# Markups Across the Income Distribution: Measurement and Implications

Kunal Sangani
Harvard University*

April 12, 2024

**Abstract**

I examine the relationship between customer income and firm markups using data on household transactions and wholesale costs. Over the observed purchases, high-income households pay 15pp higher retail markups than low-income households. Half of the markup gap is due to differences in markups paid at the same store. Markups paid by a household also increase with the income of other buyers. A model in which household search intensity depends on opportunity cost of time can account for these facts. Consistent with the model's predictions, I document that retail markups across cities rise with both per-capita income and inequality. Through the lens of the model, changes in the income distribution since 1950 account for a 12pp rise in retail markups, with 25 percent of the increase due to growing income dispersion. This rise in markups consists of both within-firm markup increases and a reallocation of sales to high-markup firms.

Keywords: Markups, search, income distribution.

*Email: 8ksangani@gmail.com. First version online: March 2022. I am grateful to Gabriel Chodorow-Reich, Andrei Shleifer, and Ludwig Straub for their excellent advice and support. This paper also benefited from comments by David Baqaee, Adrien Bilal, Xavier Gabaix, Deivy Houeix, Erik Hurst, Kiffen Loomis, Alexander MacKay, Pierfrancesco Mei, Guido Menzio, Namrata Narain, Pascual Restrepo, Peleg Samuels, Lawrence Summers, Adi Sunderam, and participants at the NBER Summer Institute 2022 and SED 2023. This paper contains my 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 author 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.

# 1 Introduction

A growing body of evidence suggests that average markups in the U.S. economy are rising (De Loecker et al. 2020; Autor et al. 2020; Barkai 2020; Gutiérrez 2017). Many of the mechanisms put forward to explain this phenomenon attribute the rise in markups to changes in the supply side of the economy, such as a decline in antitrust enforcement (Gutiérrez and Philippon 2018), the rise of superstar firms (Autor et al. 2017), or structural technological change (De Loecker et al. 2021).

How changes in the demand side of the economy may contribute to the rise in markups is less studied.[1] A large literature in industrial organization and trade finds that price sensitivity tends to decline with income (see e.g., Nevo 2001; Handbury 2021; Auer et al. 2022). If this is the case, a shift in the composition of demand toward high-income households should lead to a decline in aggregate price sensitivity and hence a rise in markups. Yet, the magnitude of this force and its contribution to the rise in markups are unclear.

This paper estimates the relationship between household income and firms' markups, and explores the role of the income distribution in shaping markups across space and time. The first part of the paper is empirical: using rich data on household purchases and wholesale prices, I quantify how retail markups vary with an individual household's income and with aggregate income. The second part of the paper develops a macroeconomic model that can account for the evidence and uses the model to analyze how retail markups evolve with changes to the income distribution.

I construct a dataset of retail markups—hereafter referred to as markups for brevity—by pairing transaction-level data on household purchases with data on wholesale costs faced by retailers. My use of price and cost data to measure retail markups follows Gopinath et al. (2011) and Anderson et al. (2018), who argue that, since rent and labor are fixed at short horizons, merchandise costs are a natural proxy for the marginal costs faced by retailers.[2] The merged dataset includes markups for over 26 million transactions made in a single year. Relative to using data from a single retailer, this merged dataset has two advantages: (1) it includes all household expenditures in tracked product categories and thus captures substitution across retailers; and (2) it includes detailed demographic information for households, typically not available in standalone retailer data.

[1] Exceptions include Bornstein (2021), who studies how aging demographics affect firms' markups, and Döpper et al. (2021), who attribute rising markups in scanner data to a secular decline in price sensitivity.

[2] Two other ways markups are measured in the literature are demand estimation and production function estimation. In Appendix E, I show that markups recovered from demand estimation are strongly correlated with retail markups in my dataset and show that markups estimated using the production function approach by De Loecker et al. (2020) also increase with buyer income.

The data indicate stark differences in markups paid across income groups. Over the observed set of purchases, the average markup paid by households increases from 29 percent for households with less than $20,000 in annual income to 44 percent for households with over $200,000 in annual income. This 15pp markup gap is due to differences in markups paid for identical products (as previously documented by Aguiar and Hurst 2007a; Broda et al. 2009) as well as differences in basket composition: high-income households tend to buy products with higher average markups. To relax the assumption that local inputs do not affect marginal cost, I also consider more conservative measures of the markup gap that control for the county and store of purchase. In both cases, a significant gap in markups paid by high- and low-income households remains: three-quarters of the overall 15pp markup gap persists within county, and half of the markup gap persists within store.

While these estimates describe how markups vary with household income in the cross-section, how markups vary with aggregate income also depends on spillovers that are absorbed in the intercept of cross-sectional regressions (Chodorow-Reich 2020; Wolf 2021). To isolate these spillovers, I exploit variation in the incomes of other customers that a household shops alongside. For example, I consider how markups paid by a household vary with aggregate income in their city, the income of other customers at the retail chains where they shop, and the income of other buyers of the products they purchase. My preferred specifications exploit variation over time in the income of a product's or retail chain's customers, controlling for both time-varying unobserved household characteristics and local costs.

Across an array of specifications that exploit different sources of variation, I find evidence of positive and large spillovers of others' incomes on markups paid. That is, households pay higher markups when shopping alongside higher-income customers. This positive dependence of markups paid on others' incomes implies that the "macro elasticity" of markups to income—that is, the relationship between markups and aggregate income—is larger than the "micro elasticity" observed in the cross-section. Compared to a micro elasticity of markups paid to household income of 3–4 percent, my estimates of the macro elasticity of markups to income range from 8 to 15 percent.

I use these moments to discipline a macroeconomic model in which markups set by firms depend on the income distribution. Households in the model differ from one another in two ways. First, households have different tastes for products, resulting in differences in basket composition across income groups. Second, households have different opportunity costs of time, leading to differences in how much search effort they exert to find low prices. Together, these ingredients generate price dispersion for identical products as well as different average markups across products. In equilibrium, all households' search decisions and the distribution of markups charged by firms are determined jointly and depend on the income distribution.

The model accounts for the patterns recorded in the data—high-income households pay higher markups, and spillovers across households generate a macro elasticity of markups to income greater than the micro elasticity—and produces additional predictions for search behavior that I test in the data. In particular, I show that search intensity in the data is decreasing in own income (as previously documented by Pytka 2018) but increasing in average county income, consistent with the prediction that household search decisions are strategic substitutes in equilibrium. An alternate class of models in which differences in markups across income groups arise due to non-homothetic preferences (e.g., Berry et al. 1995; Simonovska 2015; Auer et al. 2022) are unable to account for these patterns in search behavior or for the contribution of differences in prices paid for identical products to the markup gap across income groups.

I use the model to return to the question that opens this paper: how do changes in the income distribution affect aggregate markups? I derive conditions in the model under which a rise in real income levels or a rise in income dispersion lead to higher markups and show that both sets of conditions hold when I calibrate the model to match differences in markups paid across income groups. The predictions that markups rise with per-capita income and inequality are borne out in data on retail markups across U.S. cities. Remarkably, the model's predictions for aggregate markups across cities account for over 30 percent of the variation in markups across cities in the data, three times more than a representative agent, nested CES model that predicts markups using data on retailer market shares and concentration.

The model suggests that changes in the income distribution can play a meaningful role in the evolution of markups over time. I consider changes in the distribution of post-tax real income in the U.S. documented by Saez and Zucman (2019). Through the lens of the model, the increase in income levels and dispersion from 1950 to 2018 account for a 12pp rise in the aggregate retail markup.[3] This increase is in line with data on retail gross margins from the Census of Annual Retail Trade Survey. Increases in the aggregate markup are moderate from 1950–1980 but accelerate after 1980 due to rising income dispersion.

In the model, half of the rise in the aggregate markup over this time period is due to a reallocation of sales to high-markup firms. The expansion of high-markup firms occurs because declining search intensity and changes in demand composition lead households

[3] Rising markups do not necessarily imply rising profits. In the model, profits are static due to free entry.

to shop more often at firms with high markups. The contribution of reallocations to the increase in the aggregate markup is consistent with evidence from De Loecker et al. (2020) and Autor et al. (2020). However, while some studies interpret the increased market share of high-markup firms as the cause of rising markups, in this paper both reallocations and rising markups are consequences of changes in the demand side of the economy. In all, the calibration suggests that changes to the income distribution can be a potent force in both reshaping market structure and increasing the aggregate markup.

**Related literature.** This paper's empirical findings on differences in markups paid across income groups are closely related to a literature that documents differences in prices paid for identical products, starting with seminal papers by Aguiar and Hurst (2007a) and Broda et al. (2009). When costs are constant across retailers and over time, differences in prices paid for identical products isolate within-product differences in markups paid. My use of wholesale cost data to construct retail markups facilitates comparisons across products (as well as across periods when product costs change). Taking basket composition into account doubles the elasticity of markups paid to household income compared to the within-product component alone, because the baskets of high-income households tend to be composed of higher-markup products.[4]

This paper also adds to a literature in macroeconomics and trade that provides evidence of a relationship between income or wealth, price sensitivity, and markups, including Lach (2007), Alessandria and Kaboski (2011), Simonovska (2015), Anderson et al. (2018), Stroebel and Vavra (2019), DellaVigna and Gentzkow (2019), Jaimovich et al. (2019), Gupta (2020), Handbury (2021), Faber and Fally (2022), and Auer et al. (2022).[5] Most closely related to my study of the relationship between markups and aggregate income are Simonovska (2015), who documents how markups charged by a single e-commerce retailer vary with per-capita income across export destinations, and Anderson et al. (2018), who explore how markups within two retailers vary with local income across store locations. The dataset I construct allows me to capture patterns of substitution across retailers, as well as decompose the relationship between markups and aggregate income into partial and general equilibrium effects. Nevertheless, my estimates of the macro elasticity of markups to income (8–15 percent) are broadly consistent with results from Simonovska

[4]This paper is the first to measure the contribution of basket composition to differences in markups across income groups. Previous work (e.g., Broda et al. 2009; Griffith et al. 2009; Handbury 2021) shows that households can save by substituting to lower-quality products or cheaper stores, but do not have common units in which to compare products of different qualities or stores with different amenities. The use of cost data to construct markups provides a common unit for these cross-good comparisons.

[5]The relationship between income and price sensitivity was first conjectured by Harrod (1936), who called this relationship "The Law of Diminishing Elasticity of Demand."

(2015) and Anderson et al. (2018).

A vast literature in industrial organization estimates markups in consumer markets and often finds a negative relationship between income and price sensitivity. For example, Nevo (2001), Villas-Boas (2007), Nakamura and Zerom (2010), and Grieco et al. (2021) find that high-income consumers are less price sensitive in each of the markets they study (breakfast cereals, yogurt, coffee, and automobiles). The demand systems used in these studies attribute differences in price sensitivity to utility primitives and hence are unable to account for differences in prices paid for identical products in the same market, which I find constitute half of the markup gap across income groups. It is also difficult to infer a macroeconomic relationship between markups and income in these models because they are partial equilibrium and typically estimated within narrow product categories.

My model builds on insights from Aguiar and Hurst (2007a) and Kaplan and Menzio (2015) that search and cost of time play an important role in prices paid, and on the consumer search technology developed by Burdett and Judd (1983).[6] Pytka (2018) and Nord (2022) also develop models in which households have heterogeneous search intensities; both embed these households in an incomplete markets model to explore differences between expenditure and consumption inequality.[7] My model allows for any non-parametric distribution of household incomes, enabling analytic comparative statics with respect to the income distribution, and accommodates a more flexible search technology. The empirical analysis of retail markups, the use of markup data to calibrate the model, and the application to markups over time are also unique to this paper.

Finally, this paper relates to recent work that documents trends in markups over time and considers potential drivers. Most recently, Döpper et al. (2021) and Brand (2021) explore how markups recovered from demand estimation in retail scanner data evolve over time. Both studies provide complementary evidence that demand-side forces play an important role in the evolution of markups. Döpper et al. (2021) attribute rising markups in scanner data to declining consumer price sensitivity and incomplete pass-through of marginal cost reductions. Brand (2021) attributes rising markups to consumers becoming less price sensitive, perhaps due to increased product differentiation. Neither of these papers ties the rise in markups to changes in the income distribution, as I do in this paper.

[6]Theoretical foundations for the importance of search in explaining price dispersion go back to Stigler (1961). See also subsequent empirical work on search effort and the use of savings technologies including Griffith et al. (2009), Aguiar et al. (2013), Coibion et al. (2015), Pytka (2018), and Nevo and Wong (2019).

[7]Nord (2022), which was released after the first version of this paper, also shows that the skewness of price distributions covaries with buyer income, consistent with a search model. Nord (2022) focuses on the feedback between prices and the business cycle, while I focus on long-run trends in markups.

# 2 Constructing a Dataset on Retail Markups

This section describes two data sources—NielsenIQ Homescan and PromoData Price-Trak—that I use to construct a dataset of retail markups paid by households. Appendix A details how the data are cleaned and describes other ancillary data sources.

## 2.1 Data sources

**Consumer panel data.** NielsenIQ Homescan provides transaction data for a nationally representative group of households over the period 2004–2019. In the main text, I present results using data from 2007, covering 62 million transactions by over 60,000 households.[8]

NielsenIQ Homescan panelists use in-home scanners or a mobile application to record all purchases intended for in-home use. The data include purchases of items in all NielsenIQ-tracked categories from any retail outlet. In addition to reporting the date and store visited for each shopping trip, panelists scan the universal product code (UPC) of each item purchased, report the number of units purchased, and record savings from coupons. While NielsenIQ does not pay panelists, it offers households a variety of incentives to accurately report data. Broda and Parker (2014) provide a summary of other key features of the data.

NielsenIQ also collects demographic data on panelists including household size, the age of each household member, race, and household income. Following Handbury (2021), I exclude households with below $10K in income from my analysis.

For some analyses, I make use of NielsenIQ's product hierarchy, which organizes UPCs into product groups and modules. There are about 125 product groups and just over 1,000 highly disaggregated product modules.

**Wholesale costs.** I use data on wholesale costs from PromoData Price-Trak, a weekly monitoring service that tracks wholesale prices for over 100,000 UPCs. The PromoData come from 12 grocery wholesaler organizations that sell products to retailers across the U.S. and covers the period 2006–2012.[9] On a weekly basis, wholesalers send PromoData list prices and promotional discounts that they make available to their customers. Previous

[8]From 2006–2009, NielsenIQ separately labels households with $100K, $125K, $150K, and >$200K in income. This breakdown is not available for other years. I use 2007 to avoid the 2008–2009 recession years.

[9]A significant portion of grocery retail sales pass through wholesalers. In 2007, grocery store retail sales totalled $491B (Census of Retail Trade). Total sales by merchant wholesalers in grocery and related products in that year were $476B (Census Monthly Wholesale Trade Survey), and about half of grocery wholesalers' sales (excluding those to other wholesalers) are to retailers (2012 Economic Census).

studies using these data include Nakamura and Zerom (2010), Stroebel and Vavra (2019), and Afrouzi et al. (2021).

The PromoData report both base prices and "deal prices." Deal prices include promotional discounts and are only available to retailers during windows scheduled by the wholesaler. In the main text, I present results using deal prices as the measure of retailers' wholesale costs (results from instead using base prices are reported in Table 1).

As previously shown by Stroebel and Vavra (2019), wholesale prices from the PromoData are similar across markets: 80 percent of market-item pairs in the data have a wholesale price exactly equal to the modal price across markets in that month (see Appendix Table A1).[10] Hence, for my baseline results, I calculate a national wholesale price for each UPC in each month. Similar results obtain using the subset of transactions where a wholesale price is reported for the market where the purchase is made (see Table 1).

In all, about 67,000 UPCs purchased by Homescan panelists in 2007 are matched to wholesale costs from PromoData. These UPCs constitute 43 percent of transactions and 37 percent of expenditures in the 2007 panel.

## 2.2 Constructing retail markup estimates

I calculate the retail markup on product $g$ purchased by household $i$ in transaction $t$ as the price paid by $i$ over the wholesale cost of product $g$ in the month of transaction $t$,

$$\text{Retail Markup}_{i,g,t} = \frac{\text{Price}_{i,g,t}}{\text{Wholesale cost}_{g,t}}.$$

Here, wholesale costs in the month of purchase proxy for the replacement costs that retailers face for restocking items purchased by customers. As argued by Gopinath et al. (2011), these replacement costs are a reasonable measure of marginal costs since other components of retailers' costs, such as rent, capital, and labor, are fixed at short horizons. The approach of using data on retailers' replacement costs to measure retail markups follows Aguirregabiria (1999), Eichenbaum et al. (2011), Gopinath et al. (2011), and Anderson et al. (2018), among others.[11]

[10] Using cost data from a major grocer, DellaVigna and Gentzkow (2019) also find that wholesale costs across stores do not vary with local income (see DellaVigna and Gentzkow 2019 Appendix Figure 14).

[11] Eichenbaum et al. (2011), Gopinath et al. (2011), and Anderson et al. (2018) each use data on replacement costs provided by a single retailer. Their data include wholesale costs as well as shipping costs and net rebates. While my data allow me to study patterns of substitution across retailers, this comes at the cost of fidelity on shipping costs and rebates. I conduct robustness exercises below that control for unobserved store-level costs and exclude perishable items. Reassuringly, Gopinath et al. (2011) report that their results are unchanged if they use wholesale costs as a proxy for replacement costs, and Anderson et al. (2018) report that gross margins calculated using replacement costs versus wholesale costs are closely correlated.

While my baseline assumption is that these wholesale costs accurately measure retailers' marginal costs, the analysis below seeks to control for two ways in which list wholesale prices in the PromoData may mismeasure the marginal costs faced by retailers: (1) true replacement costs may vary from the list prices recorded by PromoData due to retailer-wholesaler or retailer-manufacturer deals, such as negotiated rebates or volume discounts, as well as shipping/transport costs, and (2) true marginal costs may differ from replacement costs, for instance due to local inputs for shelving and inventory management. In Section 3, I discuss empirical strategies that aim to control for each of these potential sources of mismeasurement.

Retail markups are winsorized at the 1 percent level for all analyses. The cost-weighted average markup in the merged dataset is 32 percent, broadly in line with evidence on retail grocery gross margins from the 2007 Census Annual Retail Trade Survey.

**Selection.** Appendix Table A3 shows that the share of transactions and expenditures that are matched to wholesale costs is similar across income groups. Appendix Table A3 also shows that relative prices of unmatched products exhibit a larger covariance with income than those in the matched sample, suggesting that differences in markups across income groups in the matched sample are likely to be conservative.

**Comparison to other markup measures.** Two other approaches used to measure markups in the literature are demand estimation and production function estimation. In Appendix E, I estimate a random coefficients model à la Berry et al. (1995) to recover marginal costs and markups in a single product category (margarine). The recovered markups exhibit a strong positive correlation ($\rho \approx 0.6$) with the retail markups in my data.[12] Since the NielsenIQ data do not include retailer names, it is more challenging to compare retail markups in my data with firm-level markups measured using the production function approach. Nevertheless, Appendix E shows that markups of public firms measured by De Loecker et al. (2020) also exhibit a positive relationship with customer income.

# 3 Empirical Evidence

In this section, I explore the relationship between markups and household income. Section 3.1 explores how markups vary with household income in the cross-section. Section 3.2 documents that, conditional on income, markups paid by a household also in-

[12]Whether markups recovered from demand estimation include retailers' markups depends on assumptions about vertical conduct between retailers and manufacturers, which I discuss in Appendix E.

crease with the income of other households. These spillovers contribute to the relationship between markups and aggregate income in an economy.

### 3.1 High-income households pay higher markups

Figure 1 plots the aggregate (cost-weighted average) markup paid by households over the income distribution for the sample of purchases matched to wholesale costs. The aggregate markup increases from 29 percent for the lowest-income households in the sample to 44 percent for households with over $200K in annual income.

To address the concerns discussed above about unobserved components of marginal cost, I document the difference in markups paid between low- and high-income households within county and within store. These controls absorb factors that may lead to systematic differences in marginal cost across counties or across stores, such as differences in shipping or local input costs, thus isolating the differences in markups.

The first specification adds demographic controls and county fixed effects:

$$\text{Markup}_{i,g,t} = \sum_{\ell} \beta_\ell 1\{i \text{ has income level } \ell\} + \gamma' X_i + \delta_{\text{County}} + \epsilon_{i,g,t}. \tag{1}$$

The markup paid by household $i$ for good $g$ in transaction $t$ is $\text{Markup}_{i,g,t}$. Demographic controls $X_i$ include fixed effects for race, ethnicity, household size, presence of a female head of household, and the age group of the female head of household; $\delta_{\text{County}}$ are county fixed effects; and $\epsilon_{i,g,t}$ is a mean-zero error. I weight the regression by costs and leave out the income level indicator for households with less than $20K in income, so that the coefficients $\beta_\ell$ are differences relative to the group with below $20K in income.

Figure 2a plots the coefficients $\beta_\ell$ for specification (1) with and without county fixed effects. Controlling for demographics, the fixed effect for the highest-income group is 18pp, slightly larger than the unconditional difference of 15pp. After adding county fixed effects, the fixed effect for the highest-income group falls to 14pp. Hence, about three-quarters of the difference in retail markups paid between the highest and lowest income groups in the sample is due to differences in markups paid within county.

For just over half of the transactions in the sample, NielsenIQ provides store IDs that identify the specific store outlet where each purchase was made. Specification (2) adds store fixed effects:

$$\text{Markup}_{i,g,t} = \sum_{\ell} \tilde{\beta}_\ell 1\{i \text{ has income level } \ell\} + \tilde{\gamma}' X_i + \alpha_{\text{Store}} + \epsilon_{i,g,t}. \tag{2}$$

**Figure 1:** Cost-weighted average markup paid by income group.

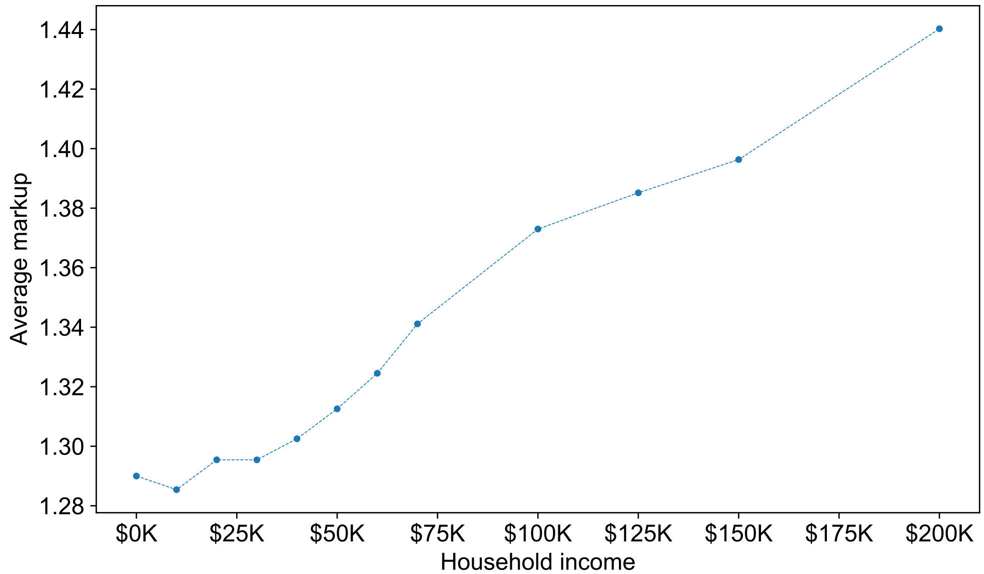



Figure 2b plots the coefficients on income level for specification (2) with and without store fixed effects for the subsample of 14.0 million transactions where a unique store ID is available. For the subsample of transactions made at an identified store outlet, the gap in markups paid across the lowest- and highest-income groups is 13pp, which is moderately smaller than in the full sample.[13] Adding store fixed effects brings the difference in markups paid by the low and high-income households to 7pp. That is, for the sample of transactions made at a NielsenIQ-identified store, over half of the markup gap is due to differences in markups paid within store.

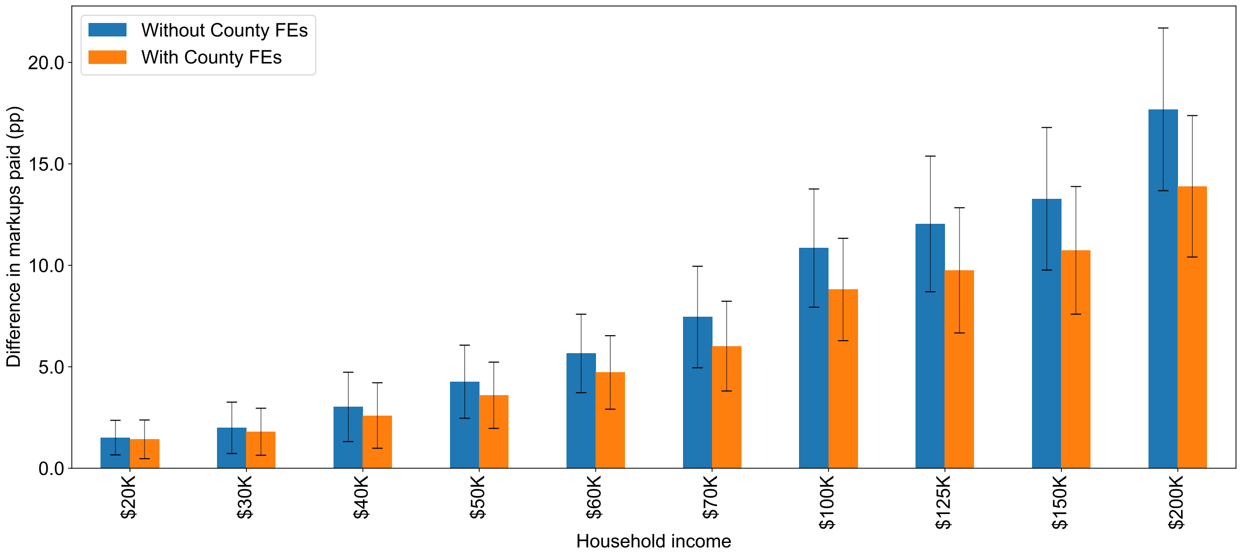
Why do households pay different markups even within the same store? Appendix Figure B2 shows that half of the within-store markup gap is due to differences in markups paid for the same UPC (due to differences in exploiting sales and coupons), and half is due to within-store differences in basket composition. I discuss these two sources of the markup gap in detail in Section 3.1.2.

