New Estimates of US Markups and Their Reactions to Aggregate Demand*

Bulat Gafarov, Tengda Gong, Jens Hilscher

March 20, 2026

Abstract

This paper proposes new estimates of time-varying US markups and investigates the contributions of aggregate demand to their dynamics. We proceed in two steps. First, we show that annual average markups can be nonparametrically estimated based on a locally-linear demand model. We apply a revenue-weighted instrumental variable approach to estimate region-specific time-varying markups using US grocery price scanner data from 2001-2024. We find a secular trend rise and unconditional procyclical variation in markups. Second, using a dynamic panel model, we quantify the contributions of four main channels of conditional variation: income (real GDP and housing prices), the real interest rate, market concentration, and demographic trends. The income and real interest rate channels explain a large fraction of markup variation across markets and over time. Market concentration had only a small impact, while the demographic trend is insignificant. From 2008-2016 negative real interest rates increased markups counteracting downward pressure from the income channel. During COVID, higher housing prices and low interest rates pushed markups above the trend while tighter monetary policy helped them return to trend in 2024.

JEL codes: C3, E43, L13, Q11

Keywords: nonparametric demand estimation, markups, market concentration, grocery store prices, dynamic panels, monetary policy, COVID-19, aggregate demand, historical decomposition.

*Bulat Gafarov and Jens Hilscher are faculty members from the Department of Agricultural and Resource Economics at the University of California, Davis. Tengda Gong is an Assistant Professor from the Department of Agribusiness and Agricultural Economics at the University of Manitoba. Contacts: bgafarov@ucdavis.edu, tengda.gong@umanitoba.ca, jhilscher@ucdavis.edu. We would like to thank John Asker, Saroj Bhattarai, Steve Cecchetti, Jim Chalfant, James Cloyne, Thibault Fally, Paul Grieco, Leonard Goff, Oscar Jorda, Madina Karamysheva, Ricardo Reis, Daniel Sumner, Sanjay Singh, Konstantin Styrin, Alex Torgovitsky, Takuya Ura, Hal Varian, and Kaspar Wüthrich as well as participants at the 2022 AAEA meetings in Anaheim, the Agricultural Economics Seminar at UC Davis (October 2023), the Macro brownbag at UC Davis (November 2023), the 2023 Latin American Meeting of the Econometric Society, Bogota, Colombia, the 2024 ASSA meetings in San Antonio, the 2025 CEA meetings in Montreal, and the 2026 ASSA meetings in Philadelphia for helpful comments and Hovhannes Mnatsakanyan for research assistance. We also would like to thank faculty and student participants at the University of Connecticut and the University of Manitoba where Tengda presented an early version of this paper as his job market paper. A previous version of the paper was circulated under the title "Secular Rise and Pro-Cyclical Variation in Markups: Evidence from US Grocery Stores." Research for this paper was supported by a grant from the Giannini Foundation of Agricultural Economics. Part of the scanner data used in this paper comes from NielsenIQ, part from IRI. Disclaimers: (i) Researchers' 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 researchers 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. (ii) All estimates and analyses in this paper, based on data provided by IRI, are by the authors and not by IRI.

1 Introduction

Since 2001, the US has experienced the Great Recession and a global pandemic. Policymakers have reacted with large fiscal and unprecedented monetary stimuli. It is clear that these interventions have had large positive effects on overall GDP. However, the impact on the structure of the economy, including the distribution of gains from growth, is less clear. Markup over marginal cost is an important indicator that affects how the gains from economic activity are divided between consumers and producers. An important open question is how policy intervention supporting aggregate demand affects markups.

Sellers set markups to maximize profits. Markups, therefore, respond to changes in demand elasticities—if consumers become less price sensitive, markups increase. Elasticities of demand and, therefore, markups primarily depend on two factors: the income and wealth of consumers, as well as the competitive landscape in which sellers operate. Demand elasticities may change as a result of income or wealth shocks on the consumer side (Phelps and Winter, 1970; Afrouzi and Caloi, 2024). A possible mechanism is that consumers become less price sensitive when housing prices increase (Stroebel and Vavra, 2019). Demand elasticities faced by stores may also be affected by changes in the competitive landscape, leading to countercyclical markups (Rotemberg and Saloner, 1986; Rotemberg and Woodford, 1992). The relative importance of these channels will determine the aggregate cyclical pattern of optimal markups. If the former channel is stronger, we expect markups to be procyclical, while the latter channel implies countercyclical markups.

In this paper, we use new estimates of markups to study the cyclical effect of aggregate demand on variations in markups. We therefore need estimates of markups that vary over time and at a frequency that allows us to identify the effects of aggregate demand shocks on markups. We assemble new data of time-varying and region-specific markups. Following Gandhi, Navarro, and Rivers (2020), we use the first-order profit optimization condition to nonparametrically identify optimal markups. We associate the demand elasticity at the optimal price with the local average treatment effects (LATE) of log prices on log quantities (Imbens and Angrist, 1994).[1]

This general approach allows us to estimate markups when there are potentially nonlinear demand functions by calculating a local derivative close to the optimal price. It also allows us to abstract from the specific competitive structure; instead, the store-level demand functions that we estimate can be interpreted as sufficient statistics for the variable of interest.

It is possible that, in our weekly data, there are temporary deviations between optimal

[1]Unlike Gandhi, Navarro, and Rivers (2020) we do not estimate the entire non-parametric production functions, which may be time varying. Instead, we focus on the time-varying demand elasticity in a neighborhood of the equilibrium price. Gaarder, de Frahan, Mogstad, Torgovitsky, and Volpe (2024) also estimate heterogeneous coefficient linear demand models.

and actual markups, for example due to informational frictions.[2] However, at the annual level, we assume that the ratio of optimal and actual markups is mean reverting toward one. Markups estimated using our LATE parameter can be interpreted as the harmonic and revenue-weighted average of underlying time-varying optimal and, therefore, actual markups within a year. By analogy with the harmonic Laspeyres price index (Balk, 2012), we refer to this average as *harmonic markup index*.

