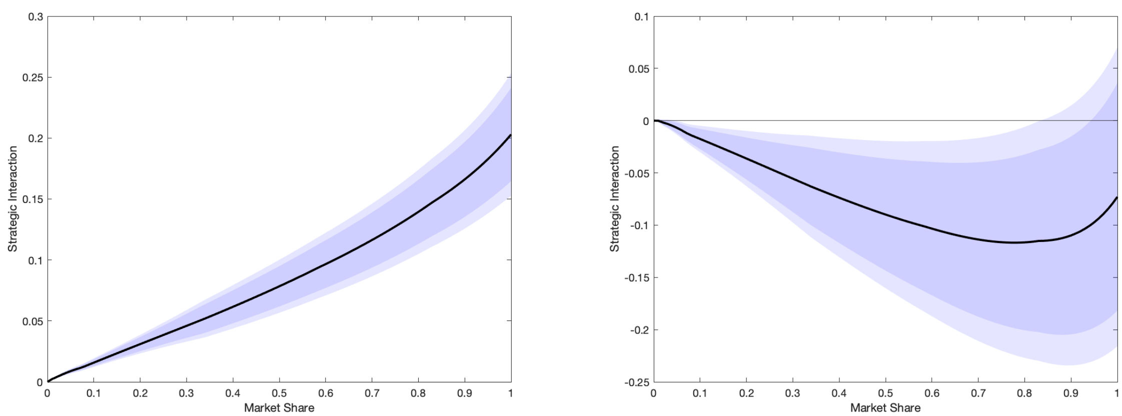

Figure 13: Group-wise Strategic Complementarity *This figure shows the results for the strategic complementarity in dependence of market shares of the competitor for the group-wise regression. The left panel shows the strategic complementarity for a firm with 1% market share, the left panel show strategic complementarity for a firm with 30% market share. Standard errors are calculated according to the delta method. Shaded regions represent the 95% and 99% confidence intervals.*