### 3.1.1 Robustness

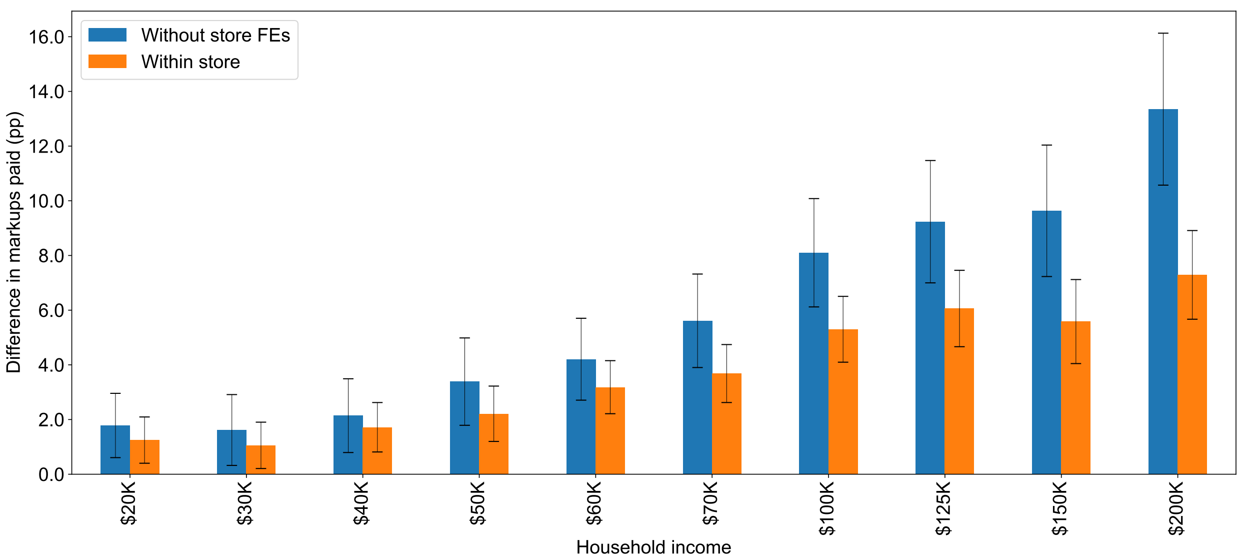
The markup gap across income groups is robust to using alternate measures of retailers' costs and to concerns about unobserved volume discounts for large retailers.

[13]Store IDs are only available for a subset of retailers. The markup gap may be smaller in the subset of transactions with Store IDs because these retailers are more homogeneous than retailers in the full sample.

**Figure 2:** Difference in markups paid relative to households with below $20K income.




**(a)** With and without county fixed effects ($N$ = 25.8 million).




**(b)** With and without store fixed effects (sample with store IDs, $N$ = 14.0 million).

*Note:* These figures plot the coefficients $\beta_\ell$ on household income dummies in a cost-weighted regression of markup paid on household income dummies and demographic controls (race, ethnicity, household size, and presence and age of female head of household). Income levels on the horizontal axis are the minimum of the income bracket provided by NielsenIQ. Standard errors are two-way clustered by product brand and household county. Figure (a) shows $\beta_\ell$ with and without county fixed effects (specification (1)), and (b) shows $\tilde{\beta}_\ell$ with and without store fixed effects for the sample of transactions where the unique store ID is observed (specification (2)).

**Table 1:** Robustness of markup gap.

| *Markup gap (pp) relative to <$20K* | Demographics $100K | Demographics $200K | Within County $100K | Within County $200K | Within Store $100K | Within Store $200K |
|---|---|---|---|---|---|---|
| Baseline | 10.8 | 17.7 | 8.8 | 13.9 | 5.3 | 7.3 |
| Weighting by sales | 10.4 | 17.8 | 7.8 | 12.7 | 4.5 | 6.1 |
| Using PromoData base price | 9.7 | 16.0 | 8.0 | 12.7 | 5.6 | 7.9 |
| Using PromoData market-level price | 8.5 | 16.3 | 6.6 | 11.7 | 5.8 | 8.9 |
| With day-of-week fixed effects | 10.6 | 17.5 | 8.6 | 13.8 | 5.1 | 7.1 |
| With supply-side controls | 10.6 | 17.4 | 8.5 | 13.6 | 5.2 | 7.2 |
| Excluding perishable items | 10.5 | 17.1 | 8.4 | 13.4 | 4.8 | 6.7 |

*Note:* Results from fixed effect specifications (1) and (2) under alternate measures. The baseline measure is the cost-weighted gap in markups constructed using deal wholesale prices from PromoData.

**Spoilage and congestion costs, and other measures of wholesale costs.** Table 1 shows that the unconditional, within-county, and within-store markup gaps are robust to a number of different measurement choices. For example, the markup gap is similar if regressions are weighted by sales rather than costs; if the PromoData base price, which excludes promotional discounts, is used as the measure of wholesale costs; and if the markup gap is measured using only the subset of transactions for which PromoData reports a wholesale price of the purchased item in the market of purchase. Including day-of-week fixed effects, which control for changes in congestion faced by retailers over the course of the week, or supply-side controls, which include sales shares of the purchased UPC/brand and sales concentration in the product module, also do not alter the results. Finally, limiting the sample to non-perishable items, and thus excluding food items that may have higher shipping and spoilage costs, does not materially affect the results.

**Differential volume discounts.** While the store fixed effects in specification (2) absorb systematic differences in wholesale, shipping, and local input costs that cause marginal costs to differ by store, they do not absorb heterogeneity in marginal costs by product-store pair. This is a problem if some retailers face lower marginal costs on a subset of products, which would be the case for example if large retailers negotiate volume discounts on commodity items but not luxury items.[14] In this case, the data would overstate the marginal cost and understate the markup on commodity items sold at large retailers. If low-income households buy more commodity items at large retailers

[14]This concern is most salient with respect to retailer size, since the Robinson–Patman Act limits wholesalers and manufacturers from selling an identical product to different retailers at different prices, but does permit differences in prices that reflect differences in cost of delivery or volume.

than high-income households, this mismeasurement would lead me to overestimate the difference in markups paid across income groups.

I address this concern by testing whether the markup gap is driven by large retailers in the sample. I rank retailers by total sales and estimate specification (2) for the subsample of transactions excluding the largest retailer, the largest three retailers, the largest five retailers, and so on. If the markup gap is partially driven by mismeasurement of marginal costs at large retailers, then the gap should attenuate as large retailers are removed.

Appendix Figure B1 shows that the estimated markup gap is stable as large retailers are removed from the sample. This suggests that the estimated difference in markups paid is not a result of differential quantity discounts at large retailers.

#### 3.1.2 Accounting for the markup gap

Two forces contribute to the gap in markups paid across income groups. First, high-income households pay higher prices for identical products, as documented in seminal papers by Aguiar and Hurst (2007a) and Broda et al. (2009). This force accounts for about half of the markup gap conditional on demographics, within county, or within store.[15] In principle, differences in prices paid for identical products could arise from differences in exploiting variation in posted prices or from differences in coupon usage. In the data, coupons play a negligible role (Appendix Figure B3 shows that savings due to coupon usage vary by less than 1pp across income groups), suggesting this component of the markup gap is predominantly due to differences in how income groups exploit spatial and intertemporal variation in posted prices.

The remainder of the markup gap is due to differences in the composition of shopping baskets across income groups: high-income households pay higher markups because their baskets contain a larger share of high-markup products. Prior work shows that households can save by substituting to lower-quality products or less expensive stores (e.g., Broda et al. 2009; Griffith et al. 2009; Handbury 2021), but lack common units in which to compare products of different qualities or stores with different amenities. The use of cost data to construct markups provides a common unit for cross-good comparisons, allowing me to measure this source of markup differences across income groups for the first time.[16]

[15] That is, the gap in markups paid for identical products is 9pp conditional on demographics, 7pp within county, and 4pp within store (Appendix Figure B4), compared to total markup gaps of 18pp, 14pp, and 7pp.

[16] This across-product component of the markup gap could be positive or negative. Classic models of quality discrimination suggest that firms set low markups on goods bought by high-income customers to deter these customers from substituting to lower quality goods (e.g., Mussa and Rosen 1978, Tirole 1988 Ch. 3). Other models (such as the one in this paper) predict that markups on goods purchased by high-income

Including this second source of markup differences across households doubles previous estimates of the elasticity of markups to household income that only exploit differences in prices paid for identical products. Appendix Table B1 shows that the elasticity of markups paid to household income is 3.2–3.8%, compared to estimates of the elasticity of prices paid for identical products to income of 1.1–1.3% in Broda et al. (2009).[17]

Some differences in basket composition across income groups could be due to differences in search behavior (e.g., if some products exhibit systematically higher variation in prices). To decompose the markup gap into a component due to search and a component due to basket composition, I construct a measure of markups that each household would pay if it instead paid the average posted price for a UPC in its county, using data on posted prices from NielsenIQ Retail Scanner data. These counterfactual markups remove any effects of consumers' ability to exploit price dispersion, thus isolating effects of basket composition. These estimates (shown in Appendix Figure B5) attribute 40 percent of the market gap to differences in basket composition and the remaining 60 percent to differences in search. I return to these moments when calibrating the model in Section 6.

## 3.2 Markups depend positively on other buyers' incomes

This section explores how markups paid by a household depend on the income of other buyers. These spillovers are important for characterizing how the *macro elasticity* of markups to income—the response of the aggregate markup to an increase in all households' incomes—relates to the micro elasticity of markups to income observed in the cross-section. To a first order, the macro elasticity is the sum of the elasticity of markups to own income and the elasticity of markups paid to others' incomes.[18]

There are several possibilities. If markups paid depend on a household's relative income rather than absolute income, then these spillovers will be negative and exactly offset the micro elasticity, leading to no aggregate relationship between markups and income. If markups depend only on a household's own income—i.e., retailers perfectly price discriminate across customers—then these spillovers will be zero, leading the macro and micro elasticities to coincide. Finally, imperfect price discrimination could lead to positive spillovers and a macro elasticity larger than the micro elasticity.

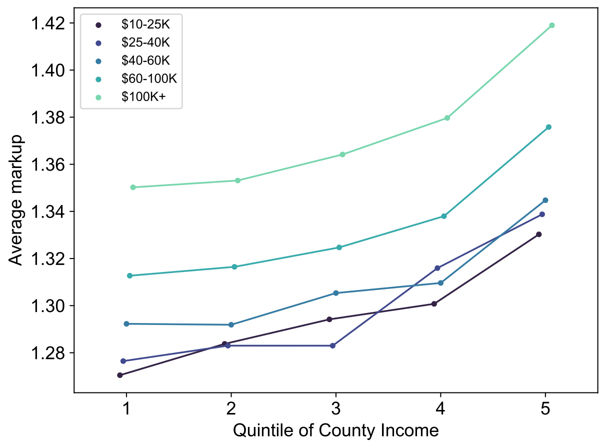
A first look at the data suggests the last case may be the relevant one. Figure 3 shows

households are higher due to differences in price sensitivity across income groups.

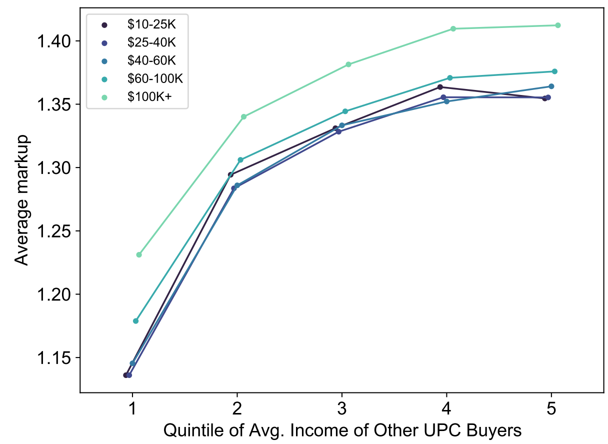
[17]The elasticity of prices paid for identical products to income in my sample is slightly higher at 1.7–2.0%.

[18]Formally, suppose markups depend on own and others' incomes, $\mu_i = \mu(z_i, z_{-i})$. If all incomes rise $d \log z$, the change in the aggregate (cost-weighted average) markup is $d \log \mathbb{E}[\mu_i] \approx (\mathbb{E}[\frac{\partial \log \mu_i}{d \log z_i}] + \mathbb{E}[\frac{\partial \log \mu_i}{d \log z_{-i}}])\, d \log z$, where $\mathbb{E}[\cdot]$ denotes the cost-weighted average.

**Figure 3:** Average retail markup by income group, split by other buyers' incomes.




**(a)** County income




**(b)** Income of UPC's other buyers

the average markup paid by five income groups split by (a) quintile of county income and (b) quintile of the income of a UPC's other buyers. Consistent with the findings of the prior section, markups paid by the highest income group lie above markups paid by the lowest income group in all cases. Conditional on income, markups paid also increase with the income of other buyers in the same county or purchasing the same product.

There are two primary concerns in interpreting this descriptive evidence. First, measured markups may be affected by unobserved components of marginal cost that vary across space. Second, households' choices of where to live or what to buy may be determined by unobserved households characteristics that are correlated with other buyers' incomes (a Manski 1993 reflection problem). To overcome these concerns, I exploit time series variation in the income of buyers that a household shops alongside. For this purpose, I compile data on retail markups for all years from 2006 to 2012. I consider three different definitions of the pool of shoppers that a household is grouped with, each of which exploits a different source of variation and allows different controls.

The first specification considers how markups paid by household $i$ for UPC $g$ at store $s$ in year $t$ on transaction $k$ depend on average income in the core-based statistical area (CBSA) where the household lives:[19]

$$\log \text{Markup}_{i,g,s,t,k} = \beta_1 \log \text{CBSA Income}_{\text{CBSA}(i,t),t} + \gamma_{i,\text{IncomeLevel}(i,t)} + \alpha_s + \delta_t + \varepsilon_{i,g,s,t,k}. \quad (3)$$

This specification isolates variation in markups paid by a household due to variation in

[19]CBSAs refer collectively to metropolitan and micropolitan statistical areas. Annual estimates of average CBSA income come from the Bureau of Economic Analysis (BEA) (see Appendix A).

CBSA income over time. The fixed effects $\gamma_{i,\text{IncomeLevel}(i,t)}$ absorb household characteristics that may be correlated with a household's choice of location and within-household changes in income, the store fixed effects $\alpha_s$ absorb potentially unobserved components of marginal cost that vary at the store level but are fixed over time, and the year fixed effects $\delta_t$ absorb secular time trends in unobserved components of marginal cost.

Table 2 column 1 reports the results from specification (3): doubling the per-capita income in the CBSA where a household lives is associated with a 7.1 percent increase in markups paid by the household. This estimate is indicative of large, positive spillovers of others' incomes on markups paid. Note that these estimated spillovers are inclusive of endogenous responses by households and firms, such as entry of new stores, changes in how households allocate spending across stores and goods, and so on.

The second specification considers how markups paid depend on the income of other buyers at the same retail chain:

$$\log \text{Markup}_{i,g,s,t,k} = \beta_2 \log \text{Income at Retailer Locations}_{\text{Retailer}(s),t} + \gamma_{i,t} + \alpha_s + \phi_{\text{County}(s),t} + \varepsilon_{i,g,s,t,k}. \quad (4)$$

To calculate average income at a retailer's locations, for each store $s$, I take the sales-weighted average of per-capita county income (from the BEA) across all stores in the same retail chain and same NielsenIQ designated market area (DMA) as store $s$.

The intuition for the variation exploited in specification (4) is best summarized by the following experiment. Suppose household $i$ shops at two retail chains A and B. Retailer A has locations in another city $C_A$ and retailer B has locations in another city $C_B$. If incomes in city $C_A$ rise relative to $C_B$, how do the markups paid by household $i$ at retailer A change compared to at retailer B? Since retailers tend to set uniform prices across cities (DellaVigna and Gentzkow 2019), markups paid may change differentially since the income of buyers household $i$ is pooled with when shopping at retailer A is changing relative to the income of buyers household $i$ is pooled with at retailer B.

Note also the addition in (4) of county-year fixed effects $\phi_{\text{County}(s),t}$, which absorb variation in county-level costs over time, and of household-year fixed effects $\gamma_{i,t}$, which absorb changes in unobserved household characteristics over time. These additional controls overcome concerns from (3) that spillovers may be biased by unobserved changes in marginal costs that are correlated with changes in local income, or by changes in unobserved household characteristics. The fact that two stores in the same county in the same year can be exposed to customer bases with different incomes—due to changes in income at other stores in the same retail chains—allows for these additional controls.

**Table 2:** Spillovers of other buyers' incomes on markups paid, exploiting variation over time from 2006–2012.

| *Log Retail Markup* | (1) | (2) | (3) |
|---|---|---|---|
| Log CBSA Income | 0.071** | | |
| | (0.013) | | |
| Log Income at Retailer's Locations | | 0.068** | |
| | | (0.030) | |
| Log Income of Other UPC Buyers | | | 0.142** |
| | | | (0.038) |
| Fixed effects | HH-Inc. Level, Store, Year | Household-Year, Store, County-Year | Household-Year, Store-Year |
| $N$ (millions) | 91.9 | 50.8 | 97.0 |
| $R^2$ | 0.19 | 0.21 | 0.21 |

*Note:* The sample includes all household transactions matched with wholesale cost data from 2006 to 2012. Regression weighted by sales (in 2007 USD), and standard errors two-way clustered by product brand and household county. * indicates significance at 10%, ** at 5%.

Column 2 of Table 2 reports that the elasticity of markups paid to average income across a retailer's locations, conditional on these controls, is 6.8 percent.

Finally, the third specification exploits variation in the income of buyers of the same UPC over time,

$$\log \text{Markup}_{i,g,s,t,k} = \beta_3 \log \text{Income of other UPC buyers}_{g,t} + \gamma_{i,t} + \psi_{s,t} + \varepsilon_{i,g,s,t,k}, \tag{5}$$

where the store-year fixed effects $\psi_{s,t}$ control for potentially time-varying unobserved components of cost at the store level. This specification measures how markups paid by household $i$ for two products at the same store change as the average income of those products' buyers change over time. I calculate the income of the other buyers of UPC $g$ in year $t$ as the sales-weighted average income of all households purchasing $g$ in year $t$ excluding household $i$. Table 2 column 3 reports an elasticity of markups paid to the income of a UPC's other buyers of 14.2 percent.

Thus, across all three specifications, Table 2 finds positive and large spillovers of others' incomes on markups paid, ranging from 6 to 14 percent. These spillovers are economically meaningful, since the elasticity of markups paid to own income is 3.2–3.8 percent. Combining the micro elasticity of markups to own income with the elasticity to others' incomes, my preferred estimates for the macro elasticity of markups to income fall between 8 and 15 percent.[20]

While this paper is the first to estimate and decompose the aggregate relationship between markups and income into own-income effects and spillovers, the 8–15 percent range aligns with previous work. Using data from an online apparel retailer, Simonovska (2015) finds an elasticity of markups on identical products to per-capita income across export destinations of 12–24 percent. Anderson et al. (2018) use data from two retailers and estimate an elasticity of store gross margins to local income of 10 and 17 percent.

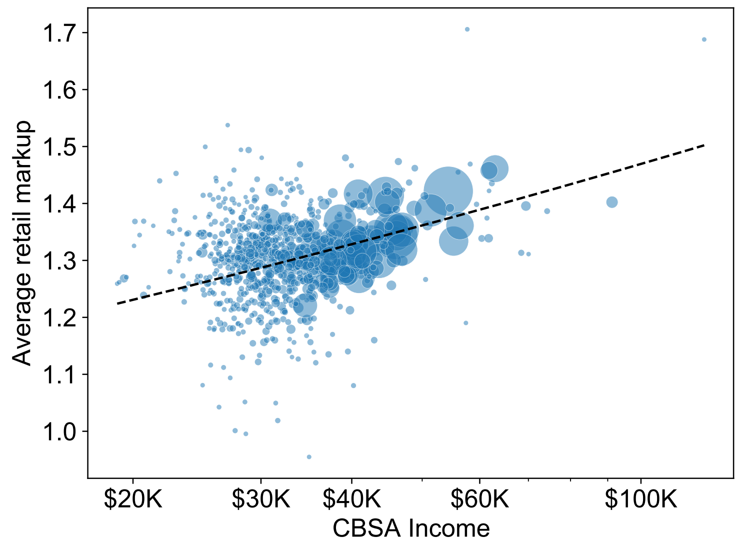
**Heterogeneity across income groups.** Appendix Table B7 explores whether the direction and magnitude of spillovers are heterogeneous across income groups. In the Varian (1980) model of sales, for example, increasing the share of uninformed buyers (i.e., price-insensitive, high-income households) increases prices paid by uninformed buyers but decreases prices paid by informed buyers (i.e., price-sensitive, low-income households), since informed buyers benefit from the entry of more stores. I find that spillovers are positive and of similar magnitude for households in all income groups, with marginally larger spillovers for high-income households. I return to this evidence when discussing pro-competitive effects in the model in Section 4.8.

**Comparison to cross-CBSA relationship.** While specifications (3)–(5) carefully control for unobserved local costs and household characteristics, we can compare the macro elasticity of markups to income estimated from this bottoms-up approach to a "naive," top-down approach that estimates the relationship between markups and aggregate income in the cross-section of cities. Figure 4 plots retail markups in each CBSA against per-capita income. The elasticity of CBSA markups to income is 11 percent, squarely in the range of macro elasticities estimated using the bottoms-up approach.

Why doesn't the elasticity of markups to aggregate income, as measured in the cross-section of CBSAs, seem to be biased by unobserved local costs? In Appendix D, I estimate the bias due to unobserved local costs using data from the Census Retail Trade Survey of Detailed Operating Expenses and data on retail wages and rents. Even if a substantial portion of labor and rent expenses are recategorized as variable costs (instead of overhead costs), the bias in the measured elasticity across CBSAs is mild, because retailers' total labor and rent costs are small relative to costs of goods sold.

[20] Analogous specifications to (3)–(5) can also be estimated in a single cross-section of data, though using a cross-section of data prevents me from including store fixed effects to absorb local variation in costs. Estimates from the cross-section are qualitatively similar (see Appendix Table B2).

**Figure 4:** CBSA per-capita income and average retail markup.




*Note:* Each bubble is a core-based statistical area (CBSA). The size of each bubble is proportional to total CBSA expenditures in the NielsenIQ Homescan data. CBSA per-capita income in 2007 is from the BEA.

**Macro elasticities across space vs. over time.** The final analysis in this section explores whether the link between markups and aggregate income differs across space *vs.* over time. Appendix Table B4 estimates the elasticity of markups to CBSA income using data from 2006–2012, first exploiting variation in income across CBSAs, and second exploiting only variation in income over time within each CBSA. The two estimates are nearly identical, suggesting that the link between markups and income over time within a CBSA is similar to the relationship between markups and income across CBSAs. Since this exercise is limited to a short sample period, I conduct similar analyses using data on unit prices in each product category as a proxy for markups from 2004–2019. Whether looking across all years from 2004 to 2019 or only looking at changes over from the beginning to end of the sample (and thus removing the influence of business cycles), the elasticities across space and over time are nearly identical.

# 4 A Search Model of Income and Markups

In this section, I develop a quantitative model of consumer search and firm pricing that can account for the empirical patterns documented in the previous section. Besides generating additional testable predictions, the model allows us to consider more ambitious counterfactuals—such as how changes in income dispersion affect markups—that are difficult to isolate in the data.

Households in the model differ in two key ways. First, households with different income levels have heterogeneous tastes, resulting in differences in basket composition across income groups. Second, households endogenously choose different search intensities. The latter draws on a large literature that documents the importance of search effort in prices paid (e.g., Aguiar and Hurst 2007a, Alessandria and Kaboski 2011, Kaplan and Menzio 2015). Together, these forces generate both patterns documented in the previous section: (1) markups increase with household income, owing to both differences in prices paid for identical products and basket composition, and (2) the macro elasticity of markups to income is greater than the micro elasticity.

## 4.1 Households

A unit measure of households purchases goods indexed by $k = 1, ..., K$. Household $i$'s utility is given by the CES preferences,

$$u(\{c_{ik}\}) = \left( \sum_{k=1}^{K} (\beta_{ik} c_{ik})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}},$$

where $c_{ik}$ is household $i$'s consumption of good $k$, and $\beta_{ik}$ is a taste shifter for good $k$ that is allowed to vary across households. In the data, these taste shifters will allow me to flexibly match differences in basket composition across income groups.[21]

When purchasing each good $k$, households know the distribution of prices set by firms, $F_k(p)$, but do not know which firm sells at which price. As a result, households retrieve price quotes from firms before making a purchase. The number of price quotes observed by household $i$ buying good $k$ prior to purchase is a random variable with probability mass function $\{q_{ik,n}\}_{n=1}^{\infty}$. That is, with probability $q_{ik,1}$ household $i$ observes only a single price quote, with probability $q_{ik,2}$ the household observes two price quotes, and so on. Each price quote is an independent draw from the distribution of prices $F_k$.