We use the food retail sector in the US as a setting to study how aggregate demand and market structure affect the markup index over time and across markets. There are several benefits to studying this sector. It is a large sector, and for many households, expenditure on food represents a significant fraction of discretionary spending. Changes in food markups, therefore, affect a large share of the population. Moreover, recent evidence suggests that consumers form their overall inflation expectations based on grocery bills (D'Acunto, Malmendier, Ospina, and Weber, 2021). The food retail sector also covers a wide variety of geographical locations, and there are many different goods, most of which consumers purchase frequently. The store panel data are from IRI (2001-2012) and NielsenIQ (2006-2024). Our sample period includes three recessions, as well as the recent high inflation period.

We estimate time-varying own-price elasticities of demand using the well-known Hausman (1996) IV strategy.[3] We choose 25 large markets in the US in order to construct a geographically diverse sample of paired markets. The idea of Hausman is that one can use market pairs, for example, New York City and Philadelphia, in order to identify cost shocks. We estimate revenue-weighted elasticities at the food-category level so that the estimated elasticity reflects the typical consumer basket. This approach leads to accurate market-category-year elasticities (e.g. cereal in Philadelphia in 2018) that vary at the annual frequency.

We find that the Hausman IV strategy works well: (i) a test for weak instruments rejects the null for about 92% of elasticity estimations; (ii) estimates are precise under strong IV—close to 96% of their standard errors are below 0.5 and more than 98% of their t-statistics are above 1.96; and (iii) estimates are reasonable, i.e. more than 93% of them lie above one. We find large variation across markets, good categories, and, importantly, over time, validating our initial assumption of time-varying elasticities.[4]

[2]If optimal markups are constant, as is assumed in many New Keynesian models (Broer et al., 2020), variation in the gap translates to variation in actual markups. This hypothesis has been present at least since Keynes (1936). In his *General Theory*, he hypothesized that the gap increases during economy-wide downturns when aggregate demand declines. See also Eichenbaum et al. (2011). Our results instead imply that optimal markups are time-varying.

[3]This instrumental variable approach was also used in a static version in DellaVigna and Gentzkow (2019). The main other methods to estimate elasticities use demand control variables with high-frequency data (Levin, Lewis, and Wolak, 2017; Brand, 2021) or impose covariance restrictions on supply and demand shocks (Döpper et al., 2025; MacKay and Miller, 2023). There have been alternative measures of IV variables, such as production-side model-implied wholesale costs in De Loecker and Scott (2022).

[4]De Loecker and Scott (2022) also find time-varying own-price elasticities for the beer market. Their paper combines production data and retail sector data.

We construct a panel of elasticity estimates across markets and good categories to analyze unconditional variation over time. We then convert estimates to harmonic markup indices and use them to measure both lower-frequency trend and higher-frequency cyclical variation. We find a statistically significant slow upward trend in gross markups of 1.4% per year.[5]

At higher frequencies, we find evidence of unconditional procyclical markups. Our data include two substantial contractions—the 2001 dot-com recession and the great recession of 2008. The annual drops in gross markups relative to the trend were about 20% and 4%, respectively; these patterns are evidence of procyclical markups. In addition, there were several other shocks to aggregate demand in our sample period. Some examples include tightening of monetary policy in 2017, the passage of the tax cut and jobs act (TCJA) in 2017, and the 2020 COVID pandemic, which also saw large changes in markups. Overall, we find that cyclical variation tended to be larger than trend variation. We note that our sample period started and ended with drops in markups that move against the trend. Average gross markups therefore increased by only 3.3% from 2001 to 2024.[6]

Our evidence so far is based on unconditional variation. The unconditional procyclical variation suggests that income and wealth channels are stronger than those driving changes in the competitive landscape. In the second half of the paper, we use a dynamic partial adjustment model (Balestra and Nerlove, 1966) to estimate the effects of different channels on markups. This structural dynamic panel provides estimates of effects of market-specific and aggregate factors from time-varying cross-sectional heterogeneity in markups. This analysis allows us to measure the effects of various channels, including those resulting from policy, on markups while controlling for other factors. We can therefore identify conditional cyclicality, i.e. measure the conditional contribution of various channels on markups.

We include four sets of variables: income, the real interest rate, market concentration, and demographic trends. We construct a structural historical decomposition of the markups conditional on one channel at a time in a manner similar to the SVAR literature (Bernanke, Gertler, Watson, Sims, and Friedman, 1997; Bachmann and Sims, 2012).

When explaining variation in markups we use measures capturing both factors affecting income and wealth of consumers and those affecting market structure. We find several effects. First, we identify a positive contribution of the household income channel, measured by per capita GDP. When people's income declines, for example, because of a more severe recession than experienced by other markets, markups decrease as demand elasticities increase. Aguiar, Hurst, and Karabarbounis (2013) find that consumers become more price sensitive in response to income losses, e.g. by shopping around more. We also find that higher housing prices result in higher markups in both the current year and the long-run. The positive effect may come

[5]Gross markup equals price divided by marginal cost or equivalently $1/(1 - LernerIndex)$.

[6]The average gross markup increased by 11.5% from 2006 to 2019. Equivalently, the Lerner Index increased by 14.1%, which is comparable to but smaller than Döpper et al. (2025).

from the wealth channel for home owners (Stroebel and Vavra, 2019).

Second, an increase in population leads to lower elasticities and therefore higher markups. An increase in population in a specific market might be a sign of improved economic performance or anticipated income growth, potentially making consumers less price sensitive. This channel, however, is not statistically significant.

Third, we estimate the contribution of market structure, measured using the Herfindahl-Hirschman index of store concentration (HHI) and product concentration (the number of product codes per category). These market channels, especially the HHI index had a countercyclical effect on markups, in particular during the COVID period.

Fourth, we find a strong effect of the real interest rate channel. Lower real rates, e.g. negative rates, lead to higher markups. This suggests a direct effect of inter-temporal substitution on aggregate demand. Importantly, the effect is present after controlling for current household incomes, proxied by regional real GDP per capita. Our finding highlights an important transmission mechanism of monetary policy that acts through changes in optimal markups. Monetary policy therefore has a direct impact on prices through markups.

Based on the estimation results of our dynamic panel regression model, we perform a structural historical decomposition into the four major groups of factors that we have identified. We find that these factors can explain variation both over time and across markets. This exercise also allows us to observe conditional cyclicality of the different channels. A major driver of markups over the sample were changes in income and wealth. Market-concentration measures put countercyclical pressure on markups, while population had only a small impact.

The real interest rate put important supportive pressure on markups during the zero-interest-rate period of 2008 to 2016. This effect is separate from that of income and wealth, which, during this period, applied significant downward pressure on markups. As such, one can interpret the accommodative monetary policy as stabilizing prices. In contrast, income and interest rates both put upward pressure on markups from 2020 to 2022. Therefore, during this period, low interest rates combined with macro factors delivered the highest markup growth during our sample period. Monetary policy, therefore, did not stabilize prices and contributed to the highest levels of food inflation in decades.

We can also use our new estimates of time-varying markups to learn about variation in marginal costs. We deflate food-at-home CPI prices by the overall CPI and then use our estimates of markups to calculate implied marginal costs. Over the sample period, food at home and overall CPI rose by almost the same amount (76.8% and 77.2%), which implies that real food prices stayed approximately constant. However, there was substantial variation in marginal costs. Consistent with Coibion and Gorodnichenko (2015); Gafarov et al. (2025), we find evidence of marginal costs changing in response to fuel costs, which are an important component of variable input costs.

Our paper is related to two main strands of the literature, demand elasticity estimation

and markup dynamics. There are two related papers that also estimate annual elasticities of demand in the US food retail sector. Brand (2021) considers a subset of nine food categories. Döpper, MacKay, Miller, and Stiebale (2025) use the Berry, Levinsohn, and Pakes (1995) (BLP) method to estimate pooled US-level elasticities for 133 categories available in the NielsenIQ scanner data, including many non-food categories. Neither paper uses an IV approach. As a result, both may underestimate elasticities and overestimate markups. Also, neither paper analyzes the effects of both aggregate demand and market concentration factors on the cyclicality of markups.[7] Another difference from the existing literature is our focus on revenue-weighted elasticities and markups.

Our paper contributes to the literature on markup dynamics. A prominent example is De Loecker, Eeckhout, and Unger (2020) who document that since the 1980s, there has been a long-run upward trend in markups. Their analysis uses average cost data from financial statements, which may result in some challenges related to time variation in production functions and aggregation when studying time patterns (Bond et al., 2021; De Ridder et al., 2024; Grieco et al., 2024). When using marginal cost data instead, it is possible to directly estimate markups. Anderson, Rebelo, and Wong (2023) proxy for marginal cost using replacement cost data for a large US retailer from 2006 to 2009. They find that during their sample period, markups are mildly unconditionally procyclical. However, they do not analyze conditional cyclicality, which is the focus of Nekarda and Ramey (2020). Based on an aggregate production function, they use the labor share in the economy to proxy for time-varying markups. In a monetary SVAR, they find conditional procyclicality of markups in response to monetary policy driven demand shocks. In contrast, our approach uses disaggregated data and controls for time-varying changes in income and wealth measures and market structure.

The remainder of the paper is organized as follows. Section 2 introduces our methodological approach to elasticity and markup estimation. We discuss the theoretical IV approach to elasticity estimation, the retail scanner data that we use, the data choices we make, and the estimation results. Section 3 presents results on unconditional trend and cyclical variation in elasticities and markups. Section 4 analyzes factors that drive markup dynamics, and their historical contributions. Section 5 discusses the implications of these results. Section 6 concludes the paper.

[7] Döpper et al. (2025) allow for variation in demand side factors through demographic factors in their random coefficient Gaussian mixture model while assuming that market concentration stays constant over time. In contrast, we model the combined effect of consumer demand and market concentration using non-parametric estimates of store-level demand elasticities.

2 Nonparametric identification of markups via demand elasticities

2.1 Optimal markups

We start by introducing a model that is sufficiently flexible to allow us to estimate time-varying demand elasticities and markups in a market with many products and competing stores.

Let $Q_{si\tau}(P_{si\tau}, P_{-si\tau}, X_{si\tau})$ be a smooth downward-sloping store-level demand curve faced by store s for product i at time τ. It depends on the price charged by competitors $P_{-si\tau}$ and the demographic characteristics of consumers $X_{si\tau}$, which will depend on the store's location. Note that stores may charge different prices since they are in different geographic locations. Suppose that all stores follow a Bertrand price competition model. In the Bertrand-Nash equilibrium, competitor stores set equilibrium price $P^*_{-si\tau}$, which in general will be different across stores. We assume constant unit costs equal to $c_{si\tau}$. A given store s then sets its own price by maximizing profits[8] according to

$$\max_{P_{si\tau}} (P_{si\tau} - c_{si\tau}) Q_{si\tau}(P_{si\tau}, P^*_{-si\tau}, X_{si\tau}). \tag{1}$$

The first order condition for the optimal price $P^*_{si\tau}$ implies

$$(P^*_{si\tau} - c_{si\tau}) \frac{\partial Q_{si\tau}(P^*_{si\tau}, P^*_{-si\tau}, X_{si\tau})}{\partial P_{si\tau}} |_{P_{si\tau}=P^*_{si\tau}} + Q_{si\tau}(P^*_{si\tau}, P^*_{-si\tau}, X_{si\tau}) = 0. \tag{2}$$

Markups $\mu_{si\tau}$ are then set according to:

$$\mu^*_{si\tau} = \frac{P^*_{si\tau} - c_{si\tau}}{c_{si\tau}} = \frac{1}{e_{si\tau}(P^*_{si\tau}, P^*_{-si\tau}, X_{si\tau}) - 1}, \tag{3}$$

where $e_{si\tau}(P^*_{si\tau}, P^*_{-si\tau}, X_{si\tau}) = -\frac{\partial \log Q_{si\tau}(P_{si\tau}, P^*_{-si\tau}, X_{si\tau})}{\partial \log P_{si\tau}} |_{P_{si\tau}=P^*_{si\tau}}$ is the absolute elasticity of residual demand that store s faces at the profit maximizing price in the Bertrand-Nash equilibrium.

Assuming that, in the data, we always observe prices in the neighborhood of $P^*_{si\tau}$ for all stores s, one can compute the markups $\mu_{si\tau}$ using a nonparametric estimate of elasticity as the local average treatment effect (LATE) of the logarithm of the own price on the logarithm of residual demand at every store. Therefore, this LATE parameter is a sufficient statistic for the optimal markup charged by the store. A similar LATE approach was recently introduced by Munro et al. (2023) to study the impact of egg subsidies in rural India. Similar to us, they do not require a structural parametric model as long as the observed prices remain close to

[8] We could include fixed costs in this formula, but we omit them because they do not affect the maximization problem.

market equilibrium. However, unlike them, we work directly with differentiated products and Bertrand-Nash competition. This setup is also related to Gallego et al. (2019), who consider optimal price setting theory for a monopolist with a general smooth demand function that can be approximated with a linear demand function in the neighborhood of the optimal price.