Upon receiving $n$ price quotes, households compare the minimum price quote received to an exogenous reservation price, $R$, and buy one unit of the good from the firm with the lowest price quote $p$ as long as $p \leq R$. If all $n$ quotes received by the household have a price greater than $R$, the household costlessly re-draws $n$ quotes. The household repeats this search process for each unit of consumption. I assume units of consumption are infinitesimal, so that integer constraints can be ignored.

The distribution $\{q_{ik,n}\}_{n=1}^{\infty}$ is determined by the household's search intensity for good $k$,

[21] As we will see below, these preferences determine basket composition, but do not affect markups set by firms—firms' residual demand curves and hence markups are pinned down by household search behavior.

denoted $s_{ik}$. Formally, the function $\mathcal{S} : s \mapsto \{q_n\}_{n=1}^{\infty}$ maps search intensity to the distribution of number of price quotes received. Assumptions about the mapping function $\mathcal{S}$ are laid out in Section 4.4, but one should have in mind that greater search intensity increases the likelihood of receiving more price quotes, thereby lowering expected prices.

Households choose these search intensities $s_{ik}$, as well as time spent working $l_i$ and consumption $c_{ik}$, to maximize utility subject to a time constraint and budget constraint:

$$\max_{l_i, \{c_{ik}, s_{ik}\}} u(\{c_{ik}\}) \qquad \text{s.t.} \qquad \begin{cases} \sum_k t_i(c_{ik}, s_{ik}) + l_i = 1, & \text{(Time constraint)} \\ \sum_k p_{ik} c_{ik} = z_i l_i, & \text{(Budget constraint)} \end{cases}$$

where $t_i(c, s)$ is the time it takes household $i$ to shop for $c$ units with search intensity $s$, $l_i$ is time spent working with labor productivity $z_i$, and $p_{ik}$ is the average price paid by $i$ for good $k$. With infinitesimal units of consumption, there is no uncertainty in the average price $p_{ik}$ or consumption $c_{ik}$ of each good $k$. The budget constraint anticipates that free entry will set firms' profits to zero, so that all earnings come from labor market work.

I assume that the amount of time it takes household $i$ to shop for $c$ units with search intensity $s$ is

$$t_i(c, s) = \frac{c}{a_i} s, \tag{6}$$

where $a_i$ is household $i$'s search productivity. Search productivity $a_i$ is allowed to vary to reflect the fact that households may differ in their access to search technologies. For example, access to a car or to a greater density of nearby stores decreases the time required to retrieve a given number of price quotes.

Equation (6) assumes time spent shopping for a good increases linearly with the amount of consumption $c$ and with search intensity $s$ and decreases with search productivity $a_i$. The assumption that time spent shopping is linear in search intensity $s$ is without loss of generality, since we could accommodate a different increasing relationship by adjusting the mapping function $\mathcal{S}$. The relationship between shopping time and consumption amount is more important: if shopping time did not increase with basket size, doubling income would double both the time cost and gains from search, leading to no change in search behavior. However, ample evidence in previous work (Aguiar and Hurst 2007a; Pytka 2018) and in my data (Appendix Tables B5–B6) supports the assumption that shopping time increases with basket size.[22]

Households increase search intensity $s_{ik}$ as long as the savings are greater than the time

[22] Appendix Table B5 shows that household shopping time (measured with proxies from Kaplan and Menzio 2015) increases with total expenditures, controlling for income, demographic characteristics, and markups paid. Similar results obtain using household size as an instrument for expenditures. Appendix Table B6 shows that shopping time also increases with basket size within-households over time.

cost of doing so. Thus,

$$-\frac{\partial p_{ik}(s_{ik}, F_k)}{\partial s_{ik}} \leq \phi_i, \tag{7}$$

where the *opportunity cost of search effort* $\phi_i = z_i/a_i$ captures the foregone labor market earnings from increasing search intensity $s_i$. If $\phi_i$ is too high, a household will choose the corner case $s_{ik} = 0$. I focus on the case where each household has an internal solution for $s_{ik}$, so that (7) holds as an equality.

**Aggregate search behavior.** Households are indexed by labor productivity $z \in (0, \infty)$, where labor productivity is distributed in the population according to the cumulative density function $H(z)$. I assume that taste shifters and search productivity are identical for all households with a given labor productivity $z$. Thus, all households with labor productivity $z$ share the same opportunity cost of search effort $\phi(z) = z/a(z)$.

Denote the aggregate consumption of good $k$ by $C_k$ and the consumption-weighted distribution of $z$ by

$$d\Lambda_k(z) = \frac{c_k(z)}{C_k} dH(z). \tag{8}$$

I refer to $\Lambda_k(z)$ as the *distribution of buyers' incomes* for good $k$ because it captures the cumulative distribution of wages $z$ over the set of purchases.

We can summarize aggregate search behavior for good $k$ using the probability mass function $\{\bar{q}_{k,n}\}_{n=1}^{\infty}$, where

$$\bar{q}_{k,n} = \int_0^{\infty} q_{k,n}(z) d\Lambda_k(z), \qquad \text{for all } n. \tag{9}$$

## 4.2 Firms

Each good $k$ is supplied by a measure $M_k$ of ex ante identical firms, which produce output with a constant-returns production technology in labor. I take the per-unit variable cost of production as the numeraire (i.e., households' labor productivities $z$ are measured relative to the cost of producing one unit of output).

Firms set prices to maximize variable profits $\pi(p)$,

$$\max_p \pi(p) = (p-1)D_k(p),$$

where the demand curve $D_k(p)$ that a firm faces depends on its price, the distribution of prices charged by other firms $F_k$, and the aggregate search behavior of households.

Following Burdett and Judd (1983), define a *dispersed-price equilibrium* as an equilibrium in which firm prices follow the distribution $F(p)$, all firms choosing a price $p \in \text{supp}(F)$ make identical profits, and charging any price $p \notin \text{supp}(F)$ results in strictly lower profits. Proving the existence of a dispersed-price equilibrium follows closely from Burdett and Judd (1983); I relegate the details to Appendix C. Given aggregate search behavior $\{\bar{q}_n\}_{n=1}^{\infty}$ with $\bar{q}_1 \in (0,1)$, the unique equilibrium price distribution $F(p)$ is

$$F(p) = \begin{cases} 0 & \text{if } p < \underline{p} \\ 1 - \Psi\left[\left(\frac{R-1}{p-1}\right)\bar{q}_1\right] & \text{if } \underline{p} \leq p \leq R \\ 1 & \text{if } p > R, \end{cases} \qquad \text{where} \qquad \underline{p} = 1 + \frac{\bar{q}_1}{\sum_{n=1}^{\infty} n\bar{q}_n}(R-1), \tag{10}$$

and $\Psi(\cdot)$ is the inverse of the strictly increasing, $C^{\infty}$ function $y(x) = \sum_{n=1}^{\infty} n\bar{q}_n x^{n-1}$.

Firms pay an entry cost of $f_e$ units of labor. Free entry and exit determines the mass of firms supplying each good, $M_k$, so that in equilibrium $\pi(p) - f_e = 0$ for all $p \in [\underline{p}, R]$.[23]

## 4.3 Equilibrium

An equilibrium is a tuple $\left(F_k, \{c_k(z), s_k(z)\}_{z=0}^{\infty}, M_k\right)_{k=1,\dots,K}$ such that (1) consumption and search intensity chosen by households for each product, $c_k(z)$ and $s_k(z)$, maximize utility; (2) $F_k$ for each $k$ is a dispersed price equilibrium given aggregate search behavior $\bar{q}_{k,n}$; (3) firms make zero profits net of the entry cost; and (4) all resource constraints are satisfied.

When search intensities are chosen endogenously, the Burdett and Judd (1983) framework produces two dispersed price equilibria, where one dispersed price equilibrium is stable and the other is unstable (see Burdett and Judd 1983 Theorem 2). I conduct all comparative statics locally around the stable equilibrium in each market.

## 4.4 Characteristics of the search mapping $\mathcal{S}$

It will be useful for the following results to define two additional conditions on the mapping $\mathcal{S}$ from search intensity to the distribution of price quotes received.

**Assumption 1.** Denote the cumulative mass function of $\{q_n\}_{n=1}^{\infty}$ by $Q_n$. The mapping $\mathcal{S} : s \mapsto \{q_n\}_{n=1}^{\infty}$ is such that $Q_n(s)$ is weakly decreasing in $s$ for all $n$ and strictly decreasing in $s$ for $n = 1$; $Q_1(0) = 1$ and $\lim_{s\to\infty} Q_1(s) = 0$; and for any non-degenerate distribution $F$,

$$\sum_{n=1}^{\infty} \frac{d^2 Q_n}{ds^2} \left[\mathbb{E}\left[p|n\right] - \mathbb{E}\left[p|n+1\right]\right] > 0, \tag{11}$$

[23] Assuming free entry or an exogenous mass of firms does not affect equilibrium markups. As we will see in Lemma 3, markups are pinned down by aggregate search behavior alone.

where $\mathbb{E}[p|n]$ is the expected value of the minimum of $n$ draws from $F$.

Assumption 1 imposes that expected price is decreasing and convex in search intensity. Expected price is decreasing in search intensity because increasing $s$ results in a FOSD shift in $\{q_n\}_{n=1}^{\infty}$; expected price is convex in search intensity because the left-hand side of (11) is the second derivative of expected price with respect to search intensity.

**Assumption 2.** For any non-degenerate distribution $F$, the mapping $\mathcal{S}$ satisfies

$$\sum_{n=1}^{\infty}\left(\frac{d^2Q_1}{ds^2}\frac{d^2Q_n}{ds^2}-\frac{dQ_1}{ds}\frac{d^3Q_n}{ds^3}\right)\left[\mathbb{E}\left[p|n\right]-\mathbb{E}\left[p|n+1\right]\right]\geq 0,$$

where $Q_n$ and $\mathbb{E}[p|n]$ are as defined in Assumption 1.

Assumption 2 guarantees that when $\phi(z)$ is increasing and convex in $z$, then the probability of receiving only one price quote, $q_1(z)$, is also increasing and convex in $z$. While Assumption 1 is imposed throughout, Assumption 2 will only be necessary when characterizing how a change in income dispersion affects markups.

While these two conditions may appear technical, I show in Appendix C.5 that both assumptions are satisfied by the two most common parameterizations of the Burdett and Judd (1983) model used in the literature: (1) a version in which households receive only one or two quotes (e.g., Alessandria and Kaboski 2011, Pytka 2018, and Nord 2022), and (2) a version in which the number of price quotes received is drawn from a Poisson distribution (e.g., Albrecht et al. 2021 and Menzio 2021).

## 4.5 Search and markups paid in the cross-section

Lemma 1 shows that how search intensity and markups paid vary over the income distribution depends on how the opportunity cost of search effort, $\phi(z)$, varies with income.

**Lemma 1** (Search intensity and markups paid)**.** *Suppose Assumption 1 holds. If the opportunity cost of search effort $\phi(z)$ is increasing (decreasing) in z, then*

1. *Search intensity $s_k(z)$ for any good k is decreasing (increasing) in z,*
2. *Average prices $p_k(z)$ and markups paid $\mu_k(z)$ for any good k are increasing (decreasing) in z.*

Whether high-income households pay higher or lower markups for identical goods depends on whether search productivity $a(z)$ increases more or less than one-for-one with labor productivity $z$. This means that the model can in principle generate either of two possibilities frequently posited in the literature: (1) if search productivity rises faster than one-for-one with labor productivity, low-income households receive fewer price quotes than high-income households in equilibrium and pay a "poverty premium" (e.g., Caplovitz 1963; Prahalad and Hammond 2002); (2) if search productivity rises less than one-for-one with labor productivity, high-income households exert less search intensity and hence pay higher average markups for identical products.

While the empirical evidence implies that the latter case is relevant for our setting, Lemma 1 explains why the relationship between markups and customer income may vary in other settings. For example, Grunewald et al. (2020) find that low-income customers pay higher markups in the auto loan market, where internet access and education likely play a heightened role in consumers' abilities to gather and compare price quotes.[24]

## 4.6 Strategic interactions in search

A household's choice of search intensity depends on its own cost of search effort and on the potential savings that can realized by increasing search intensity. Since firms' prices—and thus the returns to search—are determined by aggregate search behavior, household search decisions will be shaped by other households' search behavior. Lemma 2 shows that when search costs are sufficiently low, household search decisions are strategic substitutes.

**Lemma 2** (Strategic substitutes in search)**.** *There exists some $\phi^{cutoff}$ such that, if $\phi(z) < \phi^{cutoff}$ for all $z$, then (1) a stable dispersed-price equilibrium exists, and (2) in that equilibrium, household search decisions are strategic substitutes ($\partial s_{ik}/\partial s_{jk} \leq 0$ for any two households $i \neq j$ in market $k$).*

Intuitively, whether household search decisions are strategic complements or substitutes depends on whether an increase in search intensity by one household increases or decreases returns to search for other households. When search costs are low—and hence aggregate search intensity is initially high—an increase in aggregate search intensity decreases returns to search. Thus, an increase in search intensity by one household leads other households to decrease search intensity.[25] Section 5 below shows that search behavior in the data conforms with this prediction. Lemma 2 also provides an explanation for findings from Nevo and Wong (2019) that increases in household search effort during the Great Recession coincided with declines in the returns to shopping effort.

[24] In Pytka (2018) and Nord (2022), high-income households instead search less due to a convex disutility of shopping time. As a result, these models cannot admit a "poverty premium" in some markets.

[25] When search costs are too large, households choose not to search, and the dispersed-price equilibrium ceases to exist. There are intermediate values of $\phi$ for which the dispersed-price equilibrium exists and household search decisions are strategic complements. However, the strategic substitutes behavior in Lemma 2 is the relevant region both in the calibrated model and in data on search behavior.

As we will see, strategic substitutabilities in search behavior moderate the effect of changes in the average cost of search effort on aggregate markups. Of course, this moderating feedback loop is absent in models where household price sensitivity is determined by exogenous utility primitives.

## 4.7 Comparative statics of aggregate markup to $\Lambda(z)$

Define the *aggregate markup* $\bar{\mu}$ as the ratio of total sales to total variable costs in the economy, and $\bar{\mu}_k$ as the ratio of sales to variable costs for good $k$. Lemma 3 shows that the fraction of households receiving one price quote is a sufficient statistic for the aggregate markup.

**Lemma 3.** *In equilibrium, the aggregate markup for good k is* $\bar{\mu}_k = 1 + (R-1)\bar{q}_{1,k}$. *The aggregate markup in the economy is*

$$\bar{\mu} = 1 + (R-1)\frac{\sum_k C_k \bar{q}_{k,1}}{\sum_k C_k}.$$

Intuitively, since firms selling good $k$ must make identical profits at all prices in the support of $F_k$, and the only customers of a firm with price $R$ are those that receive no other price quotes, the fraction of customers receiving only one price quote $\bar{q}_{k,1}$ pins down the profits of all firms and hence the aggregate markup for good $k$. The aggregate markup in the economy is then a cost-weighted average over all goods.

I will now consider how changes to the income distribution affect the aggregate markup. To provide an analytic characterization, I specialize to the case where $K = 1$, so that there is a single distribution of buyers' incomes $\Lambda(z)$.

**Proposition 1** (First-Order Shift)**.** *Suppose* $K = 1$ *and Assumption 1 holds. Consider a perturbation in the household distribution such that the new* $\tilde{\Lambda}(z)$ *first-order stochastically dominates the initial* $\Lambda(z)$. *This perturbation increases the aggregate markup if* $\phi(z)$ *is increasing in z.*

**Proposition 2** (Mean-Preserving Spread)**.** *Suppose* $K = 1$ *and Assumptions 1 and 2 hold. Consider a perturbation in the household distribution such that the new* $\tilde{\Lambda}(z)$ *is a mean-preserving spread of the initial* $\Lambda(z)$. *This perturbation increases the aggregate markup if* $\phi(z)$ *is increasing and convex in z.*

Propositions 1 and 2 follow from Lemma 3. Since $\bar{q}_1$ is a sufficient statistic for the aggregate markup, a change in the distribution of buyers' incomes increases the aggregate markup if it leads to an increase in $\bar{q}_1$. For a first-order stochastic shift, this is the case when $q_1(z)$ is increasing (Proposition 1), and for a mean preserving spread, this is the case when $q_1(z)$ is increasing and convex in $z$ (Proposition 2).

**Balanced growth.** While Proposition 1 shows that a first-order stochastic shift in buyers' incomes leads to an increase in the aggregate markup, the model can generate a balanced growth path if increases in labor productivity are offset by increases in search productivity.

**Corollary 1** (Balanced Growth)**.** *Suppose $K = 1$. The aggregate markup is constant if either*

1. *Search productivity is linear in income, $a(z) = \alpha z$.*
2. *For each household $i$ in the economy, $z_i' = \gamma z_i$, and $a_i' = \gamma a_i$.*

In the first case in Corollary 1, the opportunity cost of search effort is degenerate at $\phi(z) = \phi = 1/\alpha$, and so all households pay identical markups, and changes in the income distribution have no effect on the aggregate markup in the economy. In the second case, opportunity costs of search effort differ across households, but by increasing labor productivity and search productivity at the same rate for all households, the distribution of $\phi$ is unchanged, and hence there is no change in the aggregate markup.

In other words, whether growth leads to an increase in the aggregate markup in the model depends on a *race between labor and search productivity*. If labor productivity growth outpaces search productivity growth, opportunity costs of search effort rise over time, leading to rising markups (tending in the limit toward the Diamond 1971 monopoly price equilibrium).

This result relates to a literature that asks why markups and price dispersion have not fallen over time, despite improvements in search technology. Menzio (2021) reviews an empirical literature showing constant price dispersion over time and proposes a model in which declining search frictions coincide with constant price dispersion due to increased specialization by sellers. This paper makes the complementary point that improvements in search productivity need not lead to a decline in markups and price dispersion—and can even coincide with increasing markups—if labor productivity is also rising.

## 4.8 Discussion

**Comparison to model with non-homothetic preferences.** A large literature instead attributes differences in price sensitivity to non-homothetic preferences or static utility primitives (e.g., Berry et al. 1995; Simonovska 2015; Handbury 2021; Auer et al. 2022). Such models can be calibrated to match differences in markups paid across income groups, but are unable to account for differences in prices paid for identical products within the same market or for the patterns in search behavior that I document below in Section 5. Moreover, since price sensitivity in those models are determined by utility primitives rather than endogenous search decisions, they lack the moderating feedback loop between aggregate income and price sensitivity present in the search model.[26]

**Extension with pro-competitive effects.** Jaravel (2019) and Handbury (2021) find that increasing the share of high-income households in an economy leads to lower relative prices on goods consumed by high-income households. Their results are consistent with larger markets leading to a reduction in relative prices because new firm entry has pro-competitive effects that compress markups.

The model can be extended to account for these pro-competitive effects by adding an endogenous response of search productivity to the mass of firms supplying a good:

$$a_{ik} = \bar{a}_i M_k^{\zeta}.$$

Intuitively, as more retailers enter to supply good $k$, households can collect more price quotes for $k$ in the same amount of search time. The elasticity of search productivity to the mass of firms, $\zeta$, parameterizes the strength of this effect. Appendix Figure B11 shows that $\zeta = 0.3$ can match the elasticity of prices to market size from Jaravel (2021).

My dataset on retail markups do not support a choice of $\zeta > 0$, however. Appendix Table B7 shows that, both in the cross-section of cities and in the time series, markups paid by high-income households if anything rise faster with aggregate income than markups paid by low-income households. A value of $\zeta > 0$ would instead imply that markups for high-income households rise more slowly with aggregate income, due to offsetting pro-competitive effects in goods purchased by high-income households.[27] Hence, my baseline results assume $\zeta = 0$.

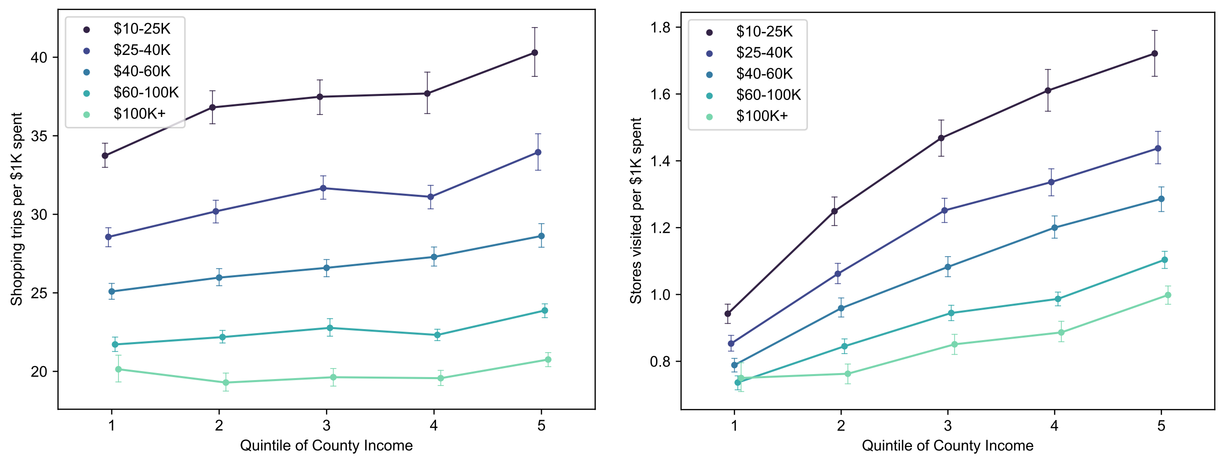
# 5 Evidence on Search Behavior

This section provides evidence for the search mechanism in the model. I show that measures of search intensity in the NielsenIQ data are consistent with two predictions

[26]A previous version of this article developed a model with non-homotheticities in elasticities of substitution and taste for quality, following Handbury (2021) and Faber and Fally (2022). When calibrated to match differences in markups paid across income groups (the micro elasticity), that model substantially overshot the macro elasticity of markups to income, due to the absence of this moderating feedback loop.

[27]The discrepancy in results may be explained by my focus on retail markups rather than price indices (compared to Handbury 2021) and my use of difference-in-differences across cities to discipline these spillovers (compared to the aggregate time series evidence in Jaravel 2021). On the latter, Sangani (2023) cautions that commodity price changes have differential effects on the markups of products consumed by low- and high-income households, potentially influencing trends in markups estimated in a short sample.

**Figure 5:** Shopping intensity by income group and county income.




**(a)** Shopping trips per $1K expenditures. **(b)** Unique stores visited per $1K expenditures.

of the model: (1) search intensity is decreasing in household income (Lemma 1), and (2) conditional on income, search intensity is increasing in high-income areas (Lemma 2).

I use two measures of shopping behavior introduced by Kaplan and Menzio (2015): the number of shopping trips a household makes and the number of unique stores it visits. I normalize both measures by dollar spent to account for the fact that total search time reflects both search intensity and the size of the consumption basket; my objective is to isolate the former. Of course, using expenditures to control for the size of the consumption basket risks confounding our results with differences in prices paid by income. Appendix Table B3 replicates the analysis instead normalizing by households' total number of transactions, number of unique UPCs purchased, or number of unique brands purchased. All three alternatives yield similar results to the ones presented here.

Figure 5 plots the two measures of shopping intensity—shopping trips per $1,000 expenditures (left panel) and unique stores visited per $1,000 expenditures (right panel)—across five income groups. The horizontal axis splits these income groups by quintile of average income in the county in which the household is based. Two patterns emerge, consistent with both predictions on search behavior in the model. First, across all county quintiles, high-income households exert less search intensity per dollar spent (Lemma 1). Second, conditional on income, households exert greater search intensity in high-income counties, consistent with strategic substitutes in search (Lemma 2). Appendix Table B3 tests these relationships formally and shows that these relationships are robust to using various measures of search intensity and to additional controls (e.g., state fixed effects and the density of grocery establishments in each county).

# 6 Calibration

I calibrate the model to match differences in basket composition and markups paid across income groups in the data. The calibrated model is able to replicate several untargeted moments in the data and from the literature.

## 6.1 Calibration procedure

Households in the model differ in two key ways: households have different tastes, generating differences in basket composition across income groups, and have different opportunity costs of search effort. I calibrate household tastes using data on households' expenditure shares across goods, and I choose opportunity costs of search to match average markups paid by income group.