If we estimate a log-log demand specification of a locally linear model, then the markups formula holds under the Bertrand-Nash competition of profit maximizing firms with a general and time-varying residual demand curve. Thus, this result can be seen as a nonparametric identification of markups using the first-order condition as proposed in Gandhi et al. (2020) for identification of the production functions. Our goal is to study heterogeneity in elasticities and corresponding markups. We next discuss how to estimate weighted averages of these parameters.

2.2 Actual markups

There are various reasons why optimal markups $\mu^*_{si\tau}$, which we will estimate, may deviate temporarily from actual markups $\mu_{si\tau}$ that stores set. For example, when the store sets prices it may face imperfect information about consumer preferences and choose a price setting strategy that simultaneously aims to maximize profits while also learning about demand (Wager and Xu, 2021). In a December 2025 article Consumer Reports states that "Instacart is quietly running pricing experiments on millions of shoppers" (Consumer Reports, 2025). These findings imply that prices change frequently (in our data, product prices change for 30% of the weeks within a store-year).[9] However, for goods that are sold less frequently, the store may also face frictions, imperfect information, or menu costs. As a result, there may be temporary differences between optimal and actual markups. Let deviations between actual and optimal markups in a week be equal to

$$\frac{\mu^*_{si\tau}}{\mu_{si\tau}} = \xi_{si\tau}. \tag{4}$$

We assume that $\xi_{si\tau}$ is a mean reverting process that has a long run average of one. Specifically, we assume that it is a weakly dependent stationary time series adapted to a natural filtration generated by $\{\xi_{si(\tau-1-\ell)}, \mu^*_{si(\tau-\ell)}\}_{\ell=0}^{\infty}$ with the property that $\mathbb{E}(\xi_{si\tau}|\mu^*_{si\tau}) = 1$. Note that we are assuming the ratio of optimal to actual markups to be a stationary mean-reverting process. The markups themselves can be non-stationary.

Intuitively, this assumption allows for the realistic possibility of persistence in deviation from optimal markups. It is satisfied, for example, if $\mu^*_{si\tau}/\mu_{si\tau}$ is an autocorrelated time series with log normal distribution. This assumption is consistent with, for example, Mankiw and

[9] This price variation can be driven, for example, by promotions imposed by wholesale suppliers (Anderson and Fox, 2019), as discussed above.

Reis (2002) and other models of price stickiness.

2.3 Harmonic Markup Index

From an estimation perspective, the presence of suboptimal prices introduces measurement error. That is, weekly optimal markup is a noisy measure of actual markups. Following Wald (1940), to mitigate this problem, we work with group averages of optimal markups as a proxy for actual markups. To reduce measurement error, we aggregate weekly markups to the annual level and propose a harmonic markup index. The harmonic mean is preferred to the simple average since we expect deviations $\xi_{si\tau}$ with higher optimal markups to have higher volatility.[10] The harmonic average automatically reduces the influence of such observations. Indeed, calculating the harmonic average based on (4) implies that

$$\Big(\frac{1}{\mathcal{T}}\sum_{\tau=1}^{\mathcal{T}}\mu_{si\tau}^{-1}\Big)^{-1} = \Big(\frac{1}{\mathcal{T}}\sum_{\tau=1}^{\mathcal{T}}(\mu_{si\tau}^{*})^{-1} + \frac{1}{\mathcal{T}}\sum_{\tau=1}^{\mathcal{T}}(\xi_{si\tau}-1)(\mu_{si\tau}^{*})^{-1}\Big)^{-1}. \tag{5}$$

Since $\xi_{si\tau}$ is a weakly dependent process, the second sum on the right hand side is approximately zero for sufficiently large $\mathcal{T}$ by the law of large numbers. At the annual level, we therefore expect the harmonic mean of actual markups to align with the harmonic mean of optimal markups.[11]

By choosing larger groups of goods for aggregation, we can further reduce the influence of suboptimal prices. We therefore aggregate up to the region-category level and compute a harmonic average across multiple stores in region r and multiple UPCs in a category j in year t to get a *harmonic markup index*,

$$\mu_{rjt}^{*} \triangleq \Big(\frac{1}{\mathcal{T}\sum_{i=1}^{N_j}\sum_{s=1}^{N_r} w_{si}^{rjt}}\sum_{\tau=1}^{\mathcal{T}}\sum_{i=1}^{N_j}\sum_{s=1}^{N_r}\frac{w_{si}^{rjt}}{\mu_{si\tau}^{*}}\Big)^{-1}, \tag{6}$$

where w_{si}^{rjt} is a positive weight for year t of UPC i in category j of store s in market r; $\mathcal{T}$ is the number of weeks in a year. For example, one can use weights proportional to the annual revenue of a given UPC for particular store. Revenue weighting can further take into

[10] If optimal markups $\mu_{si\tau}^{*}$ are low, then there is no incentive to deviate from the optimal markups. Food is perishable, so a store manager cannot afford to lose customers by withholding goods until consumer demand becomes less elastic again. On the contrary, if there is a large and persistent positive shock in demand that pushes optimal markups up, the store may take a more cautious approach and gradually adjust actual markups to match the new optimal level. This consideration implies that the conditional variance $\mathbb{E}((\xi_{si\tau}-1)^2|\mu_{si\tau}^{*})$ is an increasing function of $\mu_{si\tau}^{*}$. Similarly, we expect the conditional volatility of $\xi_{si\tau}$ to be higher if the product's revenue is lower. One reason could be limited attention. Store price setting is more likely to focus attention on reducing deviations from optimal prices for larger revenue products.

[11] For example, Anderson et al. (2017) find that on average 21% of products in a grocery store change prices every week. It is therefore reasonable to assume that the store evaluates product markups frequently and that optimal and actual markups are co-integrated (in logs) implying that equation (4) represents a reasonable description of their evolution and that averaging at the annual level will reduce measurement error.

account the fact that the (conditional) variance of the disturbance term $\xi_{si\tau}$ is likely smaller for products with larger revenue share.