**Preferences and taste shifters.** I take differences in basket composition across income groups directly from the data. To do so, I order all UPCs in the data by average buyer income and split them into $K = 10$ groups with equal sales. Appendix Figure B6 shows the fraction of expenditures by households across income levels on each of the 10 groups.[28]

I then choose parameters to exactly match these spending shares from the data. I assume preferences over goods are Cobb-Douglas ($\sigma = 1$) and choose each taste shifter, $\beta_k(z)$, to match the share of expenditures by households with income $z$ on good $k$. Note that the choice of Cobb-Douglas preferences is without loss of generality in the cross-section, since for any other value of $\sigma$, we could adjust the taste shifters $\beta_k(z)$ to continue matching observed expenditure shares. However, the assumption of Cobb-Douglas preferences can matter for counterfactuals, since these preferences will determine how spending shares evolve as relative prices across goods change. I show in Appendix Figure B12 that choosing other values for $\sigma$ has little effect on the results.[29]

**Search parameters.** I assume the mapping function $\mathcal{S} : s_i \mapsto \{q_{i,n}\}_{n=1}^{\infty}$ is Poisson, following recent work by Albrecht et al. (2021) and Menzio (2021). I take firms' unit costs as the numeraire and set households' reservation price to the 98th percentile of markups in the data, $R = 3.3$ (see Appendix Table B9 for results with alternate values for $R$). The

[28]Why choose $K = 10$? Appendix Table B8 reports results for different values of $K$ ranging from 1 to 100. For values of $K \geq 10$, there is little change in the results, suggesting that $K = 10$ adequately captures the degree of segmentation of products across income groups.

[29]I also do not find large changes in the degree of basket segmentation over time from 2004 to 2019 that would inform a different choice of functional form for preferences. Boar and Giannone (2023) also find little evidence of changes in consumption segregation over time.

**Table 3:** Calibration parameter values and sources.

| Parameter | | Value | Source |
|---|---|---|---|
| Number of products | $K$ | $10^{\dagger}$ | Increasing $K > 10$ does not change results |
| Elasticity of substitution | $\sigma$ | $1^{\dagger}$ | Cobb-Douglas |
| Quality shifters | $\beta_k(z)$ | - | Match spending shares |
| Unit wage | $w$ | 1 | Numeraire |
| Reservation price | $R$ | $3.3^{\dagger}$ | 98th percentile of markups in the data |
| Search mapping | $\mathcal{S}$ | Poisson | Albrecht et al. (2021), Menzio (2021) |
| Opp. costs of search | $\phi(z)$ | - | Match avg. markup paid by income group |
| Search productivity | $a(z)$ | - | Solved from $\phi(z) = z/a(z)$ |
| Household distribution | $H(z)$ | - | Saez and Zucman (2019) |

*Note:* † indicates that the Online Appendix reports results under alternate values.

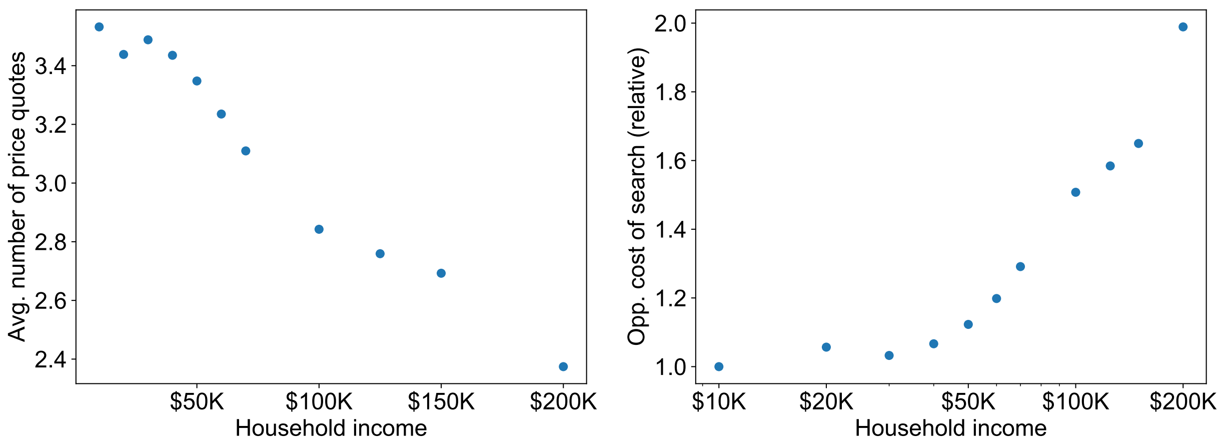
opportunity costs of search effort $\phi(z)$ can then chosen using the following two-step procedure to exactly match the average markup paid by each income group:

1. **Inner loop: Price distributions $F_k$.** Given a guess for $\phi(z)$, the price distribution $F_k$ and household search decisions $s_k(z)$ solve a fixed point for each $k$. The fixed point comes from the fact that, for a given $s_k(z)$, $F_k$ is pinned down by (10), and given $F_k$ and $\phi(z)$, search intensities $s_k(z)$ are pinned down by the household first order condition (7). I iterate the process until the price distribution $F_k$ and search intensities $s_k(z)$ converge.

2. **Outer loop: Opportunity costs of search effort $\phi(z)$.** Given an initial guess $\phi(z)$, I calculate price distributions $F_k$ for all $k$. Given $F_k$ and households' expenditure shares, there is a one-to-one mapping from households' costs of search to the average markups paid. Thus, I re-compute $\phi(z)$ to match the average markup paid by each income group. I iterate this process until $\phi(z)$ converges.

Finally, for the distribution of households across income groups, I use estimates of pre- and post-tax income by percentile of the income distribution from 1950 to 2018 from Saez and Zucman (2019). I assume markups paid and search behavior of any household with income over $200K are equal to the $200K income group, so that the results are not influenced by extrapolation beyond the income range observed in the data. My baseline results assume a unitary elasticity of expenditures to post-tax income; however, using the expenditure shares of income groups on NielsenIQ or food-at-home categories produces similar quantitative results (see Appendix Table B10).

**Figure 6:** Price quotes received and opportunity cost of search effort.




**(a)** Avg. price quotes received ($s(z) + 1$). **(b)** Opportunity cost of search effort $\phi(z)$.

## 6.2 Calibrated statistics

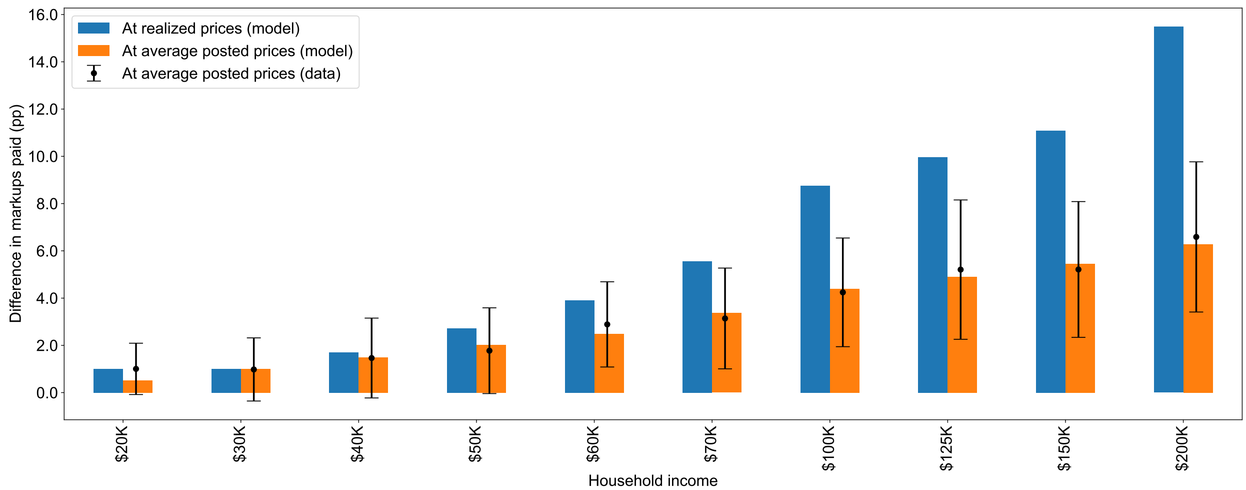
Figure 6 plots calibrated statistics on search intensity and opportunity cost of search effort by income group. The left panel shows that the expected number of price quotes, $(s(z) + 1)$, decreases with income: households with \$200K in income receive on average 30 percent fewer price quotes per purchase than low-income households. These differences in search intensity are consistent with McKenzie and Schargrodsky (2005), who find that shopping frequency for households at the ninetieth percentile of the income distribution is 30 percent lower than that of households at the tenth decile, after controlling for quantity purchased. The calibrated opportunity costs of search effort $\phi(z)$, shown in the right panel, are increasing and convex in log income, satisfying the conditions in Propositions 1 and 2.

The calibrated model is able to match several untargeted moments, including the relative contributions of search and basket composition to the markup gap across income group and spillovers of other buyers' incomes on markups paid and search intensity.

**Contributions of search and basket composition.** We can isolate the relative contributions of search and basket composition to the markup gap across households in the model by calculating the markup gap across households that would result if households instead paid average posted prices for each good. This decomposition is analogous to the one constructed in the data in Figure B5.

Figure 7 plots the overall markup gap across income groups and the counterfactual markup gap that would result if all households instead paid average posted prices for each good $k$. The overall markup gap across income groups in the model matches the data by construction, but the counterfactual markup gap from paying average posted prices is an untargeted moment. The model is able to match the contributions of search and basket composition from the data quite closely.

**Figure 7:** Contributions of search and basket composition to markup gap: Model vs. data.




*Note:* The blue bars plot differences in markups paid by each income group relative to households with below $20K income in the model. The orange bars plot differences in markups across income groups if all households instead paid the average posted price for each good. Empirical counterparts are from Appendix Figure B5.

**Spillovers and strategic interactions.** To successfully predict how markups change with aggregate income, the model must match the relationship between markups and aggregate income in the data, which depends on spillovers across households. Table 4 compares the elasticity of markups paid and search intensity to own and others' incomes in the model to the data. These moments are computed by simulating the model with the income distributions of 881 CBSAs (described in detail in Section 7) and running regressions of markups paid and price quotes received on own income and average income analogous to those in the data.

By construction, the model matches the elasticity of markups to own income in the data. The elasticity of markups paid to others' incomes, which is untargeted, is in line with the evidence on spillovers from Section 3.2 (and conservative relative to cross-sectional estimates in the data). As a result, the model produces a macro elasticity of markups to income of 10.3 percent, squarely in the 8–15 percent range estimated in the data.

The elasticity of search intensity (measured by number of price quotes received) to income in the model is −0.11, within the range of −0.08 to −0.40 estimated in the data (see

**Table 4:** Elasticity of markups paid and search intensity to own and others' incomes in the data and in the model.

| | *Log markup* | | *Search intensity* | |
|---|---|---|---|---|
| | Data | Model | Data | Model |
| Log Own Income | 0.032 | 0.032 | −0.40 | −0.11 |
| Log Others' Income | 0.102 | 0.071 | 0.09 | 0.03 |

*Note:* The first column is from Appendix Table B2, and the third column is from Table B3. Model columns are from simulations of income distributions for 881 CBSAs.

Appendix Table B3). Additionally, the model generates a positive elasticity of search intensity to others' incomes, consistent with the strategic interactions in search documented in Section 5. Finally, returns to search in the model—doubling search intensity decreases prices 7–9 percent—line up with estimates on the returns to shopping effort from Aguiar and Hurst (2007a).

# 7 Cross-Sectional Implications

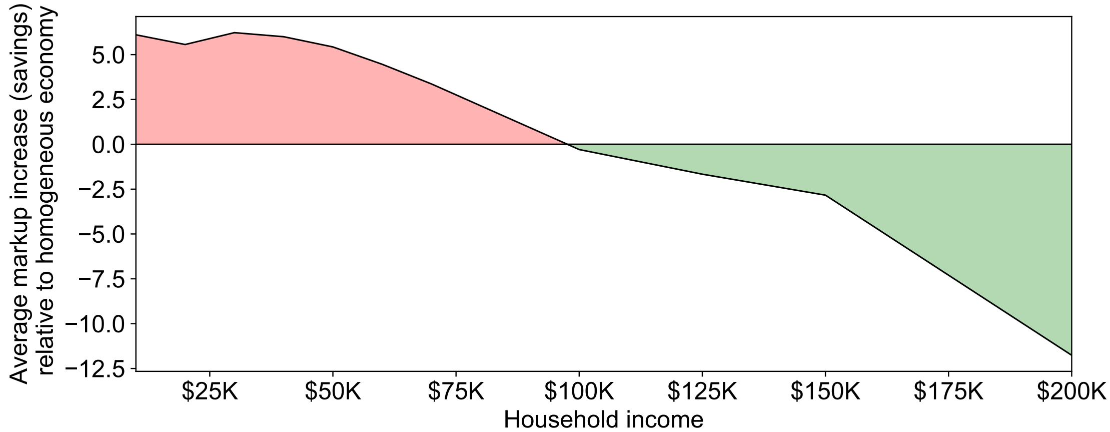
I use the calibrated model to explore spillovers across households and to predict how markups vary across cities.

**Spillovers.** In the model, households' search decisions affect the distribution of prices set by firms and hence the markups paid by other households. To quantify these spillovers, for each income level $z$, I consider an economy in which there is a representative household of that income. I compare the average markup that each household would pay in that economy to the markup it pays in the baseline calibration.

Figure 8 shows that the difference in average markups paid are large in magnitude: low-income households would pay 5pp lower markups in an economy populated with only low-income households. On the other hand, the highest-income households save over 10pp on markups paid due to the presence of low-income households. These spillovers are substantial: for example, reducing markups by 5pp would save low-income households about $250 on the $5,000 these households spend annually in the data.

These spillovers are also indicative of the distributional impact of imperfect price discrimination. If firms were able to better price discriminate across income groups (e.g., retailers abandoned the uniform pricing practices documented by DellaVigna and

**Figure 8:** Losses (savings) in markups paid relative to homogeneous income economies.




Gentzkow 2019), markup differences across income groups would increase, resulting in the gains/losses from Figure 8 at the limit of perfect price discrimination.

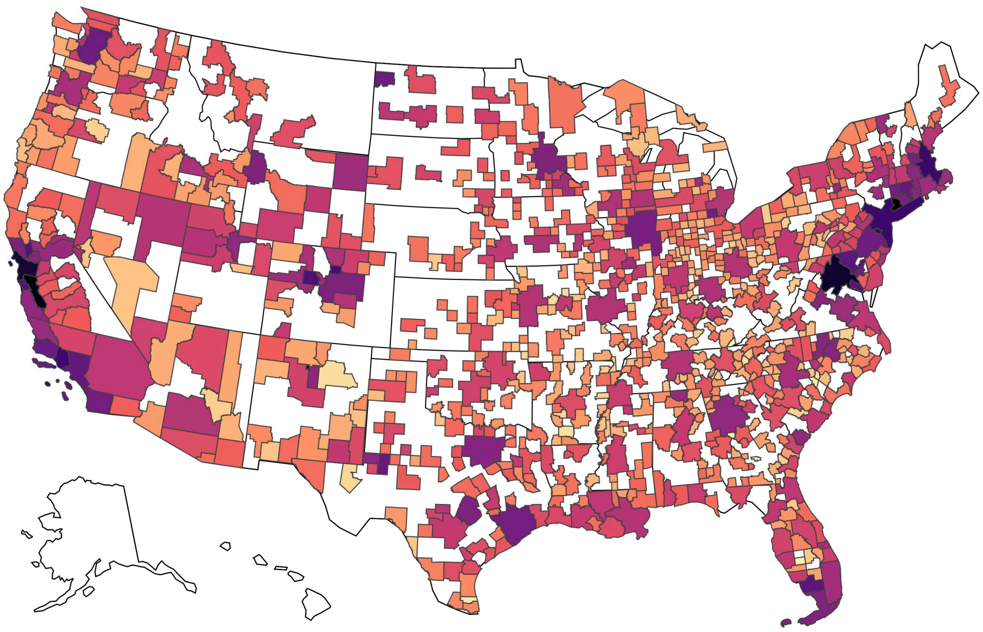
**Markups across cities.** How does the model predict markups vary across space? I use data on the income distribution of each CBSA from the American Community Survey (ACS) five-year estimates. The income bins in the ACS coincide almost one-for-one with those used by NielsenIQ, so I simulate the model pairing the income distribution of each CBSA with the search and taste parameters, $\phi(z)$ and $\beta_k(z)$, calibrated for the national data.

Figure 9 shows the predicted CBSA markups. The model predicts high markups in high-income coastal cities, such as New York, Washington, D.C., Boston, and San Francisco. Some less dense, high-income areas are also predicted to have high markups, such as Bridgeport, Connecticut; Napa, California; and Jackson, Wyoming. These patterns mirror the pattern of markups across CBSAs in the data (Appendix Figure B7).

The first two columns of Table 5 show that per-capita income and income dispersion—as measured by Gini Indices from the ACS—are important determinants of the CBSA markups predicted by the search model. These predictions are borne out in the data: the aggregate markup in a CBSA in the data rises with both per-capita income and inequality. The elasticities of CBSA markups to local income and inequality in the model and in the data line up remarkably well. Moreover, markups predicted by the model account for 31 percent of the variation in CBSA markups in the data (column 5). For comparison, a linear regression of CBSA markups on per-capita income and inequality explains only 28 percent of the variation in the data.

I also calculate CBSA markups predicted by a "supply-side" nested CES model. The nested CES model is frequently used to infer markups from data on firm market shares and

**Figure 9:** Predicted markups across CBSAs, based on CBSA income distributions.



*Note:* CBSAs are colored according to the aggregate markup predicted by the model, ranging from 29 percent (light yellow) to 46 percent (dark purple).

concentration (e.g., Smith and Ocampo 2023). Predictions from the search model isolate how demand composition affects markups, while predictions from the nested CES model focus on how supply-side factors—market shares and concentration—affect markups.[30]

Markups predicted by the nested CES model account for only 10 percent of the variation in CBSA markups and covary negatively with markups in the data. Adding these "supply-side" model markups and income measures does not greatly increase the share of variation in CBSA markups explained compared to the search model-predicted markups alone.

# 8 Changes in the Income Distribution from 1950–2018

How do changes in the income distribution over time affect markups? This section first uses moments from the data to provide model-free estimates, and then uses the calibrated search model. Through the lens of the model, changes in the income distribution from 1950–2018 account for a 12pp rise in the aggregate retail markup.

[30]To calibrate the nested CES model, I use market shares of each retailer in each CBSA for each of the $K$ product groups. The elasticity facing retailer $r$ for good $k$ is $\sigma_{rk} = \rho + s_{rk}(\eta - \rho) + s_k(1 - \eta)$, where $s_{rk}$ is retailer $r$'s market share in good $k$, $s_k$ is the market share of good $k$ across all $K$ goods, $\rho$ is the elasticity of substitution across retailers selling the same good, and $\eta$ is the elasticity of substitution across goods. I choose $\rho = 5$ and $\eta = 2$ to match the level and dispersion of markups across CBSAs in the data.

**Table 5:** Search model-predicted markups across cities rise with per-capita income and inequality, consistent with markups across cities in the data.

| | Model-Predicted | | Data | | | | |
|---|---|---|---|---|---|---|---|
| *Log CBSA Markup* | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Log CBSA Income | 0.102** | 0.096** | 0.110** | 0.102** | | | 0.018 |
| | (0.001) | (0.002) | (0.006) | (0.007) | | | (0.015) |
| Gini Index | | 0.113** | | 0.153** | | | 0.067 |
| | | (0.014) | | (0.057) | | | (0.058) |
| Log Model-Predicted Markup | | | | | 1.056** | | 0.824** |
| | | | | | (0.053) | | (0.143) |
| Log Nested CES Markup | | | | | | -0.720** | -0.122* |
| | | | | | | (0.072) | (0.074) |
| *N* | 881 | 881 | 881 | 881 | 881 | 881 | 881 |
| $R^2$ | 0.84 | 0.85 | 0.27 | 0.28 | 0.31 | 0.10 | 0.31 |

*Note:* The dependent variable in columns 1–2 is the aggregate markup for the CBSA predicted by the search model, and in columns 3–7 is the aggregate (cost-weighted average) CBSA markup in the data. Regressions weighted by total CBSA expenditures. ** is significant at 5%, * at 10%.

## 8.1 Model-free estimates

First, consider the case of complete price discrimination, in which case each household pays a certain markup irrespective of the incomes of other households. Using the mapping from income to markups paid in Figure 1, changes in the income distribution from 1950–2018 result in a 6.4pp increase in the average retail markup from 1950–2018.

Section 3.2, however, presents evidence of spillovers across households and estimates a macro elasticity of markups to income between 8–15 percent. Since post-tax per-capita real income grew 3.5 times from 1950 to 2018, starting at an average markup of 32 percent, a back-of-the-envelope calculation suggests a change in the aggregate markup of $1.32 \times \log(3.5) \times (0.08 \text{ to } 0.15) \approx 13.2 \text{ to } 23.2$pp.

This estimate does not account for potential non-linearities in the aggregate relationship or for rising income dispersion.[31] Hence, I now turn to the model.

## 8.2 Model estimates

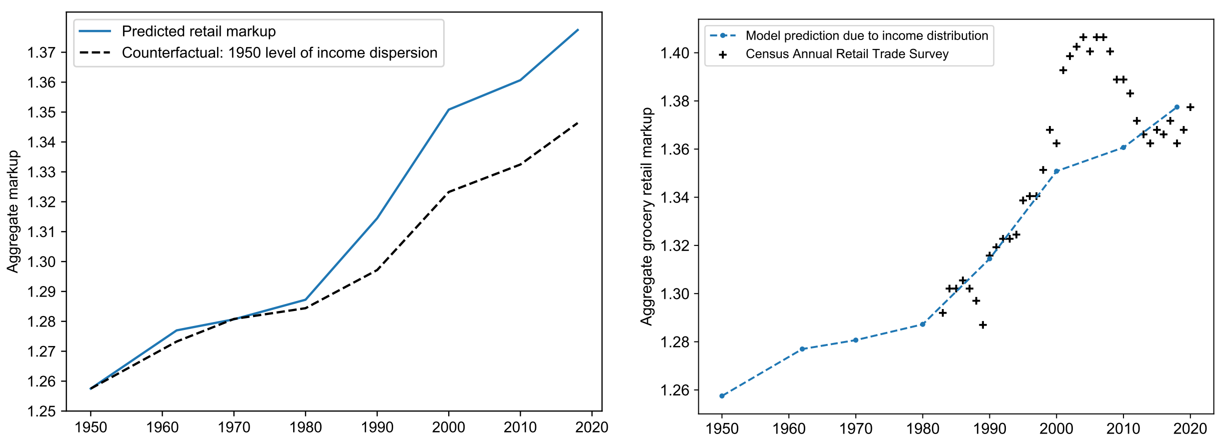
In order to simulate how changes in the income distribution affect markups over time, we need to take a stand on how search productivity evolves over time. Since the empirical

[31] According to estimates from Saez and Zucman (2019), the share of post-tax income earned by the top 10% increased 14pp from 1950 to 2018. In the cross-section of CBSAs, this increase in inequality is associated with a ≈2pp increase in markups.

evidence in Section 3.2 suggests that elasticities of markups to aggregate income in the time series and cross section are nearly identical, my baseline assumption is that the relationship between search productivity and income in the time series $a(z)$ is identical to the one estimated in the cross section.[32] Thus, for income distributions from 1950 to 2018, I assume that households with real post-tax earnings $z$ have the same search productivity $a(z)$ and opportunity cost of search effort $\phi(z)$ as households with those real post-tax earnings in 2007.

The solid blue line in the left panel of Figure 10 plots the predicted aggregate retail markup over time using the income distributions from 1950–2018 from Saez and Zucman (2019). Over this period, the model predicts a 12pp increase in the aggregate retail markup. The rise in markups is mild from 1950–1980 but accelerates significantly from 1980–2000.

**Figure 10:** Predicted aggregate retail markup under income distributions from 1950–2018.




*Note:* In the left panel, the solid line shows the predicted aggregate retail markup from 1950 to 2018, and the dotted line shows the predicted markup holding income dispersion constant at 1950 levels. In the right panel, scatter points are markups for retail grocery stores from the Census Annual Retail Trade Survey, calculated from gross margins under the assumption of constant returns.

To isolate the contribution of rising income dispersion to the increase in retail markups, the dotted black line in the left panel of Figure 10 plots the predicted retail markup holding income dispersion fixed at 1950 levels. The change in the predicted markup before 1980 is nearly identical to the change predicted under the actual income distribution. However, the two series diverge in 1980 as income dispersion rises. In 2018, the predicted markup at the 1950 level of income dispersion is 3.1pp lower than at the 2018 level of income

[32] Since $a(z)$ is increasing in $z$, search productivity improvements are "baked in" to this exercise. Appendix Figure B8 shows that long-term trends in search time in the model align with time use survey evidence from Aguiar and Hurst (2007b), and that price dispersion in the model is slightly declining over time, which is likely conservative relative to the data (see the discussion in Menzio 2021).

**Table 6:** Predicted change in aggregate retail markup from 1950–2018.

| Period | Predicted Δ in markup | Due to: Δ Income level | Due to: Δ Income dispersion | Due to: Changes in markups | Due to: Reallocations of sales |
|---|---|---|---|---|---|
| 1950–2018 | 12.0pp | 8.9pp | 3.1pp | 6.7pp | 5.3pp |
| 1950–1980 | 3.0pp | 2.7pp | 0.3pp | 1.7pp | 1.3pp |
| 1980–2018 | 9.0pp | 6.2pp | 2.8pp | 5.3pp | 3.7pp |

dispersion. Table 6 summarizes these predictions.

How does the rise in markups compare to data? The right panel of Figure 10 compares the path of the aggregate markup in the model to data on gross margins for grocery stores from the Census of Annual Retail Trade Survey.[33] An aggregate markup calculated from grocery store gross margins reported by the Census increases from 29 percent in 1983 to 38 percent in 2020.[34] The path of the aggregate retail markup predicted by the model (the dashed blue line) appears to fit the secular trend in markups in the data quite well.