We note that our harmonic markup index with revenue weights is analogous to the value-share weighted "Harmonic Laspeyres" price index (value-share is measured by the revenue share in our construction; see Coggeshall (1886) and Balk (2012)).

2.4 Panel-IV regression approach for demand elasticity estimation

In order to take our conceptual model to the data, we provide a mapping from the model variables to the variables in the data, which we discuss in more detail in the next section. We observe multiple UPC (barcode) varieties i of good category j, for example, different types of cereal. We assume that within a year the actual price $P_{si\tau}$ fluctuates in the neighborhood of the optimal price $P^*_{si\tau} = P^*_{si\tau}(P_{-si\tau}, X_{si\tau})$. Using a first order Taylor expansion of the demand function (Cauchy's mean value theorem),

$$\log Q_{si\tau}(P_{si\tau}, P_{-si\tau}, X_{si\tau}) = -e_{si\tau}(\log P_{si\tau} - \log P^*_{si\tau}) + \log Q_{si\tau}(P^*_{si\tau}, P_{-si\tau}, X_{si\tau}), \quad (7)$$

where $e_{si\tau}$ is the absolute elasticity of demand at a point between $P_{si\tau}$ and $P^*_{si\tau}$. For simplicity we assume that the elasticity $e_{si\tau}$ is approximately constant for a specific store, UPC, and week. The remaining terms in equation (7) can be captured through fixed effects and an error term in a saturated panel representation,

$$\log Q_{si\tau} = -e_{si\tau} \log P_{si\tau} + upc_{rit} + store_{sjt} + week_{rj\tau} + \varepsilon_{si\tau}. \quad (8)$$

The saturated regression model (see, for example, Abadie et al., 2020; Bugni et al., 2023) controls for the majority of demand factors $X_{si\tau}$ and the prices of substitutes using group fixed effects. The idea is that a sufficiently large set of fixed effects controls for market conditions and consumer income nonparametrically, which is more flexible and potentially more robust than including those controls directly and assuming a specific functional form. Three sets of fixed effects $\{upc_{si\tau}, store_{si\tau}, week_{si\tau}\}$ capture various demand effects associated with particular UPC, store, and week.

The most granular empirical approach would be to calculate store-level elasticities for each product variety i in a given week τ in equation (8). This is the level at which we observe the data, and so it is trivially impossible. We therefore aggregate the data across varieties and stores and calculate an average elasticity for a specific group of observations. Naturally, aggregation and higher power reduce resolution. This means that we need to choose a sample over which we are averaging and then choose weights in order to calculate the within-sample average. We next discuss grouping size and the associated estimating equation; after that we turn to the weighting scheme.

Since we are interested in cyclical variation at the national level, we choose to estimate average demand elasticities at the market-good category-year level. That is, for example, we estimate the average elasticity of cereal in Sacramento in 2020.[12] We therefore estimate:

$$\log Q_{si\tau} = -e_{rjt} \log P_{si\tau} + upc_{rit} + store_{sjt} + week_{rj\tau} + \varepsilon_{si\tau}. \tag{9}$$

On the LHS, we see that the resolution of our data is at the store-UPC-week level with subscript $si\tau$. On the RHS, we note that the elasticity e_{rjt} varies across regions r, good categories j and years t, which is our chosen level of aggregation. We assume that demand depends on (i) the UPC, i.e. the specific good, and that this effect varies across regions r and over time t, (ii) the store, an effect that varies across good categories j and time t, and (iii) the week τ, an effect that varies across regions r and good categories j. Our fixed effects therefore capture differences in demand due to quality (UPCs), location (stores), and seasonal variation (time). Our approach strikes a balance between being sufficiently specific to capture important differences between goods, across regions, and, importantly, over time, while delivering enough power to estimate elasticities with a relatively high degree of accuracy.

The remaining variation in demand shocks not captured by the fixed effects is included in the residual term $\varepsilon_{si\tau}$. However, shocks to demand may be correlated with price changes since they are determined together with product quantities in market equilibrium (Working, 1927). This correlation may therefore cause endogeneity bias of elasticities towards zero. The standard solution to the endogeneity problem when estimating demand elasticities is to use an instrumental variable that moves together with the supply shock, but is orthogonal to any remaining demand shocks in equation (9). We describe our IV approach in Section 2.5.

Revenue weighting. We now discuss the weighting scheme when estimating the weighted average elasticity e_{rjt} for a specific market, food category, and year (r, j, t) in equation (9). Our goal is to estimate markups of products that a typical consumer buys in a typical store in a typical week. In Section 2.3 below equation (6) we discussed the role revenue weighting plays in improving alignment between actual and optimal harmonic markups in the presence of suboptimal price adjustments.

We use the annual UPC-store-market revenue share to estimate LATE elasticity e_{rjt}.[13]

[12] Alternatively, one may further subdivide good categories into smaller groups like organic and conventional goods (Gafarov et al., 2024). Other options are to divide markets r into store chains (DellaVigna and Gentzkow, 2019), divide big categories into subcategories by brand or food labels, or work with a quarterly (or monthly) resolution instead of the annual one we choose. There is a recent literature concerned with automatic selection of groupings for time-varying fixed effects, for example, Bonhomme and Manresa (2015); Bonhomme et al. (2022); Chetverikov and Manresa (2022). These procedures are computationally prohibitive for large datasets such as ours. Such automatic procedures could be more promising when one is estimating demand for one individual product category as opposed to the entire national market, which is the focus of our paper.

[13] See the third paragraph in Appendix B for further details on the estimation strategy.

These revenue weights represent a typical food basket purchased by consumers. We apply revenue weights to both stages of each IV regression. Similar to Theorem 1 in Angrist et al. (2000), the revenue-weighted LATE parameter takes the form

$$e_{rjt} = \frac{1}{\mathcal{T}\sum_{i=1}^{N_j}\sigma_{rit}^2\sum_{s=1}^{N_r}\omega_{si}^{rjt}}\sum_{i=1}^{N_j}\sigma_{rit}^2\sum_{s=1}^{N_r}\omega_{si}^{rjt}\sum_{\tau=1}^{\mathcal{T}}e_{si\tau}, \tag{10}$$

where ω_{si}^{rjt} is year t revenue of UPC i in category j of store s in market r; σ_{rit}^2 is the UPC-level cost shock variance component of IV for market r in year t conditional on the demand shocks controls; $\mathcal{T}$ is the number of weeks in a year. We can use the LATE parameter (10) to compute the harmonic markup index for market r, category j, and year t with weights $w_{si}^{rjt} = \omega_{si}^{rjt}\sigma_{rit}^2$ as

$$\mu_{rjt} = \frac{1}{e_{rjt} - 1}. \tag{11}$$

That is, the harmonic markup index has an interpretation of optimal markups for a monopolist facing demand with a price elasticity equal to the LATE parameter e_{rjt} (Cournot, 1838). Lerner (1934) suggested using this formula as a measure of market power; hence, this index is widely referred to as *Lerner's index*. Lerner's index has been used extensively in recent work (for example, DellaVigna and Gentzkow, 2019, who assume time-invariant elasticity of demand). We provide an interpretation of Lerner's index in the case of heterogeneous and time-varying demand elasticity using LATE framework.

It is useful to think about implications and potential advantages of revenue-weighting the data rather than using an equally-weighted approach. There are two important differences relative to equally-weighting. First, sale quantities tend to be concentrated in a small share of UPCs. This means that an equally-weighted approach will effectively overweight UPCs that have low sales. Some customers may have inelastic demand for highly specialized products, e.g. specific organic products or fair trade labels (Gafarov, Hasbany, and Hilscher, 2025). Thus, an equally-weighted estimation approach results in lower average elasticities than the revenue-weighted alternative.[14] Second, due to limited attention of a store, the UPCs that bring the most revenues are more likely to be priced optimally and thus reflect our theoretical model of profit maximization. Prices of such products are also more likely to reflect cost shocks. Therefore, one can think of revenue weighting as a scheme to identify the effects for "compliers" (Abadie, 2003), i.e. UPCs with larger cost variance component σ_{rit}^2 common with the IV.

[14]We discuss results from such an approach in the Appendix. Estimates are reported in Table B.5. The revenue-weighted estimates are reported in Table 1.

2.5 Hausman's IV for cost shocks

We follow Hausman (1996) and Nevo (2001) to address the endogeneity problem inherent in demand estimation. We instrument the log of a store-specific weekly product price in a market of interest by the quantity-weighted average of log weekly prices of the same product sold at all stores in the paired market(s), i.e. markets that are geographically close.

Most products sold in neighboring markets (but not distant markets) are the same. This enables us to instrument product prices in a market of interest by those in its neighboring market(s). We select 12 pairs of 25 large markets that are from different US regions, thus covering the East, South, West and Midwest of the country (Figure 1). By focusing on geographically close markets we can better capture common local cost shocks such as labor and transportation costs.[15]

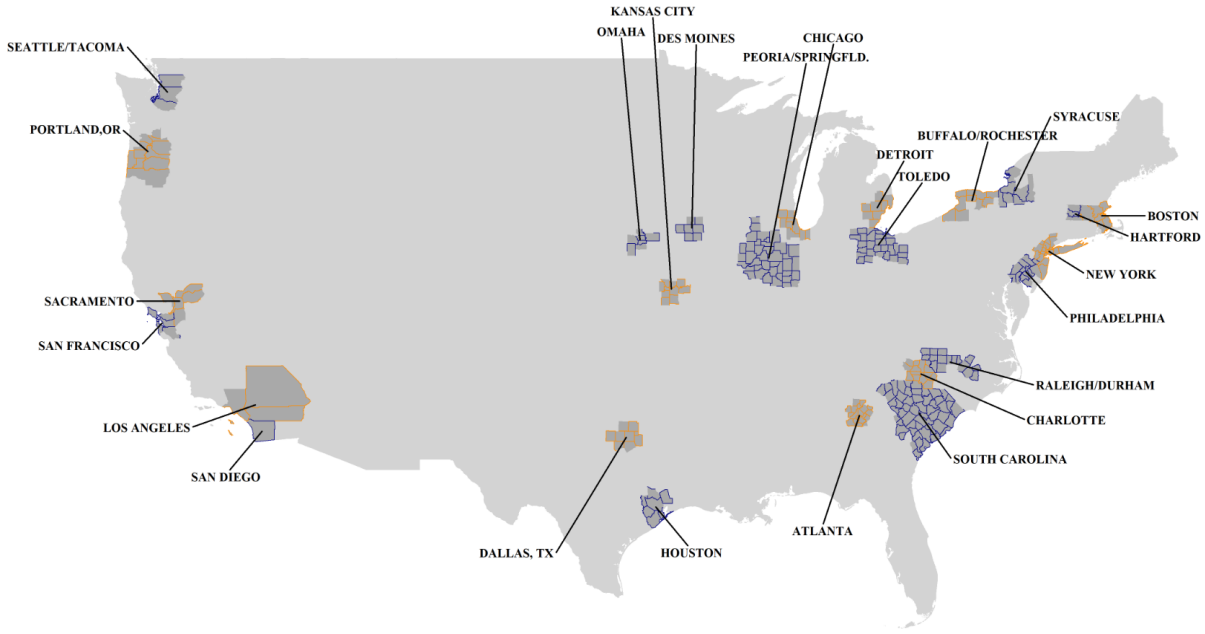


Figure 1: Paired Markets for IV estimation.

Note: Paired markets have county borders colored in blue and orange. See specific pairs in Appendix Table B.1.

Note that product UPCs are manufacturer-specific, i.e. a product with the same UPC sold in the paired market comes from the same manufacturer, which delivers the relevance of our price instrument. For example, suppose that a wholesale supplier introduces promotional coupons that temporarily lower the price of a particular UPC (e.g. a specific type of cereal) in a given week. These coupons would be passed to all the sellers of this product and constitute a positive supply shock. The coupon then will lead to a price decline for that UPC relative

[15]Our implementation of the Hausman IV strategy is different from DellaVigna and Gentzkow (2019) who used the average price for a given UPC across the entire nationwide store chain.

to the whole product category as required for the estimation of the own-price elasticity of demand. In contrast, if producer input costs change, it will impact the whole good category and will be captured by the weekly fixed effect together with category-level demand shocks.