Notably, markups in the data exhibit a rise and fall in the late 2000s. Changes in the income distribution over this period are not sufficient to explain the boom and bust in markups. Stroebel and Vavra (2019) document a link between retail markups and housing wealth; applying their estimates of the elasticity of markups to house prices bridges the gap between the model predictions and the data (shown in Appendix Figure B10).

**The role of reallocations.** Changes in the aggregate markup may reflect both changes in the markups set by individual firms and a compositional shift reallocating sales toward high-markup firms. Autor et al. (2020) and Kehrig and Vincent (2021) suggest that reallocation across firms has played the dominant role in increasing markups (and decreasing the labor share) in the U.S. economy.

Table 6 reports that 60 percent of the increase in the aggregate retail markup over time in the model comes from changes in the distribution of markups set by firms, while 40 percent is due to reallocations of sales across firms and products. The reallocation of sales occurs for two reasons: preferences lead households to shift their baskets to high-

[33]Grocery stores are most analogous to the set of retailers captured in the NielsenIQ data. However, Appendix Figure B13 show that other retailer categories in the Census of Annual Retail Trade Survey have seen a similar upward trend in gross margins over time. Historical data on retail gross margins also suggests that the upward trend in markups extends farther back in time, as discussed in Section 9.

[34]Anderson et al. (2018) find that average gross margins for retail firms in Compustat rose from 0.27 to 0.31 from 1982 to 2014. Under constant returns to scale, these estimates correspond to a rise in markups from 37 to 45 percent, which is of a similar magnitude to the rise in the Census ARTS data.

markup products as incomes grow, and high-income households search less and therefore buy more often from retailers with relatively high markups. Appendix Figure B9 shows that these changes in search behavior cause sales of firms at the bottom of the markup distribution for a good to fall from 1950 to 2018, while sales of firms at the top of the markup distribution for a good grow by over 10 percent. Hence, changes in the income distribution generate substantial reallocations even across firms supplying the same good.

**Implications for consumption inequality.** As stressed by Aguiar and Hurst (2007a), differences in expenditures can be a poor proxy for differences in consumption when comparing households with varying search behavior. Appendix Table B11 reports that the Gini index of costs of goods purchased—a proxy for consumption inequality—is 2.5 percent lower and has grown more slowly than the Gini index for post-tax income in the model. These estimates should be interpreted with caution, since they assume that markups paid for fast-moving consumer goods extend to the entire consumption bundle.

**Robustness to calibration choices.** Calibrating the model to within-county differences in markups paid, using household expenditure shares on NielsenIQ categories, relaxing the assumption of Cobb-Douglas preferences across goods, using different values for the number of goods $K$, or allowing for pro-competitive effects do not meaningfully change the model's predictions. Appendix Table B8 reports results for values of $K$ from 1 to 100 and inclusive of pro-competitive effects (using $\zeta = 0.3$). In all cases, the predicted rise in markups over time is between 10–16pp. Calibrating the model to match the within-county markup gap across income groups results in a predicted rise in markups of 11.2pp. Appendix Table B10 shows that the rise in markups is similar if we instead use households' expenditure shares on NielsenIQ-tracked categories or food-at-home products. Finally, Appendix Figure B12 shows that the predicted rise in markups is not sensitive to the choice of $\sigma$, since relative prices across products do not change dramatically.

# 9 Broader Questions

The above analysis prompts several broader questions. Here, I enumerate two that are addressed in the Online Appendix and outline areas for future work.

**When did the rise in retail markups begin?** The model predicts that retail markups were rising before 1980, though the rise in markups accelerated in 1980 due to growing inequality. In Appendix Figure B13, I compile historical estimates from Barger (1955) and digitized copies of the 1969–1977 Census Annual Retail Trade Surveys. These estimates suggest markups were indeed lower in level and rising far before 1980 for several retail sectors. The historical time series should be interpreted with caution, since industry definitions, methods, and samples differ across these sources.

**Markups beyond retail.** What about rising markups beyond retail? In some models of vertical supply chains, a decline in consumer price sensitivity leads markups to increase along the entire producer chain (e.g., Tirole 1988 Ch. 3; Wu 2022). In Appendix E.1.1, I provide suggestive empirical evidence for this channel, showing that De Loecker et al. (2020) markups of upstream firms are higher when they supply to retailers with high-income customers. These patterns suggest that the transmission of declining consumer price sensitivity to upstream firms merits further investigation.

# 10 Conclusion

Since Harrod (1936) first conjectured his "Law of Diminishing Elasticity of Demand," several studies have documented a relationship between income and price sensitivity. This paper extends that literature by showing that income levels and income inequality can be important determinants of markups in an economy. The relationship between markups and aggregate income arises both because of partial equilibrium effects—high-income households exert less search effort and buy a higher share of high-markup goods—and general equilibrium effects due to firm pricing decisions. Combining these effects, I find that doubling aggregate income raises markups between 8 and 15 percent.

A model that accounts for the micro evidence illuminates several macroeconomic implications, such as how changes in the income distribution shape prices faced by individual households, and why markups exhibit considerable variation across cities. The model can also account for the substantial rise in markups observed in the data over the past several decades and the reallocation of sales to high-markup firms. This evidence suggests that markups are not a purely "supply-side" phenomenon: changes in the composition of demand may play a potent role in the evolution of markups over time.

# References

Afrouzi, H., A. Drenik, and R. Kim (2021). Growing by the masses. Working paper.

Aguiar, M. and E. Hurst (2007a). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

Aguiar, M. and E. Hurst (2007b). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics 122*(3), 969–1006.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013). Time use during the great recession. *American Economic Review 103*(5), 1664–96.

Aguirregabiria, V. (1999). The dynamics of markups and inventories in retailing firms. *The Review of Economic Studies 66*(2), 275–308.

Albrecht, J., G. Menzio, and S. Vroman (2021). Vertical differentiation in frictional product markets. ' 29618, National Bureau of Economic Research.

Alessandria, G. and J. P. Kaboski (2011). Pricing-to-market and the failure of absolute ppp. *American Economic Journal: Macroeconomics 3*(1), 91–127.

Anderson, E., S. Rebelo, and A. Wong (2018). Markups across space and time. Technical Report 24434, National Bureau of Economic Research.

Auer, R. A., A. Burstein, S. Lein, and J. Vogel (2022). Unequal expenditure switching: Evidence from switzerland. Technical Report 29757, National Bureau of Economic Research.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2017). Concentrating on the fall of the labor share. *American Economic Review 107*(5), 180–85.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics 135*(2), 645–709.

Baker, S. R., B. Baugh, and M. C. Sammon (2020). Measuring customer churn and interconnectedness. 27707, National Bureau of Economic Research.

Barger, H. (1955). *Distribution's Place in the American Economy Since 1869*, Chapter "Trends in Margins", pp. 80–90. NBER.

Barkai, S. (2020). Declining labor and capital shares. *The Journal of Finance 75*(5), 2421–2463.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Boar, C. and E. Giannone (2023). Consumption segregation. Technical Report 31133, National Bureau of Economic Research.

Bornstein, G. (2021). Entry and profits in an aging economy: The role of consumer inertia. Technical report, Working Paper.

Brand, J. (2021). Differences in differentiation: Rising variety and markups in retail food stores. Working Paper.

Broda, C., E. Leibtag, and D. E. Weinstein (2009). The role of prices in measuring the poor's living standards. *Journal of Economic Perspectives 23*(2), 77–97.

Broda, C. and J. A. Parker (2014). The economic stimulus payments of 2008 and the aggregate demand for consumption. *Journal of Monetary Economics 68*, S20–S36.

Burdett, K. and K. L. Judd (1983). Equilibrium price dispersion. *Econometrica 51*(4), 955–969.

Caplovitz, D. (1963). *The Poor Pay More: Consumer practices of low-income families*. The Free Press.

Chodorow-Reich, G. (2020). Regional data in macroeconomics: Some advice for practitioners. *Journal of Economic Dynamics and Control 115*(103875).

Cohen, L. and A. Frazzini (2008). Economic links and predictable returns. *The Journal of Finance 63*(4), 1977–2011.

Coibion, O., Y. Gorodnichenko, and G. H. Hong (2015). The cyclicality of sales, regular and effective prices: Business cycle and policy implications. *American Economic Review 105*(3), 993–1029.

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics 51*(4), 1108–1161.

De Loecker, J., J. Eeckhout, and S. Mongey (2021). Quantifying market power and business dynamism in the macroeconomy. Technical Report 28761, National Bureau of Economic Research.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics 135*(2), 561–644.

DellaVigna, S. and M. Gentzkow (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics 134*(4), 2011–2084.

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory 3*, 156–68.

Döpper, H., A. MacKay, N. H. Miller, and J. Stiebale (2021). Rising markups and the role of consumer preferences. Technical report, Working Paper.

Eichenbaum, M., N. Jaimovich, and S. Rebelo (2011). Reference prices, costs, and nominal rigidities. *American Economic Review 101*(1), 234–262.

Faber, B. and T. Fally (2022). Firm heterogeneity in consumption baskets: Evidence from home and store scanner data. *The Review of Economic Studies 89*(3), 1420–1459.

Gopinath, G., P.-O. Gourinchas, C.-T. Hsieh, and N. Li (2011). International prices, costs, and markup differences. *American Economic Review 101*(6), 2450–86.

Grieco, P. L., C. Murry, and A. Yurukoglu (2021). The evolution of market power in the us auto industry. Technical Report 29013, National Bureau of Economic Research.

Griffith, R., E. Leibtag, A. Leicester, and A. Nevo (2009). Consumer shopping behavior: How much do consumers save? *Journal of Economic Perspectives 23*(2), 99–120.

Grunewald, A., J. A. Lanning, D. C. Low, and T. Salz (2020). Auto dealer loan intermediation: Consumer behavior and competitive effects. Technical Report 28136, National Bureau of Economic Research.

Gupta, A. (2020). Demand for quality, variable markups and misallocation: Evidence from india. Working paper.

Gutiérrez, G. (2017). Investigating global labor and profit shares. Working Paper.

Gutiérrez, G. and T. Philippon (2018). How eu markets became more competitive than us markets: A study of institutional drift. Technical Report 24700, National Bureau of Economic Research.

Handbury, J. (2021). Are poor cities cheap for everyone? non-homotheticity and the cost of living across u.s. cities. *Econometrica 89*(6), 2679–2715.

Harrod, R. F. (1936). *Trade Cycle. An essay.* London: Oxford University Press.

Jaimovich, N., S. Rebelo, and A. Wong (2019). Trading down and the business cycle. *Journal of Monetary Economics 102*, 96–121.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics 134*(2), 715–783.

Jaravel, X. (2021). Inflation inequality: Measurement, causes, and policy implications. *Annual Review of Economics 13*.

Kaplan, G. and G. Menzio (2015). The morphology of price dispersion. *International Economic Review 56*(4), 1165–1206.

Kehrig, M. and N. Vincent (2021). The micro-level anatomy of the labor share decline. *The Quarterly Journal of Economics 136*(2), 1031–1087.

Kesavan, S., B. R. Staats, and W. Gilland (2014). Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science 60*(8), 1884–1906.

Lach, S. (2007). Immigration and prices. *Journal of Political Economy 115*(4), 548–587.

Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *The Review of Economic Studies 60*(3), 531–542.

McKenzie, D. and E. Schargrodsky (2005). Buying less, but shopping more: Changes in consumption patterns during a crisis. Technical Report 92, Bureau for Research and Economic Analysis of Development.

Menzio, G. (2021). Optimal product design: Implications for competition and growth under declining search frictions. Technical Report 28638, National Bureau of Economic Research.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*(2), 301–317.

Nakamura, E. and D. Zerom (2010). Accounting for incomplete pass-through. *Review of Economic Studies 77*, 1192–1230.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Nevo, A. and A. Wong (2019). The elasticity of substitution between time and market goods: Evidence from the great recession. *International Economic Review 60*(1), 25–51.

Nord, L. (2022, August). Shopping, demand composition, and equilibrium prices. Working paper.

Prahalad, C. K. and A. Hammond (2002). Serving the world's poor, profitably. *Harvard Business Review 80*(9), 48–59.

Pytka, K. (2018). Shopping effort in self-insurance economies. Working Paper.

Saez, E. and G. Zucman (2019). *The triumph of injustice: How the rich dodge taxes and how to make them pay*. WW Norton and Company.

Sangani, K. (2023). Pass-through in levels and the unequal incidence of commodity shocks. Working paper.

Simonovska, I. (2015). Income differences and prices of tradables: Insights from an online retailer. *The Review of Economic Studies 82*(4), 1612–1656.

Smith, D. A. and S. Ocampo (2023). The evolution of u.s. retail concentration. Working paper.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy 69*(3), 213–225.

Stroebel, J. and J. Vavra (2019). House prices, local demand, and retail prices. *Journal of Political Economy 127*(3), 1391–1436.

Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.

Varian, H. R. (1980). A model of sales. *American Economic Review 70*(4), 651–659.

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *Review of Economic Studies 74*(2), 625–652.

Wolf, C. K. (2021). The missing intercept: A demand equivalence approach. Technical Report 29558, National Bureau of Economic Research.

Wu, L. (2022). Network lerner index: Demand and distortions across industries. Working paper.

# Online Appendix to *Markups Across the Income Distribution: Measurement and Implications*

Kunal Sangani

(Not for publication)

# Appendix A Data Cleaning and Construction

## A.1 NielsenIQ Homescan

**Continuous measure of household income.** NielsenIQ reports income in discrete categories. For analyses that use household income as a continuous variable, I follow Broda et al. (2009) and recode each household's income as the midpoint of the income bracket. For the group with over $200,000 in annual income, I assign an income of $225,000. Since incomes are topcoded at $100,000 for years outside 2006–2009, for all analyses that span other years, I assign an income of $150,000 to all households with incomes over $100,000.

**Treatment of retailer IDs.** While retailer identities are anonymized, NielsenIQ provides a retailer ID for each transaction in the data. Some retailer IDs capture multiple small retail chains. (Large retail chains, even those that are not part of NielsenIQ Retail Scanner program, are given unique retailer IDs.) For analyses that use retailer fixed effects, I use the subsample of transactions for which retailer IDs uniquely identify a retail chain. For the robustness analysis removing the largest retailers from the sample (Figure B1), I rank uniquely identified retail chains by total sales in the NielsenIQ Homescan data, assuming all retailers without unique identifiers are smaller than those identified.

## A.2 PromoData Price-Trak

The PromoData include a list of active categories and inactive product categories. I use both the active and inactive databases and drop duplicated observations in the inactive database. Since UPCs may be available in multiple pack sizes, I call each unique UPC-pack size available to a retailer an "item" in the following description.

**Data construction.** I calculate monthly wholesale prices as the minimum reported base and deal price for each item in each market in each month. Consistent with Stroebel and Vavra (2019), I find that wholesale prices are surprisingly uniform across markets: Table A1 shows that over 80 percent of items in a given month have a wholesale cost exactly equal to the modal price across markets.[35] Aggregating across different items, I assume that retailers purchase UPCs at the minimum price available to them, and so I calculate wholesale base and deal prices for each UPC in each month by taking the minimum price at which the UPC is offered across items (pack sizes) in that month. Since

[35] Stroebel and Vavra (2019) conduct a similar analysis at a quarterly level across all years using a subset of 32 markets in the wholesale cost data and find a similar figure of 78 percent.

the PromoData lack information on the quantities sold, this is a more principled approach than taking an unweighted average across items.

**Merge with NielsenIQ Homescan.** I merge monthly wholesale costs into the Homescan data using the date of the shopping trip and the product UPC. Table A2 shows summary statistics on the distribution of sales-weighted average markups across UPCs.

Table A3 reports the share of purchases by each income group matched to wholesale costs. The share of matched transactions (expenditures) increases (decreases) slightly with income. To check whether unmatched transactions are similar to the matched transactions, I compare the average relative price for matched and unmatched transactions by income group. These relative prices are calculated as the percent difference in the unit price of the purchased product relative to the average unit price for all transactions in the product module. The final two columns of Table A3 show that relative prices for unmatched products grow with income faster than relative prices of matched products, suggesting that markup differences in our matched sample will be conservative relative to the true differences in markups across income groups.

To check whether assuming uniformity of wholesale prices across markets materially affects the results, I replicate analyses by matching wholesale costs only to transactions made by households living in the same market as the wholesaler. I use a hand-constructed crosswalk from Scantrack Market IDs in the NielsenIQ Homescan panel to PromoData Price-Trak market areas. The sample matched at the wholesaler market level includes 3.0 million transactions (1.8 million at stores with unique NielsenIQ IDs). The analyses replicated using this subset of the data are reported in Table 1.

## A.3 Other Data Sources

- **BEA estimates of county and CBSA income.** Annual estimates of per-capita income by CBSA and county from 2000 to 2019 are from the BEA CAINC1 series. These income statistics use metropolitan areas delineated by the Office of Management and Budget in bulletin 20-01, which I use to map county FIPS codes to CBSAs throughout.
- **CBSA income distribution and inequality measures.** I use the number of households by income bucket (variable BE19001) and Gini indices (B19083) by CBSA from the 2009–2013 American Community Survey 5-year estimates.
- **Retail grocery establishments.** I use annual data on the number of NAICS 445 establishments (includes grocery stores, supermarkets, liquor stores, and specialty food stores) by county from the Census County Business Patterns.

**Table A1:** Uniformity of wholesale prices across markets.

| | *Measure of wholesale cost* Base Price | Deal Price |
|---|---|---|
| *Percent of items listed:* | | |
| At modal price ($\hat{w}^x_{i,m,t} = 1$) | 80.3 | 78.5 |
| Within 5% of modal price ($\lvert\hat{w}^x_{i,m,t} - 1\rvert \leq 0.05$) | 90.7 | 86.4 |
| Within 10% of modal price ($\lvert\hat{w}^x_{i,m,t} - 1\rvert \leq 0.10$) | 95.1 | 90.9 |

**Table A2:** Summary statistics for markup distribution.

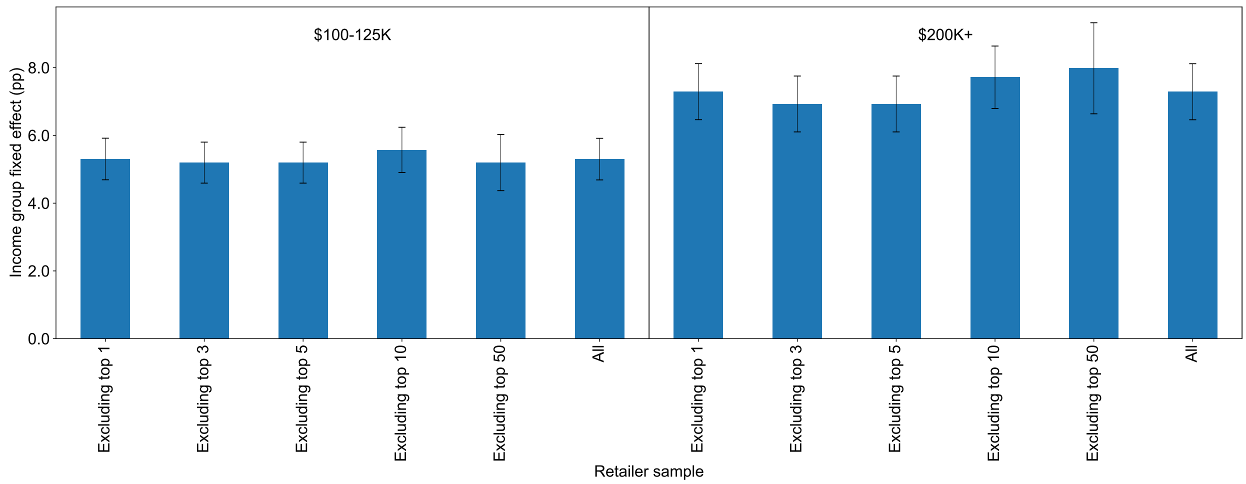
| | *Measure of wholesale cost* Base Price | Deal Price |
|---|---|---|
| *Percentiles of distribution:* | | |
| 10 | 1.053 | 1.119 |
| 25 | 1.204 | 1.288 |
| 50 | 1.382 | 1.470 |
| 75 | 1.600 | 1.694 |
| 90 | 1.911 | 2.002 |
| *Percent below $\mu = 1$:* | | |
| By count | 6.96 | 4.72 |
| By sales | 12.63 | 6.35 |

**Table A3:** Coverage of UPC wholesale cost data by income level.

| | Percent matched to wholesale cost data | | Log unit price, relative to product module average | |
|---|---|---|---|---|
| Income group | Transactions | Expenditures | Matched | Unmatched |
| \$10–25K | 41 | 38 | -0.02 | -0.05 |
| \$25–40K | 42 | 38 | 0.00 | -0.02 |
| \$40–60K | 43 | 38 | 0.04 | 0.02 |
| \$60–100K | 44 | 37 | 0.09 | 0.09 |
| Over \$100K | 44 | 35 | 0.17 | 0.17 |
| All | 43 | 37 | 0.06 | 0.05 |

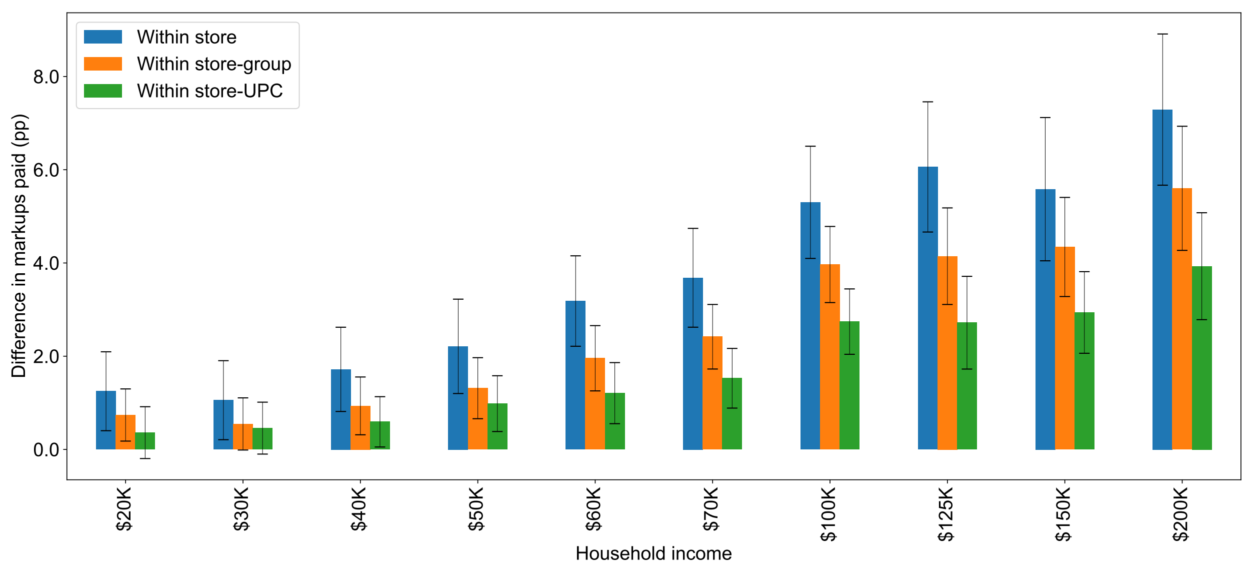
# Appendix B Additional Tables and Figures

**Figure B1:** Stability of markup gap excluding largest retail chains.



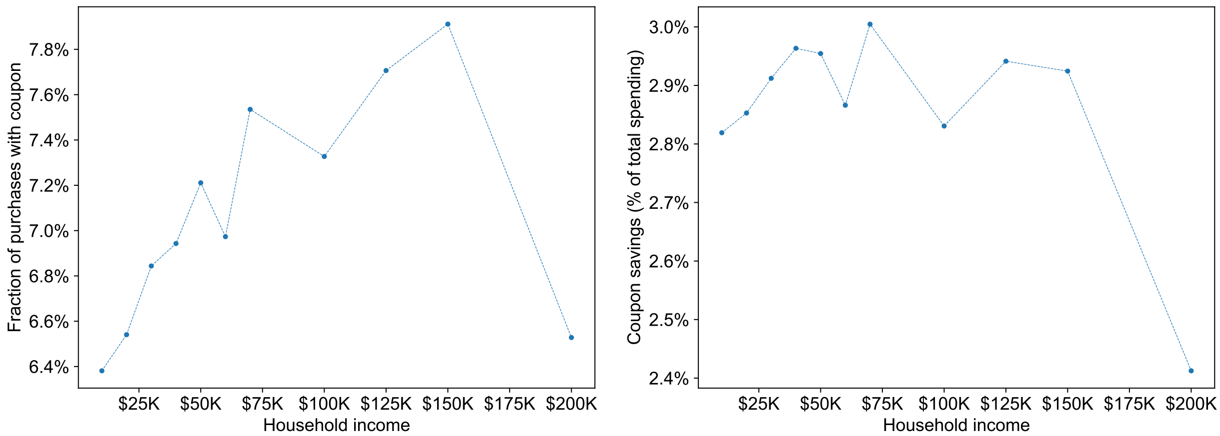

*Note:* The figure shows the fixed effect estimated in (2) for two high-income groups as large retailers (ranked by total sales in the Homescan data) are removed from the sample.

**Figure B2:** Difference in markups paid within store (blue), within store-product group (orange), and within store-UPC (green), relative to households with below $20K income.




*Note:* Blue bars include store FEs, orange bars add store-product group FEs, and the green bars add store-UPC FEs. Standard errors two-way clustered by product brand and household county.

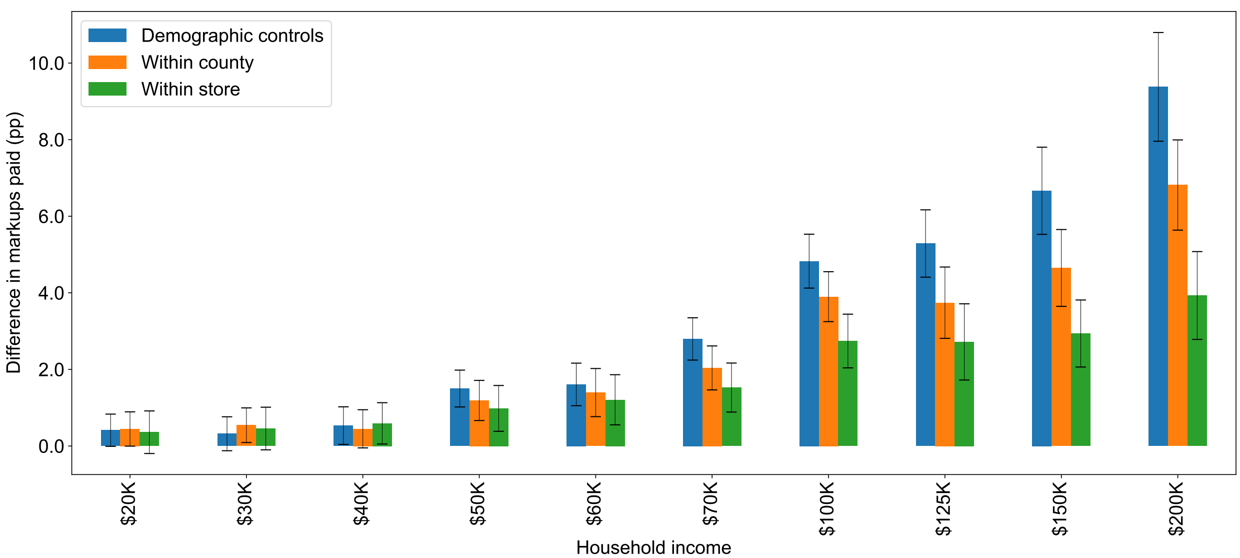
**Figure B3:** Coupon usage and savings by income.




**(a)** Fraction of purchases made with a coupon.

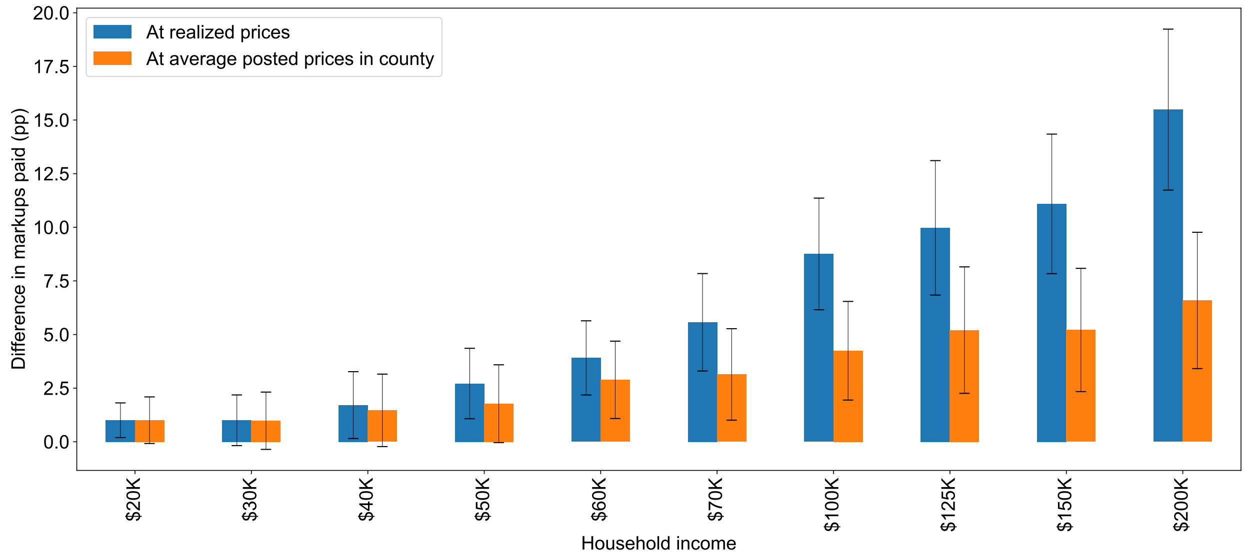
**(b)** Savings due to coupon use.

**Figure B4:** Difference in markups paid for identical products, relative to households with below $20K income.



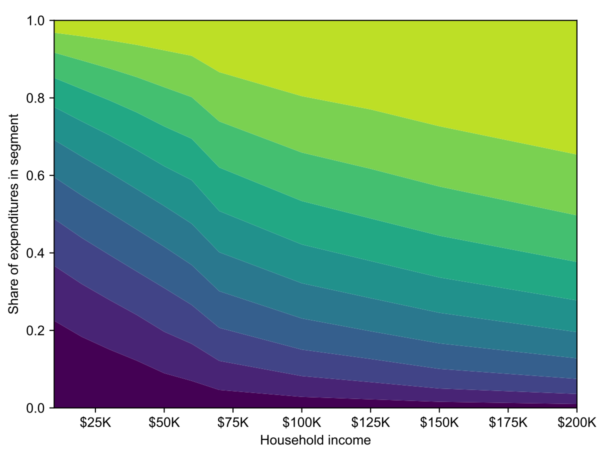

*Note:* The blue bars plot the coefficients $\beta_\ell$ from a cost-weighted regression of markup paid on household income dummies, demographic controls, and UPC fixed effects. The orange bars add county-UPC fixed effects, and the green bars add store-UPC fixed effects. Standard errors are two-way clustered by product brand and household county.

**Figure B5:** Counterfactual markup gap if households instead paid average posted prices within county for each UPC.

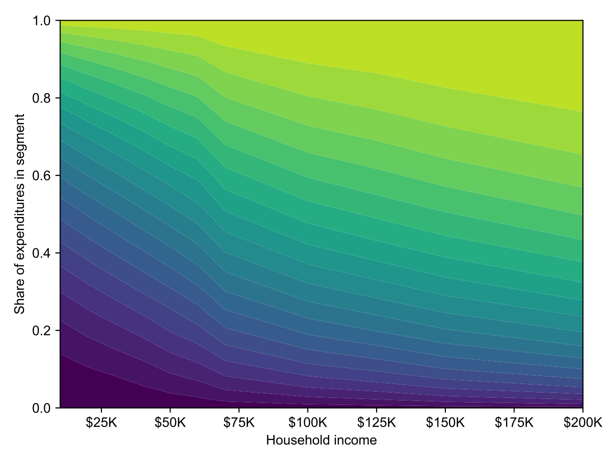



*Note:* The figure plots coefficients from a cost-weighted regression of markup paid on income dummies. The blue bars use markups constructed from households' transaction prices, while the orange bars use counterfactual markups constructed from the average NielsenIQ Retail Scanner posted price for the product in the county of purchase. Standard errors are two-way clustered by product brand and household county.

**Figure B6:** Segmentation of UPCs by household income.

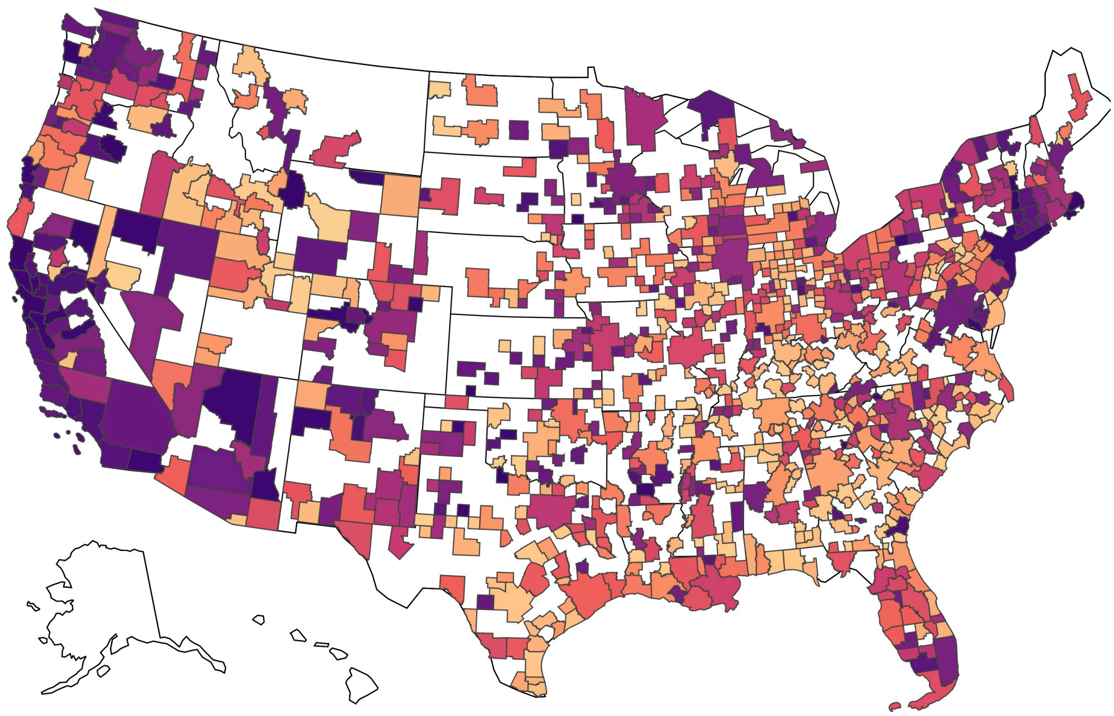



**(a)** $K = 10$ groups.



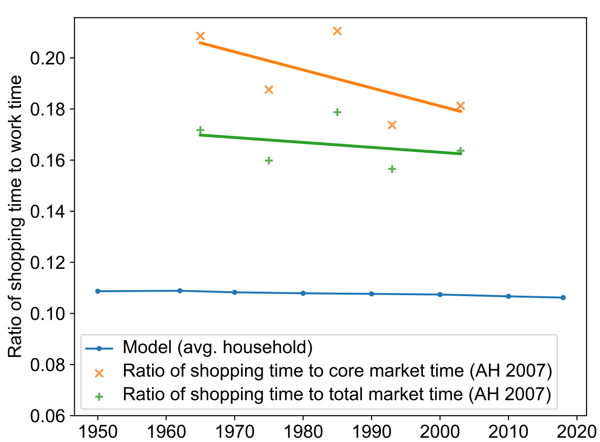

**(b)** $K = 20$ groups.

**Figure B7:** Markups across CBSAs in the data.

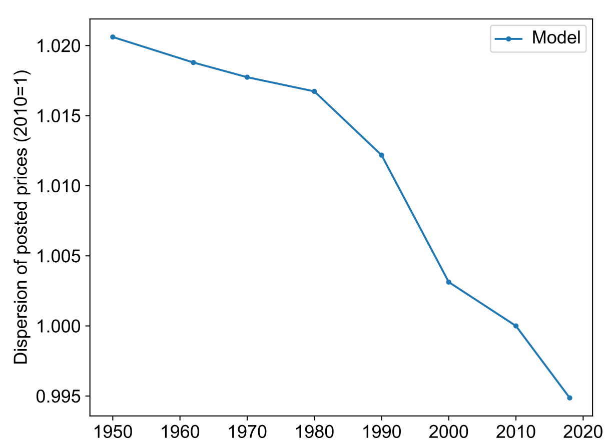


*Note:* CBSAs are colored according to expenditure-weighted percentile of aggregate markups across CBSAs in the data. Colors range from a percentile of zero (light yellow, corresponding to an aggregate markup of 0.95) to percentile of 1 (dark purple, corresponding to an aggregate markup of 1.70).

**Figure B8:** Shopping time and price dispersion in the model.

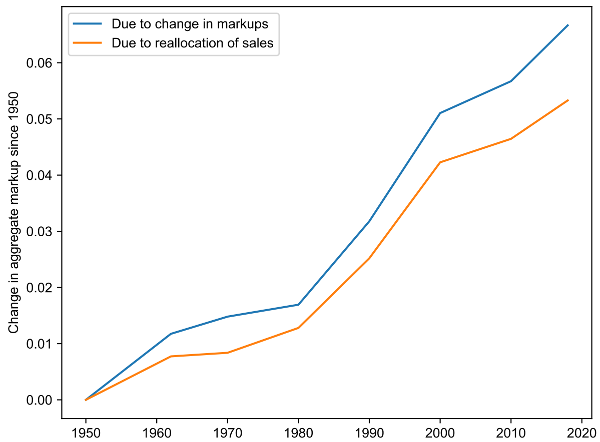



**(a)** Ratio of total shopping time to work time.

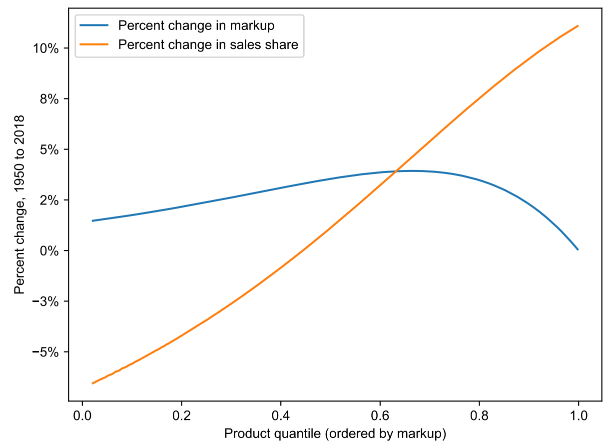



**(b)** Dispersion in posted prices.

**Figure B9:** Role of reallocations and change in the offer price distribution, 1950–2018.

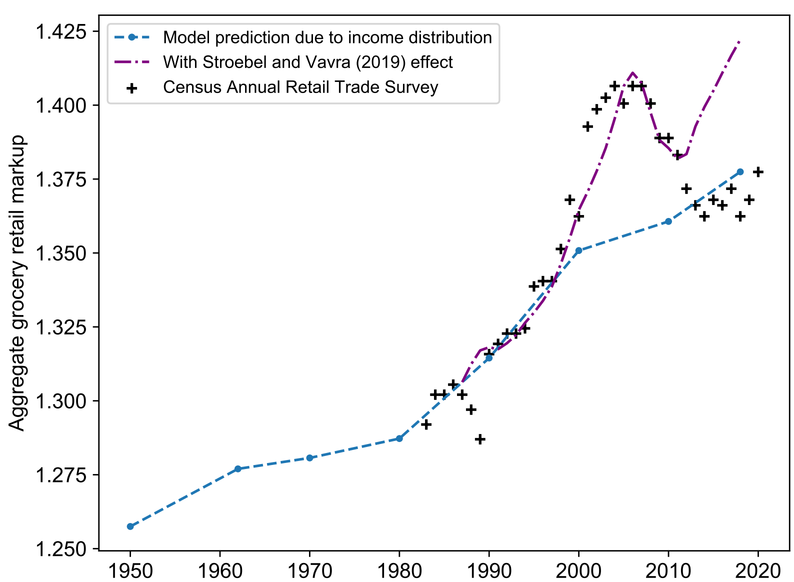



**(a)** Decomposition of change in agg. markup.



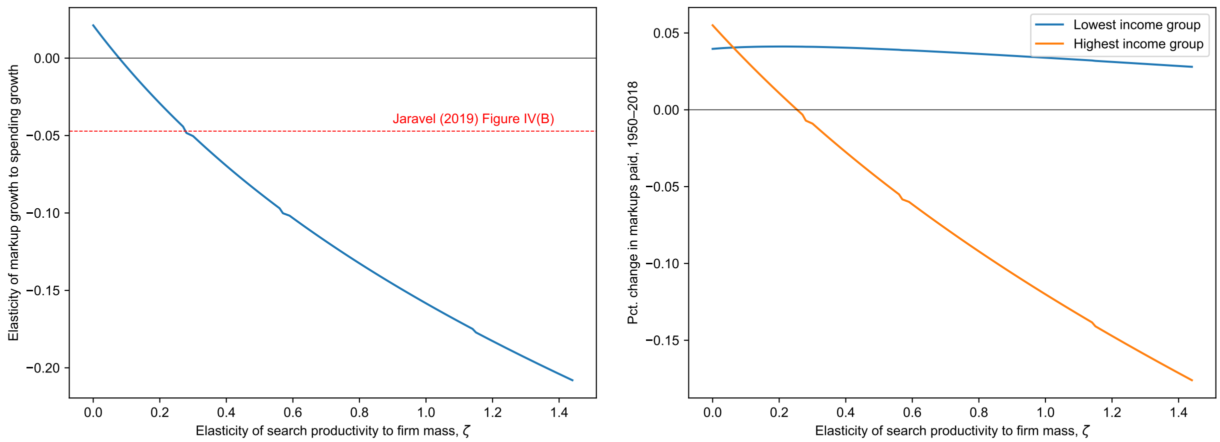

**(b)** Percent change in markups and sales shares by quantile for $k = 1$, 1950 to 2018.

**Figure B10:** Predicted aggregate retail markup from 1950–2018, adding Stroebel and Vavra (2019) effect of housing wealth.




*Note:* Scatter points are markups for retail grocery stores, converted from gross margins in the Census Annual Retail Trade Survey under the assumption of constant returns. The dashed blue line plots markups predicted by the model, and the dash-dotted purple line adds the effect of housing wealth on retail markups. I apply an elasticity of retail markups to house prices of 7% (midpoint of the range of OLS estimates from Stroebel and Vavra 2019) to changes in the S&P / Case-Shiller national home price index since 1987.

**Figure B11:** Calibrating the pro-competitive effect parameter, $\zeta$.

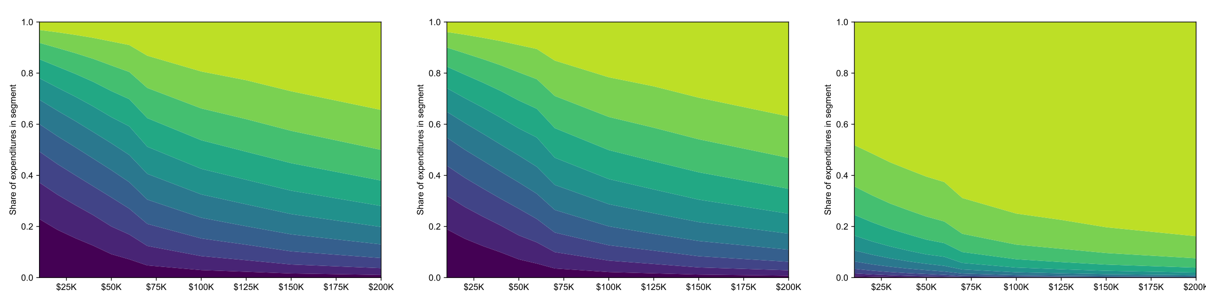



**(a)** Elasticity of markup growth to spending growth across $K = 10$ goods, 1950 to 2018.

**(b)** Change in markups paid by lowest- and highest-income households, 1950 to 2018.

*Note:* The left panel plots the coefficient estimated in a regression of markup growth on real spending growth across $K = 10$ products, simulating the model under the income distributions from 1950 and 2018 with various values of $\zeta$. The red dotted line corresponds to the OLS estimate of $-0.047$ from Jaravel (2019) Figure IV(B). The right panel plots changes in markups paid by households with \$10K and \$200K in income (2007 USD) from 1950 to 2018 under various values of $\zeta$.

**Figure B12:** Robustness to choice of $\sigma$: Model predicted spending shares in 1950 and predicted change in markups over time.

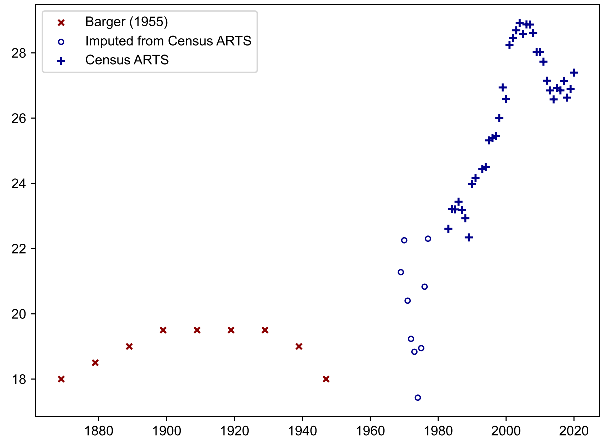



**(a)** $\sigma = 0.01$.
$\Delta\bar{\mu}_{1950-2018} = 11.995$.

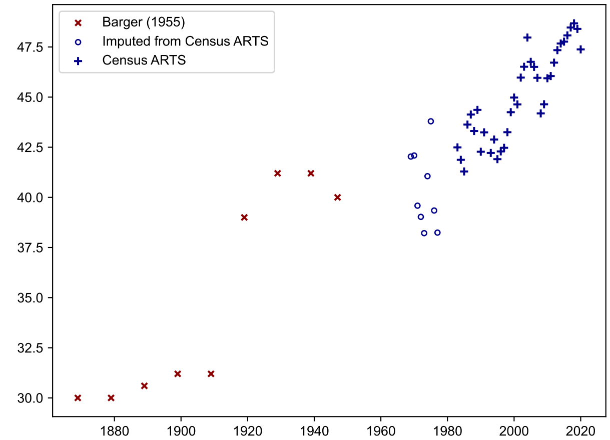
**(b)** $\sigma = 10$.
$\Delta\bar{\mu}_{1950-2018} = 11.998$.

**(c)** $\sigma = 100$.
$\Delta\bar{\mu}_{1950-2018} = 12.026$.

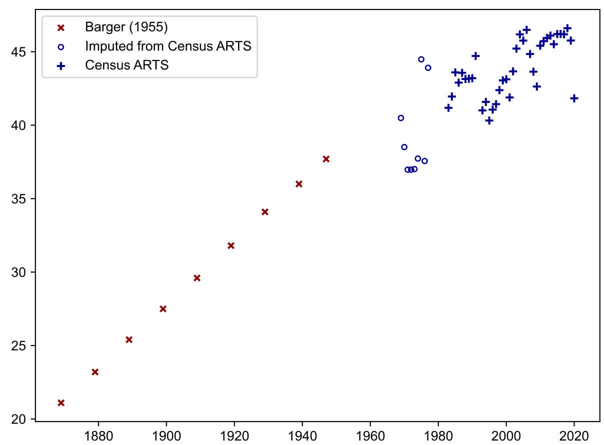
**Figure B13:** Data on retail gross margins over time by subsector.



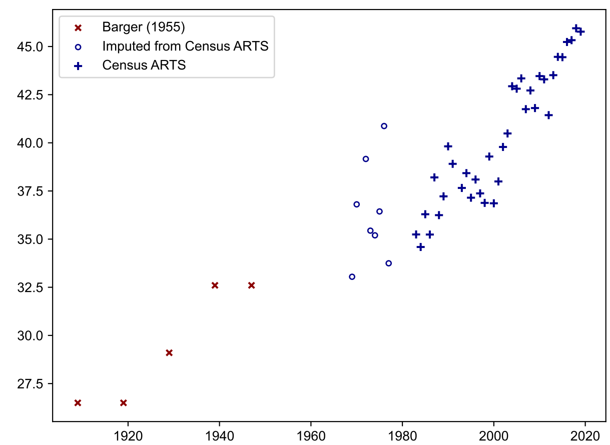

**(a)** Grocery stores




**(b)** Furniture stores




**(c)** Apparel stores




**(d)** Automobile accessory stores

*Note:* Gross margin estimates are available for selected years from 1869 to 1947 from Barger (1955), and annually from the Census Annual Retail Trade Survey from 1983 to 2020. Additionally, gross margins can be imputed from annual data on sales, purchases, and changes in inventories from the Census Annual Retail Trade Survey from 1969 to 1977. Gross margins are reported as total sales less total costs of goods sold as a percent of sales. For ARTS estimates, grocery stores include SIC 541 until 1992 and NAICS 4451 after 1993. Furniture stores include SIC 571 until 1992 and NAICS 442 after 1993. Apparel stores include SIC 56 until 1992 and NAICS 448 after 1993. Automobile parts and accessory stores include SIC 553 until 1992 and NAICS 4413 after 1993.

**Table B1:** Impact of own income on markups paid and prices paid for identical products.

| | *Log Markup* | | | | | *Log Unit Price* | |
|---|---|---|---|---|---|---|---|
| | Full Sample | | | Store ID Sample | | Full Sample | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Log Household Income | 0.032** | 0.038** | 0.029** | 0.021** | 0.017** | 0.017** | 0.020** |
| | (0.004) | (0.005) | (0.004) | (0.002) | (0.002) | (0.001) | (0.001) |
| Demographic Controls | | Yes | Yes | Yes | Yes | | Yes |
| County FEs | | | Yes | Yes | Yes | | |
| Store FEs | | | | | Yes | | |
| UPC FEs | | | | | | Yes | Yes |
| $N$ (millions) | 25.8 | 25.8 | 25.8 | 14.0 | 14.0 | 25.8 | 25.8 |
| $R^2$ | 0.00 | 0.00 | 0.02 | 0.03 | 0.09 | 0.96 | 0.96 |

*Note:* Demographic controls include race, ethnicity, household size, presence of female head of household, and age group of female head of household. Regression weighted by sales, and standard errors two-way clustered by product brand and household county. * indicates significance at 10%, ** at 5%.