We briefly discuss two assumptions that support our exclusion restriction. First, there may be idiosyncratic UPC-specific demand shocks in geographically close areas resulting in changed prices that are not driven by supply shocks. For example, in a specific location consumers may stop purchasing a specific type of lettuce because of a reported e-coli outbreak. Stores may reduce prices for that UPC in response to such a demand shock. We assume that such common UPC-specific demand shocks have a negligible impact on our estimates. Since we estimate an average price elasticity of hundreds of UPCs in a market, e_{rjt}, such an occurrence would not materially affect the final estimated elasticity unless it affects the majority of the UPCs in the category.

Second, there may be some commonality across stores in cost shocks, resulting in prices moving in more than one store at a time in response to the instrument. If consumers could perfectly observe all prices in competing stores, their reaction to a price change in a given store may be reduced. However, we make the reasonable assumption that most consumers shop at the weekly frequency of our data and therefore become aware of price co-movement only with a lag. In the context of our model, we assume that costly search results in consumers observing prices in the store they are shopping in at time τ and know the average prices from other stores from last week, $\tau - 1$ (Burdett and Judd, 1983). Thus we assume that sales at store $Q_{si\tau}$ depend on $P_{si\tau}$ and $P_{-s,i,\tau-1}$. In other words, demand at a specific store reacts primarily to changes in the average regional price with a lag. We can therefore capture such competitor-driven demand changes, $P_{-s,i,\tau-1}$, together with the consumer-side demand shocks, $X_{si\tau}$ using weekly fixed effects. Therefore, contemporaneous cross-store price movements do not violate the exclusion restriction.

Consistent with our assumptions and as discussed in Section 2.7 below, we find that this Hausman-type price IV is strong for most of our demand elasticity estimates and substantially reduces the downward simultaneity bias.

2.6 Retail scanner data

We estimate demand elasticities using data on prices and quantities for various food categories, markets, and years. We use retail scanner data from the IRI Marketing Dataset (Bronnenberg et al., 2008) and the NielsenIQ Retail Scanner Dataset.[16] The IRI scanner data covers 12 years, from 2001 to 2012 while the NielsenIQ scanner data span 19 years, from 2006 to 2024. Both have weekly transaction records of products sold by physical retail

[16]We obtain the NielsenIQ Retail Scanner Dataset from the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

stores across markets. We focus on IRI and NielsenIQ panel stores with continuous data coverage throughout the two sample periods, ensuring that our results are not driven by changes in store coverage. In the following, we describe the key features of the data that are relevant to our empirical estimation strategy.

Markets. As noted above, we choose 12 pairs from 25 markets (Figure 1).[17] There are 50 IRI markets in total. The other 25 markets, which are not used in this paper, are mostly small or do not have clear neighboring markets for pairing. An IRI market consists of one or multiple adjacent counties; we match NielsenIQ grocery store locations to find the IRI markets to which they belong. We note that the selection of products sold in neighboring markets is almost the same each year. In total, we successfully match 90.7% and 91.2% of price observations with their IVs in the IRI and NielsenIQ samples, respectively.

Markets vary significantly in size. The largest markets are New York and Los Angeles. In total, all markets had 123.8 million residents in 2010, more than 40% of the 2010 US population. Taken together, panel stores in the 25 markets represent 33% of the observations in the IRI scanner data, and contain a large but smaller proportion of observations in the NielsenIQ scanner data. Table B.2 shows a list of all markets, their populations, number of observations and number of stores. Overall, these markets are broadly representative. Table B.4 reports summary statistics for observations per store, including the number of categories and UPCs.

Food categories. We group UPC products into 16 food categories defined in the IRI sample. In the NielsenIQ sample, we estimate elasticities for 58 food categories, defined by using the NielsenIQ food groups. We include the complete list of food categories in Table B.6 and discuss additional implementation details in Appendix A. We estimate demand elasticities at the food category level. While sets of UPCs sold are not the same across all markets, they are similar for neighboring markets (see Table B.3).

2.7 Estimates of market-category-year elasticities

We estimate absolute demand elasticities at the market-category-year level—25 markets, 12 (IRI) and 19 (NielsenIQ) years, and 16 (IRI) and 58 (NielsenIQ) food categories.

When implementing our estimation strategy, we find that the proposed Hausman-type IV works for most weekly product prices at the UPC level across markets and years. We find that the price instrument is relevant (strong) for the vast majority of the market-category-year combinations. Following the rule of thumb of (Staiger and Stock, 1997), we drop estimates with an F-stat below 10, which represents about 7.9% of the data, with a relatively higher

[17] We combine Omaha and Des Moines and pair them with Kansas City. Table B.1 outlines market pairs.

percentage within the IRI sample possibly due to its smaller number of observations for each elasticity estimation.

Among the strong-IV demand elasticity estimates, less than 1% have an incorrect sign, most likely due to estimation errors in a typical IV regression. We drop these estimates. Our final sample has 29,353 demand elasticity estimates with 13.1% in the IRI sample and 86.9% in the NielsenIQ sample.

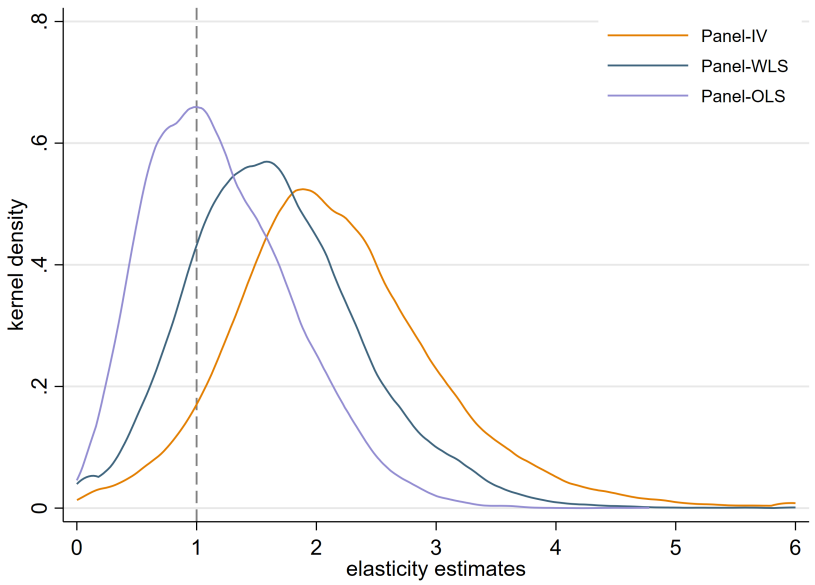


Figure 2: Smoothed density estimates of the distributions of OLS, WLS, and IV demand elasticity estimates.