**Table B2:** Impact of own and other buyers' incomes on markups paid in the cross section.

| | OLS | IV | OLS | IV | OLS | IV |
|---|---|---|---|---|---|---|
| *Log Retail Markup* | (1) | (2) | (3) | (4) | (5) | (6) |
| Log Household Income | 0.032** | 0.054** | 0.021** | 0.037** | 0.013** | 0.022** |
| | (0.004) | (0.010) | (0.003) | (0.006) | (0.001) | (0.003) |
| Log CBSA Income | 0.102** | 0.089** | | | | |
| | (0.011) | (0.011) | | | | |
| Log Income at Retailer Locations | | | 0.171** | 0.163** | | |
| | | | (0.053) | (0.051) | | |
| Log Income of Other UPC Buyers | | | | | 0.146** | 0.143** |
| | | | | | (0.043) | (0.043) |
| Demographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Store County FEs | | | Yes | Yes | Yes | Yes |
| Store FEs | | | | | Yes | Yes |
| $N$ (millions) | 23.8 | 23.8 | 9.0 | 9.0 | 14.0 | 14.0 |
| $R^2$ | 0.01 | 0.01 | 0.03 | 0.03 | 0.10 | 0.10 |

*Note:* In columns 2, 4, and 6, log household income is instrumented using the education and occupation of the male and female heads of household. Demographic controls include race, ethnicity, household size, presence of female head of household, and age group of female head of household. Regression weighted by sales, and standard errors two-way clustered by product brand and household county. * indicates significance at 10%, ** at 5%.

**Table B3:** Effect of own income and county income on search intensity.

| *Dependent variable* (Measure of search intensity) | Log Own Income Coefficient | SE | Log County Income Coefficient | SE |
|---|---|---|---|---|
| *OLS estimates*: | | | | |
| Log Shopping trips per $1K spent | -0.26** | (0.00) | 0.03 | (0.02) |
| Log Shopping trips per transaction | -0.12** | (0.00) | 0.11** | (0.03) |
| Log Shopping trips per brand bought | -0.12** | (0.00) | 0.10** | (0.04) |
| Log Shopping trips per UPC bought | -0.12** | (0.00) | 0.10** | (0.04) |
| Log Unique stores visited per $1K spent | -0.22** | (0.00) | 0.20** | (0.06) |
| Log Unique stores visited per transaction | -0.07** | (0.00) | 0.29** | (0.06) |
| Log Unique stores visited per brand bought | -0.07** | (0.00) | 0.27** | (0.05) |
| Log Unique stores visited per UPC bought | -0.08** | (0.00) | 0.28** | (0.05) |
| Log Unique retailers visited per $1K spent | -0.15** | (0.00) | 0.06** | (0.02) |
| Log Unique retailers visited per transaction | -0.01** | (0.00) | 0.14** | (0.03) |
| Log Unique retailers visited per brand bought | -0.01** | (0.00) | 0.13** | (0.03) |
| Log Unique retailers visited per UPC bought | -0.02** | (0.00) | 0.13** | (0.03) |
| *Instrumenting for household income*: | | | | |
| Log Shopping trips per $1K spent | -0.40** | (0.01) | 0.09** | (0.02) |
| Log Shopping trips per transaction | -0.30** | (0.01) | 0.19** | (0.03) |
| Log Shopping trips per brand bought | -0.27** | (0.01) | 0.16** | (0.03) |
| Log Shopping trips per UPC bought | -0.29** | (0.01) | 0.18** | (0.03) |
| Log Unique stores visited per $1K spent | -0.34** | (0.01) | 0.26** | (0.07) |
| Log Unique stores visited per transaction | -0.24** | (0.01) | 0.36** | (0.06) |
| Log Unique stores visited per brand bought | -0.21** | (0.01) | 0.33** | (0.06) |
| Log Unique stores visited per UPC bought | -0.23** | (0.01) | 0.34** | (0.06) |
| Log Unique retailers visited per $1K spent | -0.26** | (0.01) | 0.11** | (0.02) |
| Log Unique retailers visited per transaction | -0.15** | (0.01) | 0.21** | (0.03) |
| Log Unique retailers visited per brand bought | -0.12** | (0.01) | 0.18** | (0.02) |
| Log Unique retailers visited per UPC bought | -0.14** | (0.01) | 0.19** | (0.03) |

*Note:* The table reports coefficients $\beta$ and $\gamma$ estimated from a regression of various measures of search intensity (for household $i$ located in county $c$) on own income and average county income, controlling for the number of retail establishments in the county and state fixed effects:

$$\text{SearchIntensity}_{i,c} = \beta \text{Log Income}_i + \gamma \text{Log County Income}_c + \delta \text{Log Grocery Estabs}_c + \kappa_{\text{State}(c)} + \varepsilon_{i,c}.$$

Average county income is from the BEA Personal Income by County Area release. Grocery Estabs are a count of NAICS 445 establishments from Census Business Patterns (includes grocery stores, supermarkets, liquor stores, and specialty food stores.) In the second half of the table, household income is instrumented for with the education and occupation of the male and female heads of household. Standard errors clustered by household county. * indicates significance at 10%, ** at 5%.

**Table B4:** Cross section vs. time series elasticities of markups to aggregate income.

| Outcome | *Log Markup* | | *Log Avg. Unit Price* | | | |
|---|---|---|---|---|---|---|
| Sample | 2006–2012 | | 2004–2019 | | 2004 and 2019 | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log CBSA Income | 0.104** | 0.099** | 0.127** | 0.128** | 0.097** | 0.096* |
| | (0.017) | (0.026) | (0.026) | (0.030) | (0.045) | (0.052) |
| Year FEs | Yes | Yes | | Yes | | Yes |
| CBSA FEs | | Yes | Yes | | Yes | |
| Year-Product Module FEs | | | Yes | | Yes | |
| CBSA-Product Module FEs | | | | Yes | | Yes |
| *N* (millions) | 133 | 133 | 18.3 | 18.3 | 2.2 | 2.2 |
| $R^2$ | 0.03 | 0.04 | 0.99 | 0.99 | 0.99 | 0.99 |

*Note:* The sample for columns 1–2 includes all household transactions matched with wholesale cost data from 2006 to 2012. Columns 3–6 use quantity-weighted average unit prices for each product module in each CBSA, using NielsenIQ Homescan data from 2004–2019. Regression weighted by sales (in 2007 USD), and standard errors two-way clustered by CBSA and year. * indicates significance at 10%, ** at 5%.

**Table B5:** Search time increases with basket size: Cross-sectional evidence.

| | *Log Shopping Trips* | | *Log Unique Stores* | | *Log Unique Retailers* | |
|---|---|---|---|---|---|---|
| | OLS | IV | OLS | IV | OLS | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log Expenditures | 0.588** | 0.617** | 0.176** | 0.044** | 0.432** | 0.159** |
| | (0.004) | (0.013) | (0.004) | (0.012) | (0.004) | (0.013) |
| Log Avg. Markup Paid | −0.688** | −0.694** | −0.172** | −0.144** | −0.344** | −0.285** |
| | (0.025) | (0.025) | (0.028) | (0.028) | (0.025) | (0.025) |
| Demographic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Income Level FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| County FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| *N* | 63 314 | 63 314 | 63 314 | 63 314 | 63 314 | 63 314 |
| $R^2$ | 0.39 | 0.39 | 0.38 | 0.36 | 0.33 | 0.25 |

*Note:* Expenditures are total household expenditures in the NielsenIQ data, and avg. markup paid is the aggregate markup paid by the household over all observed purchases. Demographic controls include race, ethnicity, and presence and age group of female head of household. In columns 2, 4, and 6, log household size is used as an instrument for log expenditures. Standard errors clustered by household county. * indicates significance at 10%, ** at 5%.

**Table B6:** Search time increases with basket size: Time series evidence.

| | *Log Shopping Trips* | | *Log Unique Stores* | | *Log Unique Retailers* | |
|---|---|---|---|---|---|---|
| | OLS | IV | OLS | IV | OLS | IV |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Log Expenditures | 0.745** | 0.143** | 0.527** | 0.203** | 0.193** | 0.115** |
| | (0.010) | (0.029) | (0.006) | (0.032) | (0.002) | (0.029) |
| Household FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| $N$ | 917 692 | 917 692 | 917 692 | 917 692 | 917 692 | 917 692 |
| $R^2$ | 0.92 | 0.87 | 0.87 | 0.85 | 0.77 | 0.77 |

*Note:* Expenditures are total household expenditures in the NielsenIQ data. In columns 2, 4, and 6, log household income is used as an instrument for log expenditures. Standard errors two-way clustered by household and year. * indicates significance at 10%, ** at 5%.

**Table B7:** Elasticity of markups paid to CBSA income is larger for high-income households, both in cross-section and time series.

| | 2007 | All years, 2006–2012 | |
|---|---|---|---|
| *Log Retail Markup* | (1) | (2) | (3) |
| Log Avg. CBSA Income | 0.089** | 0.073** | 0.065** |
| | (0.011) | (0.018) | (0.016) |
| Log Avg. CBSA Income × Mid-Income Group | 0.016* | 0.021** | 0.006 |
| | (0.008) | (0.008) | (0.005) |
| Log Avg. CBSA Income × High-Income Group | 0.030** | 0.034** | 0.010 |
| | (0.013) | (0.013) | (0.008) |
| Year FEs | Yes | Yes | Yes |
| Demographic controls | Yes | | |
| Household FEs | | Yes | Yes |
| County FEs | | Yes | Yes |
| Store FEs | | | Yes |
| $N$ (millions) | 23.8 | 133 | 92 |
| $R^2$ | 0.01 | 0.16 | 0.18 |

*Note:* Demographic controls include race, ethnicity, and presence and age group of female head of household. Mid-Income Group is an indicator equal to one for households with incomes between $50K and $100K, and High-Income Group is an indicator equal to one for households with incomes over $100K. Regression weighted by sales (in 2007 USD), and standard errors two-way clustered by product brand and household county. * indicates significance at 10%, ** at 5%.

**Table B8:** Calibration results varying no. of products ($K$) and pro-competitive effect ($\zeta$).

| No. groups ($K$) | Predicted Δ markup 1950–2018 $\zeta = 0$ | Predicted Δ markup 1950–2018 $\zeta = 0.3$ | Markup externality low-income $\zeta = 0$ | Markup externality low-income $\zeta = 0.3$ | Markup externality high-income $\zeta = 0$ | Markup externality high-income $\zeta = 0.3$ |
|---|---|---|---|---|---|---|
| 1 | 15.7pp | 12.8pp | +10pp | +8pp | -19pp | -11pp |
| 3 | 12.6pp | 10.9pp | +7pp | +6pp | -13pp | -6pp |
| 5 | 12.2pp | 10.7pp | +6pp | +6pp | -12pp | -6pp |
| 10 | 12.0pp | 10.6pp | +6pp | +6pp | -12pp | -5pp |
| 20 | 11.9pp | 10.6pp | +6pp | +6pp | -11pp | -5pp |
| 50 | 11.8pp | 10.5pp | +6pp | +6pp | -11pp | -5pp |
| 100 | 11.8pp | 10.5pp | +6pp | +6pp | -11pp | -5pp |

**Table B9:** Change in aggregate markup from 1950–2018, varying reservation price $R$.

| $R$ | Predicted Δ in markup | Due to Δ Income level | Due to Δ Income dispersion | Due to Change in markups | Due to Reallocations of sales |
|---|---|---|---|---|---|
| 3.0 | 12.3pp | 9.0pp | 3.3pp | 6.9pp | 5.4pp |
| 3.3 (Baseline) | 12.0pp | 8.9pp | 3.1pp | 6.7pp | 5.3pp |
| 4.0 | 11.5pp | 8.6pp | 2.9pp | 6.2pp | 5.3pp |
| 5.0 | 11.1pp | 8.4pp | 2.7pp | 5.7pp | 5.3pp |

**Table B10:** Change in agg. markup from 1950–2018 for alternative expenditure shares.

| | $K = 1$ | $K = 10$ |
|---|---|---|
| Baseline (constant shares) | 15.7pp | 12.0pp |
| *Using expenditure shares by income:* | | |
| NielsenIQ-tracked categories | 14.3pp | 10.9pp |
| Food-at-home, housekeeping supplies, and personal care | 14.2pp | 10.8pp |
| Food-at-home | 13.9pp | 10.6pp |

*Note:* Expenditure shares by income from the 2007 Consumer Expenditure Survey (Table 55). The following categories are included in "NielsenIQ-tracked": food-at-home, alcoholic beverages, housekeeping supplies, small appliances and miscellaneous housewares, and personal care.

**Table B11:** Evolution of inequality in costs of goods purchased in the model over time.

| *Gini indices* | Baseline | | 1950 | 2018 | Change | |
|---|---|---|---|---|---|---|
| Post-tax income | 46.6 | – | 34.0 | 48.7 | +14.7 | – |
| Costs of goods purchased (baseline, $\zeta = 0$) | 45.5 | *-2.5%* | 33.6 | 47.5 | +14.0 | *-5.1%* |
| Incl. pro-competitive effects ($\zeta = 0.3$) | 45.5 | *-2.5%* | 33.6 | 47.5 | +13.9 | *-5.6%* |

# Appendix C Proofs

## C.1 Households

Let $p(s, F)$ denote the average price paid for a good with price distribution $F$ given search intensity $s$. Note that $p(s, F)$ can be written as

$$p(s, F) = \sum_{n=1}^{\infty} q_n(s)\mathbb{E}[p|n], \tag{12}$$

where $q_n(s)$ is the result of the search mapping $\mathcal{S} : s \mapsto \{q_n\}_{n=1}^{\infty}$ and $\mathbb{E}[p|n]$ is the expected price paid having received $n$ independent price quotes from $F$. We can write

$$\mathbb{E}[p|n] = \int_{\underline{p}}^{R} pd\left[1 - (1 - F(p))^n\right] = \int_{\underline{p}}^{R} pn\,(1 - F(p))^{n-1}\,dF(p) = \underline{p} + \int_{\underline{p}}^{R} (1 - F(p))^n dp. \tag{13}$$

**Lemma 4** (Expected price with $n$ quotes)**.** *For any nondegenerate distribution $F$, $\mathbb{E}[p|n]$ is strictly decreasing and convex in $n$. Thus, $\mathbb{E}[p|n] - \mathbb{E}[p|n+1] > \mathbb{E}[p|n+1] - \mathbb{E}[p|n+2]$ for any $n$.*

*Proof.* The first part of the lemma is immediate from derivatives of (13) with respect to $n$. Since $\mathbb{E}[p|n]$ is strictly decreasing and convex in $n$ everywhere, the function $g(n) \equiv \mathbb{E}[p|n] - \mathbb{E}[p|n+1]$ is strictly positive and decreasing in $n$, and hence $g(n) > g(n+1)$. ■

Denote the cumulative mass function of $\{q_n\}_{n=1}^{\infty}$ by $\{Q_n\}_{n=0}^{\infty}$ to rewrite (12) as:

$$p(s, F) = \sum_{n=1}^{\infty} Q_n(s)(\mathbb{E}[p|n] - \mathbb{E}[p|n+1]).$$

We can use this expression to take the first and second derivatives of the price function with respect to search effort $s$, which we denote $p_s$ and $p_{ss}$:

$$p_s = \sum_{n=1}^{\infty} \frac{dQ_n(s)}{ds}(\mathbb{E}[p|n] - \mathbb{E}[p|n+1]),$$
$$p_{ss} = \sum_{n=1}^{\infty} \frac{d^2Q_n(s)}{ds^2}(\mathbb{E}[p|n] - \mathbb{E}[p|n+1]).$$

**Lemma 5** (Properties of the price function)**.** *Under Assumption 1, for any nondegenerate price distribution $F$, $p_s < 0$ and $p_{ss} > 0$.*

*Proof.* From Lemma 4, $\mathbb{E}[p|n] - \mathbb{E}[p|n+1] > 0$, and we have assumed that $Q_n(s) \leq 0$ for all $n$ and $Q_1(s) < 0$. Thus, $p_s > 0$. Assumption 1 requires exactly that $p_{ss} > 0$. ■

Recall the households' maximization problem:

$$\max_{l_i,\{c_{ik},s_{ik}\}} u(\{c_{ik}\}) = \left(\sum_{k=1}^{K} (\beta_{ik}c_{ik})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \qquad \text{s.t.} \qquad \begin{cases} \sum_k c_{ik}s_{ik}/a_i + l_i = 1, \\ \sum_k p_{ik}c_{ik} = z_i l_i. \end{cases}$$

Combining the first order condition with respect to $l_i$ with any of the $K$ first order conditions with respect to $s_{ik}$ yields:

$$\frac{c_{ik}}{a_i}z_i + c_{ik}\frac{\partial p_{ik}}{\partial s_{ik}} = 0.$$

Dividing through by $c_{ik}$ and defining $\phi_i \equiv z_i/a_i$ yields (7) in the main text, which we can write more simply (and without subscripts) as $\phi = -p_s(s,F)$. By Lemma 5, $-p_s$ is strictly decreasing, so we can invert this equation to write $s = s(\phi, F)$. It will be helpful to note:

$$\frac{\partial s}{\partial \phi} = \frac{1}{-p_{ss}} < 0, \qquad \text{and} \qquad \frac{\partial^2 s}{\partial \phi^2} = \frac{p_{sss}}{-(p_{ss})^3}. \tag{14}$$

*Proof of Lemma 1.* Using (14), we have for any $k$,

$$\frac{\partial s}{\partial z} = \frac{\partial s}{\partial \phi}\frac{d\phi}{dz} = \frac{1}{-p_{ss}}\frac{d\phi}{dz}.$$

Thus, $\frac{\partial s}{\partial z} \propto -\frac{d\phi}{dz}$. Next, using $p(z) = p(s(z))$, we get $\frac{\partial p}{\partial z} = p_s\frac{\partial s}{\partial z} \propto \frac{d\phi}{dz}$, using the result from Lemma 4 that $p_s < 0$. ■

## C.2 Firms

For convenience, I drop subscripts $k$ except where necessary below. Given $\{\bar{q}_n\}_{n=1}^{\infty}$, recall that $\bar{q}_1$ quotes are retrieved by households receiving only one quote, $2\bar{q}_2$ quotes are retrieved by households receiving two quotes, and so on. Hence, the demand curve facing a firm charging price $p \leq R$ is

$$D(p) = \frac{C}{M}\left[\bar{q}_1 + 2\bar{q}_2(1-F(p)) + 3\bar{q}_3(1-F(p))^2 + ...\right] = \frac{C}{M}\sum_{n=1}^{\infty} n\bar{q}_n(1-F(p))^{n-1},$$

and zero for any firm charging a price $p > R$. Accordingly, variable profits at any price $p \leq R$ are

$$\pi(p) = \frac{C}{M}(p-1)\sum_{n=1}^{\infty} n\bar{q}_n(1-F(p))^{n-1}. \tag{15}$$

Our equilibrium condition for $F(p)$ is that all firms charging $p \in \text{supp}(F)$ make equal profits $\pi$, and any firm charging some $p \notin \text{supp}(F)$ will make profits strictly less than $\pi$. A firm charging the maximum price in the support of $p$ (assuming $\bar{p} \leq R$) makes profits

$$\pi(\bar{p}) = \frac{C}{M}(\bar{p} - 1)\bar{q}_1.$$

As long as $\bar{q}_1 > 0$, profits of this firm are monotonically increasing in the price it charges in the region $p \leq R$, so it is clear that $\bar{p} = R$ as long as $\bar{q}_1 > 0$. Hence, profits of all firms must be

$$\pi = \frac{C}{M}(R - 1)\,\bar{q}_1. \tag{16}$$

Setting profits in (15) equal to (16) and solving yields the expressions for $F$ and the minimum price $\underline{p}$ in (10).

*Proof of Lemma 3.* Total sales over total variable costs for good $k$, $\bar{\mu}_k$, is:

$$\bar{\mu}_k = \frac{\int_{\underline{p}_k}^{R} pD_k(p)M_k dF_k(p)}{\int_{\underline{p}_k}^{R} D_k(p)M_k dF_k(p)} = 1 + \frac{\int_{\underline{p}_k}^{R} \pi(p)M_k dF_k(p)}{C_k} = 1 + (R-1)\bar{q}_{k,1}.$$

The overall agg. markup is the cost-weighted average across the $K$ products. ■

## C.3 Equilibrium stability and strategic interactions

The equilibrium is described by a fixed point, $\bar{Q}_1 = \int_{\mathcal{I}} Q_1(s(\phi_i, \bar{Q}))di$. In response to a perturbation in any $\phi_i$, the change in the $\bar{Q}_1$ can be written as

$$d\bar{Q}_1 = \underbrace{\int_{\mathcal{I}} \frac{dQ_1(s_i)}{ds_i}\frac{\partial s_i}{\partial \phi_i} d\phi_i di}_{\text{Direct effect}} + \underbrace{\left(\int_{\mathcal{I}} \frac{dQ_1(s_i)}{ds_i}\frac{\partial s_i}{\partial \bar{Q}_1} di\right) d\bar{Q}_1}_{\text{Indirect effect through price distribution}},$$

We have written (with some abuse of notation) $\frac{\partial s}{\partial \bar{Q}_1}$ to indicate how $s$ responds to a shift in $\bar{Q}$. Since $Q_1(s)$ is strictly decreasing in $s$, a change in $\bar{Q}$, due to any one household changing its search intensity, must be accompanied by a change in $\bar{Q}_1$.

Solving for $d\bar{Q}_1$ yields

$$d\bar{Q}_1 = \frac{\int_{\mathcal{I}} \frac{dQ_1(s_i)}{ds_i}\frac{\partial s_i}{\partial \phi_i} d\phi_i di}{1 - \beta}, \qquad \text{where} \qquad \beta \equiv \int_{\mathcal{I}} \frac{dQ_1(s_i)}{ds_i}\frac{\partial s_i}{\partial \bar{Q}_1} di. \tag{17}$$

Stability requires that $\beta < 1$. Note that for $\beta < 1$, the sign of $d\bar{Q}_1$ is given by the sign of the direct effect, $\int_{\mathcal{I}} \frac{dQ_1(s_i)}{ds_i}\frac{\partial s_i}{\partial \phi_i} d\phi_i di$.

*Proof of Lemma 2.* Starting with the household first order condition, $-p_s(s_i, \bar{Q}) = \phi_i$, take the implicit derivative with respect to the change in some other household $j$'s search behavior $s_j$ to get,

$$\frac{\partial s_i}{\partial s_j} = \frac{-p_{s\bar{Q}} \frac{\partial \bar{Q}}{\partial s_j}}{p_{ss}}.$$

From Assumption 1, $p_{ss} > 0$ (Lemma 5). The numerator $-p_{s\bar{Q}}\frac{\partial \bar{Q}}{\partial s_j}$ has a natural interpretation as how returns to search change when aggregate search intensity increases.

Note that at the limit where $\phi_i \to 0$ for all $i$, $\bar{Q}_1 \to 0$, $F(p)$ approaches a degenerate distribution with point mass at $p = 1$, and since $F$ is degenerate, returns to search $-p_s = 0$.

Now suppose $\phi_i = \varepsilon > 0$ for some $i$, where $\varepsilon$ is a small number close to zero. This perturbation results in a decrease in $s_i$, and thus a strict increase in $\bar{Q}_1 > 0$. Following (10), the resulting $F(p)$ is no longer degenerate and the returns to search $-p_s > 0$. Thus, in the neighborhood of $\bar{Q}_1 = 0$, the derivative $-p_{s\bar{Q}}\frac{\partial \bar{Q}}{\partial s_j} < 0$. Hence, we can conclude that there exists some $\phi^{\text{cutoff}}$ such that if $\phi_i < \phi^{\text{cutoff}}$ for all $i$, then $\frac{\partial s_i}{\partial s_j} < 0$.

It remains to show that the equilibrium is stable. We can write $\beta$ from (17) as

$$\beta = \int_{\mathcal{I}} \frac{dQ_1(s_i)}{ds_i}\frac{-p_{s\bar{Q}_1}}{p_{ss}} di < 0 \qquad \text{for} \qquad \phi_i < \phi^{\text{cutoff}}.$$

Hence, $\beta < 0$, which satisfies our condition for stability $\beta < 1$. ■

## C.4 Comparative statics

Recall from Lemma 3 that the aggregate markup when $K = 1$ is

$$\bar{\mu} = 1 + (R-1)\bar{Q}_1.$$
$$\Rightarrow \qquad d\bar{\mu} = (R-1)d\bar{Q}_1.$$

Thus, to characterize the response of $\bar{\mu}$ to a perturbation in $\Lambda(z)$, we need to characterize the response of $\bar{Q}_1$. From (17), we need only solve for the sign of the direct effect of a change in the distribution of $\phi$'s on $\bar{Q}_1$, since equilibrium stability ensures that indirect effects do not cancel out the direct effect.

*Proof of Proposition 1.* Since $\bar{Q}_1 = \int_0^\infty Q_1(z)d\Lambda(z)$, a first-order stochastic shift in $\Lambda(z)$ increases $\bar{Q}_1$ if $Q_1(z)$ is increasing in $z$. We can write,

$$\frac{dQ_1}{dz} = \frac{dQ_1}{ds}\frac{ds}{d\phi}\frac{d\phi}{dz} = \frac{dQ_1}{ds}\frac{1}{-p_{ss}}\frac{d\phi}{dz}.$$

Note that $\frac{dQ_1}{ds} < 0$ for all $s$, and under Assumption 1, $p_{ss} > 0$. Hence, if $\frac{d\phi}{dz} > 0$, then $\frac{dQ_1}{dz} > 0$, and thus a first-order stochastic shift in $\Lambda(z)$ increases $\bar{Q}_1$ and $\bar{\mu}$. ∎

*Proof of Proposition 2.* Since $\bar{Q}_1 = \int_0^\infty Q_1(z)d\Lambda(z)$, a mean-preserving spread in $\Lambda(z)$ increases $\bar{Q}_1$ if $Q_1(z)$ is increasing and convex in $z$. We already have from above that $Q_1(z)$ is increasing in $z$ if $\phi$ is increasing in $z$ and Assumption 1 holds. Hence, we need to now find conditions under which $Q_1(z)$ is convex in $z$.

The second derivative of $Q_1$ with respect to $z$ is

$$\frac{d^2Q_1}{dz^2} = \frac{d^2Q_1}{ds^2}\left(\frac{1}{-p_{ss}}\frac{d\phi}{dz}\right)^2 + \frac{dQ_1}{ds}\frac{p_{sss}}{-(p_{ss})^3}\left(\frac{d\phi}{dz}\right)^2 + \frac{dQ_1}{ds}\frac{1}{-p_{ss}}\frac{d^2\phi}{dz^2}.$$

Again using Assumption 1, we can see that if $\frac{d^2\phi}{dz^2} > 0$, then a sufficient condition for $\frac{d^2Q_1}{dz^2} > 0$ is that

$$\frac{d^2Q_1}{ds^2} + \frac{dQ_1}{ds}\frac{p_{sss}}{-p_{ss}} \geq 0.$$

Rearranging yields,

$$\sum_{n=1}^{\infty}\left(\frac{d^2Q_{i,1}}{ds_i^2}\frac{d^2Q_{i,n}}{ds_i^2} - \frac{dQ_{i,1}}{ds_i}\frac{d^3Q_{i,n}}{ds_i^3}\right)\left[\mathbb{E}\left[p|n\right] - \mathbb{E}\left[p|n+1\right]\right] \geq 0,$$

which is exactly the condition guaranteed by Assumption 2. ∎

## C.5 Application to two-quote and Poisson cases

I show that Assumptions 1 and 2 both hold under two common parameterizations of the search mapping function $\mathcal{S}$: a two-quote case and the Poisson case.

**Two-quote.** Suppose that households always receive only one or two quotes, and that the probability of receiving two quotes is increasing in $i$'s effort according to $q_{i,2} = 1 - \exp(-s_i)$.

Assumption 1 becomes $\exp(-s_i)\left[\mathbb{E}\left[p|1\right] - \mathbb{E}\left[p|2\right]\right] > 0$, which holds since $\exp(-s_i) > 0$ and $\mathbb{E}\left[p|n\right]$ is strictly decreasing in $n$. Assumption 2 becomes:

$$\left((-\exp(-s_i))^2 - (\exp(-s_i))^2\right)\left[\mathbb{E}\left[p|1\right] - \mathbb{E}\left[p|2\right]\right] = 0 \geq 0.$$

So, we verify that the two-quote mapping satisfies both Assumption 1 and Assumption 2.

**Poisson.** Under the Poisson distribution, the mapping from $s_i$ to the probability mass function of price quotes is

$$q_{i,n+1} = e^{-s_i} \frac{s_i^n}{n!}.$$

Dropping the $i$ subscripts for convenience, some algebra yields:

$$\frac{dQ_{n+1}}{ds} = -e^{-s} \frac{s^n}{n!}, \qquad \frac{d^2Q_{n+1}}{ds^2} = \begin{cases} e^{-s} & n = 0, \\ e^{-s} \frac{s^{n-1}}{n!} (s-n) & n \geq 1. \end{cases} \qquad \frac{d^3Q_{n+1}}{ds^2} = \begin{cases} -e^{-s} & n = 0, \\ -e^{-s} (s-2) & n = 1, \\ -e^{-s} \frac{s^{n-2}}{n!} \left( (s-n)^2 - n \right) & n \geq 2. \end{cases}$$

We can verify that both Assumption 1 and Assumption 2 hold:

$$\begin{aligned} &\sum_{n=1}^{\infty} \frac{d^2Q_n}{ds^2} \left[ \mathbb{E}\left[p|n\right] - \mathbb{E}\left[p|n+1\right] \right] \\ &\qquad = e^{-s} \left( \sum_{n=1}^{\infty} \frac{s^n}{n!} \left( \left[ \mathbb{E}\left[p|n\right] - \mathbb{E}\left[p|n+1\right] \right] - \left[ \mathbb{E}\left[p|n+1\right] - \mathbb{E}\left[p|n+2\right] \right] \right) \right) > 0. \\ &\sum_{n=1}^{\infty} \left( \frac{d^2Q_1}{ds^2} \frac{d^2Q_n}{ds^2} - \frac{dQ_1}{ds} \frac{d^3Q_n}{ds^3} \right) \left[ \mathbb{E}\left[p|n\right] - \mathbb{E}\left[p|n+1\right] \right] \\ &\qquad = e^{-2s} \sum_{n=1}^{\infty} \frac{s^n}{n!} \left( \left[ \mathbb{E}\left[p|n+1\right] - \mathbb{E}\left[p|n+2\right] \right] - \left[ \mathbb{E}\left[p|n+2\right] - \mathbb{E}\left[p|n+3\right] \right] \right) \geq 0. \end{aligned}$$

# Appendix D Bias from Unobserved Local Costs

If some labor and rent costs are in fact variable, these unobserved local costs can bias the elasticity of markups to aggregate income measured across space. In this appendix, I gauge the potential magnitude of this bias using data on retail operating expenses and data on retail wages and rents.

Suppose that variable costs for a retailer with output $Y$ are given by

$$VC(Y) = cY + wL(Y) + rA(Y),$$

where $cY$ is the costs of goods sold, $wL(Y)$ are variable wage costs, and $rA(Y)$ are variable rent costs. I assume that production is constant returns and Leontief in merchandise, labor, and store space, so that labor $L(Y)$ and store space $A(Y)$ are linear in $Y$.

The aggregate markup constructed from cost of goods sold alone $\mu^{\text{COGS}}$ and the "true" aggregate markup $\mu^{\text{true}}$ are given by $\mu^{\text{COGS}} = pY/cY$ and $\mu^{\text{true}} = pY/VC(Y)$, where $pY$ are total sales. Taking the elasticity with respect to aggregate income $I$ yields:

$$\frac{d \log \mu^{\text{COGS}}}{d \log I} = \underbrace{\frac{d \log \mu^{\text{true}}}{d \log I}}_{\text{True elasticity}} + \underbrace{\frac{wL}{VC} \frac{d \log w}{d \log I} + \frac{rA}{VC} \frac{d \log r}{d \log I}}_{\text{Bias}}. \tag{18}$$

The bias in the measured elasticity of markups to income depends on two sets of statistics: the shares of wages and rent in total variable costs, and the elasticities of factor prices (e.g., wages and rents) to local income. If either the share of labor and rent in variable costs is zero, or if factor prices do not covary with local income, the bias in (18) disappears.

**Table D1:** Retail labor and rent costs from the 2007 Census Annual Retail Trade Survey.

| | All Retail | Retail Excl. Auto | Food and Beverage | Grocery stores |
|---|---|---|---|---|
| **Sales** | 3,995,182 | 3,085,043 | 547,837 | 491,360 |
| – Gross margin | 1,105,515 | 941,583 | 160,068 | 141,848 |
| **Cost of goods sold** | 2,889,667 | 2,143,460 | 387,769 | 349,512 |
| **Total operating expenses** | 873,400 | 736,072 | 128,985 | 116,175 |
| **Total labor expenses** | 468,089 | 387,610 | 74,663 | 68,375 |
| **Total rent expenses** | 90,391 | 79,666 | 11,647 | 9,806 |
| Labor cost / (OPEX + COGS) (%) | 12.4 | 13.5 | 14.4 | 14.7 |
| Rent cost / (OPEX + COGS) (%) | 2.4 | 2.8 | 2.3 | 2.1 |

*Note:* Sales and gross margins from the 2007 Census Annual Retail Trade Survey, and operating expenses from the 2007 Census Retail Trade Survey of Detailed Operating Expenses. Labor expenses include payroll, fringe benefits, contract labor costs, and commission expense. Rent expenses include lease/rental payments for stores/offices and repairs and maintenance for stores/offices.

Table D1 lists costs of goods sold, labor expenses, and rent expenses for four retail sectors from the 2007 Census Annual Retail Trade Survey of Detailed Operating Expenses. Even with expansive definitions that include all payroll, fringe benefits, commissions, and contractor costs for labor, and lease and repair expenses for both stores and offices/other buildings for rents, these are small portions of plausibly variable costs (15 and 3 percent).

I estimate the elasticities of factor costs to local income using retail wages from the Bureau of Labor Statistics Occupational Employment and Wage Statistics (OEWS) and retail rents from Moody's REIS platform; these estimates are reported in Table D2.

**Table D2:** Elasticities of retail wages and rents to CBSA income.

| Variable: | *Log Retail Wages (OEWS)* Cashiers (1) | Retail Salespersons (2) | *Log Retail Rents (REIS)* Asking Rent (3) | Effective Rent (4) |
|---|---|---|---|---|
| Log Avg. CBSA Income | 0.285** | 0.265** | 1.226** | 1.265** |
| | (0.040) | (0.027) | (0.203) | (0.208) |
| $N$ | 330 | 330 | 68 | 68 |
| $R^2$ | 0.19 | 0.23 | 0.42 | 0.43 |

*Note:* OEWS refers to the Bureau of Labor Statistics Occupational Employment and Wage Statistics. REIS refers to Moody's REIS platform. All estimates use data from 2007. Robust standard errors in parentheses. ** indicates significance at 5%.

When all labor and rent costs are fixed, $d\log\mu^{\text{true}}/d\log I = d\log\mu^{\text{COGS}}/d\log I \approx 11.0\%$ (from Table 5). At the extreme where all labor and rent costs are variable, the elasticity of markups to income falls to $d\log\mu^{\text{true}}/d\log I \approx 0.110-(0.147)(0.285)-(0.021)(1.265) = 4.2\%$. What is a reasonable proportion of costs to call variable? Using data from a large retail chain with seasonal demand, Kesavan et al. (2014) report that labor hours are 15 percent higher during peak months. If 15 percent of labor costs are variable, the true elasticity of markups to income is around 10%, implying a small bias of 1pp.

# Appendix E Comparison to Other Markup Measures

## E.1 Production Function Estimation

De Loecker et al. (2020) estimate markups for public firms in Compustat and find that average markups rose dramatically from 1980 to 2016. In this section, I show that markups of public firms estimated by De Loecker et al. (2020) are also positively associated with buyer income, using data on the income firms' customers from Baker et al. (2020).

Baker et al. (2020) estimate the distribution of buyers' incomes for firms using data on debit and credit card spending by two million individuals from 2010–2015. I calculate the average buyer income for each firm in each year as the expenditure-weighted average over the distribution of buyers' incomes in their data. Merging these measures with markups estimated by De Loecker et al. (2020) yields a set of 378 firms, including 192 retail firms (NAICS industries 44–45). I also merge data on sales concentration in NAICS-6 industries from the 2012 Economic Census, available for 152 of the 192 retail firms.

Table E1 reports estimates of the elasticity of firms' markups to average buyer income,

**Table E1:** Impact of buyer income and concentration on De Loecker et al. (2020) markups.

| *Log Production Function Markup* | All (1) | Retail Firms (NAICS 44–45) (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Log Avg. Buyer Income | 0.229** | 0.358** | 0.439** | 0.444** | 0.443** | 0.444** |
| | (0.085) | (0.067) | (0.094) | (0.094) | (0.094) | (0.095) |
| Top 4 Firms Sales Share | | | -0.101 | | | |
| | | | (0.123) | | | |
| Top 8 Firms Sales Share | | | | -0.067 | | |
| | | | | (0.117) | | |
| Top 20 Firms Sales Share | | | | | -0.082 | |
| | | | | | (0.136) | |
| Top 50 Firms Sales Share | | | | | | -0.071 |
| | | | | | | (0.139) |
| Year × NAICS-4 FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| $N$ | 1706 | 898 | 693 | 693 | 693 | 693 |
| $R^2$ | 0.76 | 0.71 | 0.68 | 0.68 | 0.68 | 0.68 |
| Within $R^2$ | 0.02 | 0.17 | 0.21 | 0.21 | 0.21 | 0.21 |

*Note:* Firm markups are calculated using replication code from De Loecker et al. (2020). Average buyer income is calculated using data on the income distribution of each firm's customers from Baker et al. (2020). The sales shares of top firms in each NAICS-6 industry are from the 2012 Economic Census. Regressions weighted by consumer spending from Baker et al. (2020), and standard errors are two-way clustered by firm and year. * indicates significance at 10%, ** at 5%.

including year by NAICS-4 fixed effects. Estimates of the elasticity of markups to buyer income range from 23 to 44 percent, whether in the full sample or limited to the sample of retail firms. The link between markups and buyer income is robust to including measures of sales concentration in NAICS-6 industries from the 2012 Economic Census.

### E.1.1 Upstream firm markups and downstream buyer income

The analysis in the main text opens the question of whether changes in consumer behavior could be responsible for changes in markups in other sectors besides retail. In standard models of vertical supply chains (e.g., Tirole 1988 Ch. 3), a reduction in the elasticity of consumer demand uniformly increases markups along the entire producer chain, and Wu (2022) shows that similar intuitions can hold in a general production network.

I use the De Loecker et al. (2020) markups to provide suggestive empirical evidence for this channel. In particular, I show that markups of upstream firms are increasing in the average income of buyers at downstream firms they supply to.

I construct a sample of matched upstream–downstream firm pairs using data from

**Table E2:** Relationship between buyer income at downstream firms with markups of upstream suppliers.

| *Markup at Upstream Firm* | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log Avg. Buyer Income of Downstream Firm | 0.103 | 0.078** | 0.085** | 0.076** |
| | (0.119) | (0.033) | (0.031) | (0.033) |
| Year FEs | Yes | Yes | Yes | Yes |
| Year-Upstream Industry FEs | | Yes | Yes | |
| Year-Downstream Industry FEs | | | Yes | |
| Year-Upstream Industry-Downstream Industry FEs | | | | Yes |
| $N$ | 9092 | 8919 | 8484 | 7765 |
| $R^2$ | 0.00 | 0.74 | 0.76 | 0.80 |

*Note:* Firm production function markups are calculated using replication code from De Loecker et al. (2020). Average buyer income is calculated using data on the income distribution of each firm's customers from Baker et al. (2020). Upstream and downstream firms are matched using data from Compustat Customer Segments. * indicates significance at 10%, ** at 5%.

Compustat Customer Segments (previously used e.g., by Cohen and Frazzini 2008), which compiles firm disclosures about customers that comprise over 10 percent of sales. I identify cases where the downstream customer is a firm (rather than geography or industry group) and construct a crosswalk from downstream firm names in the Compustat Customer Segments data to firm names used by Baker et al. (2020). Finally, I merge these linked firm pairs with data on markups of the upstream firms estimated by De Loecker et al. (2020).

Table E2 reports how markups at the upstream firms relate to average buyer income at downstream firms. Controlling for the industry of the upstream firm, the industry of the downstream firm, or both leads to a significant positive elasticity of upstream markups to income at downstream firms. For example, within a given upstream industry, doubling the income of buyers at an important retail customer is associated with an 8 percent increase in the markup of the upstream firm. Of course, these regressions measure correlations in the cross-section of firms. Further work is required to isolate causal effects of changes in price sensitivity downstream on upstream firms' markups.

## E.2 Demand Estimation

A vast literature in industrial organization estimates structural models of demand to recover marginal costs and markups from data on prices. In this section, I estimate elasticities of demand and markups in a single product module (margarine). I find that unit marginal costs and markups recovered from demand estimation are positively correlated with unit wholesale costs and retail markups.

When estimating markups in this setting, it is important to specify the form of conduct between manufacturers and retailers, since different forms of conduct affect whether the markups recovered from demand estimation are comparable to the retail markups used in the main text. For this exercise, I assume that manufacturers and retailers maximize joint surplus and then divvy that surplus according to (unmodeled) two-part tariffs or rebate arrangements. This assumption on conduct means that retail margins should form part of the overall markup recovered from demand estimation.

**Data construction.** For data on prices and market shares, I use NielsenIQ Retail Scanner data from 2006 to 2009, which includes weekly prices and sales of margarine at each store participating in NielsenIQ Retailer Scanner program.[36] I define geographic markets using designated market areas (DMAs) provided by NielsenIQ. For parsimony, I consider the 40 DMAs with the largest volume of margarine sold from 2006 to 2009: these markets account for over half of all margarine sold in the NielsenIQ Scanner data over this period.

For consumer demographics in each market, I use demographics of NielsenIQ Homescan panelists in each DMA from 2006 to 2009, weighted by NielsenIQ projection weights. Since the NielsenIQ Homescan data records the DMA in which households reside, these demographics are consistent with the DMA boundaries that I use to define markets.

I define a product as a unique retailer-UPC combination. As noted by Broda et al. (2009), retail chains offer different shopping amenities, and hence consumers may see the same UPC offered at two different retail chains as distinct products. To avoid overfitting to retailers and products with very small sales, I limit my analysis to the largest forty retailers by sales in the sample. I also isolate the top 150 UPCs by total sales and combine the remaining products with the outside good. These top 150 UPCs account for over 80 percent of the total volume of margarine sold in the DMAs in my sample. The final sample includes 338,726 observations for prices and volume across 40 DMAs in 48 months.

I construct ownership matrices using brand identifiers from NielsenIQ. The UPCs in my sample belong to 20 distinct brands. Using GS1 parent companies to construct ownership matrices does not qualitatively change the results I present here.

Since prices are endogenous, estimating the demand system requires instruments that are orthogonal to movements in demand. Following Villas-Boas (2007), I use input prices—i.e., monthly prices of soybeans, corn, oil-producing crops, and milk from the USDA National Agricultural Statistics Service—as instruments for prices.

[36] I choose margarine as the product module for demand estimation since the PromoData has especially high coverage in margarine: more than 70 percent of margarine sales in the Homescan panel in 2007 are matched to wholesale costs in PromoData.

Finally, to define market sizes, I use the NielsenIQ Homescan data to calculate average margarine consumption per household. I take the market size in each DMA in each month to be 0.95 pounds of margarine per household per month, times the number of households in the DMA. Constructed in this manner, the market share of the outside option displays considerable variation across time and space, ranging between 10 and 90 percent.

**Demand model.** I estimate a random coefficients discrete choice model of demand (Berry et al. 1995). I assume that the utility received by a household $i$ for purchasing a (retailer-UPC) product $j$ in DMA $m$ at time $t$ is

$$u_{ijmt} = \alpha_i^0 + \alpha_i^p p_{jmt} + \kappa_j + \delta_{jmt} + \varepsilon_{ijmt}, \tag{19}$$

where $\alpha_i^0$ is household $i$'s mean taste for purchasing any product, $\alpha_i^p$ determines household $i$'s price sensitivity, $p_{jmt}$ is the price of product $j$ in market $m$ at time $t$, $\kappa_j$ reflects utility from time-invariant product characteristics, $\delta_{jmt}$ is an unobserved demand shifter that varies across across products, markets, and time periods, and $\varepsilon_{ijmt}$ is an idiosyncratic draw from a Gumbel distribution. Households in market $m$ and time $t$ purchase exactly one unit of the product that gives them the highest utility (which may be the outside option).

I assume that the coefficients $\alpha_i^0$ and $\alpha_i^p$ are given by

$$\alpha_i^0 = \bar{\alpha}^0 + \Pi_{z0} z_i + \Sigma_0 \nu_i^0,$$
$$\alpha_i^p = \bar{\alpha}^p + \Pi_{zp} z_i + \Sigma_p \nu_i^p,$$

where $\bar{\alpha}^0$ and $\bar{\alpha}^p$ are means across households, $z_i$ is log of household $i$'s income, and $\Pi_{z0}$ and $\Pi_{zp}$ are interactions that allow the coefficients to vary systematically with income. The random draws $\nu_i^0$ and $\nu_i^p$, which are drawn from standard normal distributions and scaled by standard deviations $\Sigma_0$ and $\Sigma_p$, allow more flexibility for the model to match patterns of substitution across products.

**Results.** I use the `pyblp` package developed by Conlon and Gortmaker (2020) to estimate the demand system. To gain intuition, I start by estimating a logit model with the specification,

$$\log s_{jmt} = \alpha^0 - \alpha^p p_{jmt} + \gamma_j + \phi_m + \delta_{jmt}, \tag{20}$$

where $s_{jmt}$ is the market share of product $j$ in DMA $m$ at time $t$, $\gamma_j$ are product (retailer-UPC) fixed effects, which absorb other product characteristics, and $\phi_m$ are DMA fixed effects.

Column 1 of Table E3 reports results from estimating this logit model using OLS. The

**Table E3:** Demand system estimates.

| | Logit OLS<br>(1) | Logit IV<br>(2) | Random coefficients<br>(3) |
|---|---|---|---|
| Price | -0.622 | -1.063 | -1.330 |
| | (0.023) | (0.053) | (0.159) |
| $\Pi_{z0}$ | | | -10.036 |
| | | | (0.131) |
| $\Pi_{zp}$ | | | 1.324 |
| | | | (0.274) |
| Median price elasticity | 1.12 | 1.91 | 3.77 |
| Median estimated markup | 2.01 | 1.81 | 1.41 |
| Product (retailer-UPC) FEs | Yes | Yes | Yes |
| DMA FEs | Yes | Yes | Yes |
| $N$ | 338726 | 338726 | 338726 |

*Note:* Reported coefficients are estimated using the `pyblp` package developed by Conlon and Gortmaker (2020), with standard errors clustered by product. Columns 1–2 report results from the logit model specification in (20), and column 3 reports results from the random coefficients specification in (19).

coefficient on price $\alpha^p$ is negative, yielding a median price elasticity of 1.12. Column 2 uses soybean, corn, oil-producing crop, and milk prices as instruments. The resulting coefficient on price is larger in magnitude, yielding a median price elasticity of demand of 1.91 and a median markup of 1.81.

Column 3 of Table E3 reports results from estimating the random coefficients model. The estimated coefficient on price is significantly negative and similar in magnitude to the logit specification (column 2). The interaction of the value of the outside option with income, $\Pi_{z0}$, is negative, which means that high-income households are estimated to gain less utility from consuming margarine relative to an outside option. The estimated interaction of the price coefficient with income, $\Pi_{zp}$, is significantly positive, suggesting that high-income consumers are less price sensitive.[37] Incorporating consumer heterogeneity results in a higher median price elasticity of demand of 3.77.

I calculate the average marginal cost for each UPC in each month from the demand model and merge these monthly cost estimates into the NielsenIQ Homescan data for comparison with the wholesale costs and retail markups in the main text. All markup and unit marginal costs in the merged dataset are winsorized at the 1 percent level.

Panel A of Table E4 show that marginal costs recovered from demand estimation and

[37] The estimation procedure also yields estimates of $\Sigma_0$ and $\Sigma_p$, which allow for idiosyncratic variation in the utility of purchasing margarine and in price sensitivity across households. The estimates of $\Sigma_0$ and $\Sigma_p$ are not significantly different from zero.

**Table E4:** Relationship between estimated marginal costs and wholesale costs.

| *Panel A. Correlation coefficients.* | Log unit wholesale cost | | Log retail markup | |
|---|---|---|---|---|
| Using PromoData price: | Base | Deal | Base | Deal |
| Log unit marginal cost (demand est.) | 0.92 | 0.90 | | |
| Log markup (demand est.) | | | 0.54 | 0.62 |

| *Panel B. Marginal costs.* | Unit marginal cost ($/lb.) (demand est.) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| PromoData unit wholesale cost (base) | 1.224** | 0.248* | | |
| | (0.091) | (0.139) | | |
| PromoData unit wholesale cost (deal) | | | 1.183** | 0.303** |
| | | | (0.085) | (0.067) |
| Constant | -0.338** | | -0.092 | |
| | (0.106) | | (0.092) | |
| UPC FEs | | Yes | | Yes |
| $N$ | 261552 | 261552 | 261552 | 261552 |
| $R^2$ | 0.92 | 0.99 | 0.87 | 0.99 |

| *Panel C. Markups.* | Log markup (demand est.) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Log retail markup (using base wholesale cost) | 0.743** | 0.970** | | |
| | (0.111) | (0.015) | | |
| Log retail markup (using deal wholesale cost) | | | 0.794** | 0.960** |
| | | | (0.112) | (0.016) |
| Constant | 0.327** | | 0.134 | |
| | (0.121) | | (0.110) | |
| UPC FEs | | Yes | | Yes |
| $N$ | 261552 | 261552 | 261552 | 261552 |
| $R^2$ | 0.29 | 0.90 | 0.38 | 0.90 |

*Note:* Standard errors are clustered by brand. * indicates significance at 10 percent, ** at 5 percent.

wholesale costs from PromoData are closely correlated (both are measured in dollars per pound of margarine, so that differences in package size across products do not affect the results). Accordingly, the retail markups and the markups recovered from demand estimation also exhibit a strong positive correlation ($\rho \approx 0.6$).

Panel B of Table E4 shows that marginal costs of margarine recovered from demand estimation are about 30 cents per pound lower than the base wholesale costs and 10 cents per pound lower than deal wholesale costs in the PromoData. These marginal costs move slightly more than one-for-one with wholesale costs. Within-UPC variation in wholesale costs and retail markups (in panel C of Table E4) also predicts within-UPC variation in the costs and markups recovered from demand estimation.