In order to assess the validity of the price instrument, we compare OLS (equally-weighted), WLS and IV (revenue-weighted) estimates. Figure 2 reports the estimated kernel densities for the three elasticity distributions. There is a strong shift in the distribution from OLS to WLS, suggesting that there are a lot of UPCs with fairly low market shares that appear to not be representative of what people purchase. One reason could be that more differentiated, niche products with low market shares are less elastic. The strong shift from WLS to IV indicates validity and underscores the need for our IV estimation strategy (Rossi, 2014). One way to get a sense of the magnitude is to compare medians; the OLS median is 1.12, the WLS median is 1.60 while it is 2.08 for the IV estimates. The fraction of estimates below one is equal to 41.9% for OLS, 17.9% for WLS, and 6.9% for IV.[18] We conclude that our instrument corrects for a significant part of the downward endogeneity bias, while the revenue weighting adjusts for the composition of consumers' shopping patterns.

[18]We note that equation (3) implies that if we observe an estimated (absolute) demand elasticity below one for the specific food category and time period, the observed price will not have been set based solely on profit maximization.

Another way to show the validity of our approach is to compare our elasticity estimates with estimates from the literature. We note that the comparison is not exact since we are estimating a revenue-weighted average elasticity. Cereal is commonly studied in the literature and therefore an appropriate category for comparison. For the IRI sample, our mean elasticity estimate is about 2.9 for the Los Angeles market during 2007-2010. For the same market and time period, Richards and Hamilton (2015) obtain an elasticity mean of 2.2 based on a BLP model, which is comparable with our mean elasticity estimate. For the NielsenIQ sample, our median elasticity estimate is about 2.5 for all 25 markets during 2007-2017. For the same time period, Backus et al. (2021) have a median elasticity estimate of 2.7, although they include fewer diverse markets in their sample of interest. In sum, our cereal elasticity estimates are comparable to those obtained using BLP for both the IRI and NielsenIQ datasets.

Table 1: Summary Statistics of Demand Elasticity Estimates.

	IRI sample[a]			NielsenIQ sample[b]		
	mean	p10	p90	mean	p10	p90
point estimate	2.92	1.86	3.98	2.06[c]	1.11	3.06
standard error	0.24	0.10	0.42	0.17	0.06	0.33
t statistics	16.13	7.04	26.92	18.35	5.32	33.67
obs in an estimation (thousands)	120.0	18.7	259.7	1129.2	58.4	2845.0
Cragg-Donald F statistics	115.2	15.0	228.4	1740.2	33.3	2874.4

Note: [a]The IRI sample has 3,833 elasticity estimates. [b]The NielsenIQ sample has 25,520 elasticity estimates. These samples include elasticity estimates that are obtained under the strong price IV and exclude those that have negative values. See detailed descriptions in the main text. [c]The NielsenIQ sample has lower elasticity estimates than the IRI sample. The downward trend in demand elasticities over time, as explained in Section 3, largely contributes to the lower elasticity estimates in the NielsenIQ sample of more recent years. The NielsenIQ sample also has a small proportion of price-inelastic goods such as baby food, spices, seasoning, extracts and eggs.

Table 1 reports summary statistics for the final sample. We report statistics separately for the IRI and NielsenIQ samples. We note that the average demand elasticity estimate in the IRI sample is higher than that in the NielsenIQ sample. This pattern is mainly due to the downward trend in demand elasticities, which is one of the main findings of our paper and which we start to discuss in more detail in Section 3.

We compute standard errors clustered at the store and week levels in two ways to allow for arbitrary correlations of the error term $\varepsilon_{si\tau}$ across products within each store and across stores within each week. The standard errors are small, implying that the elasticities are estimated precisely. Close to 99% of the t statistics are above 1.96. Indeed, 90% of the t-statistics are above 7.04 (IRI) and 5.32 (NielsenIQ). This is because we have sufficient observations and strong price IVs in nearly every elasticity regression (see Table 1).

Our demand elasticity estimates exhibit notable heterogeneity across markets and goods, which is in line with the empirical literature (DellaVigna and Gentzkow, 2019). For example,

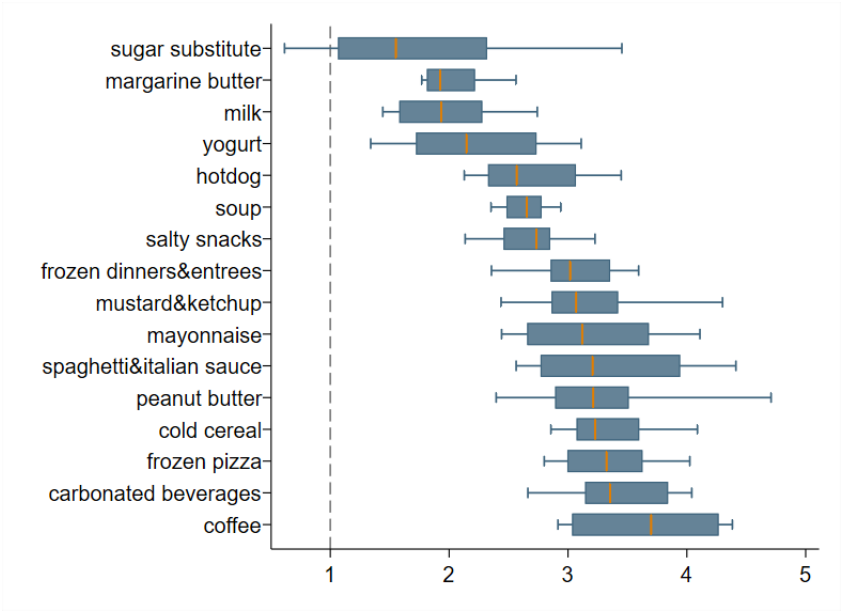


Figure 3: IRI Demand Elasticity Estimates by Food Categories in 2010.

Note: This boxplot shows the elasticity distribution across 25 markets for the 16 food categories in the 2010 IRI sample. We order food categories by their median elasticity estimates colored in dark orange. Each boxplot consists of two parts—one box and two whiskers. A box ranges from the lower quartile to the upper quartile of elasticity estimates with the median in the middle. The two whisker ends represent the 10th and 90th percentiles of the elasticity distribution, respectively.

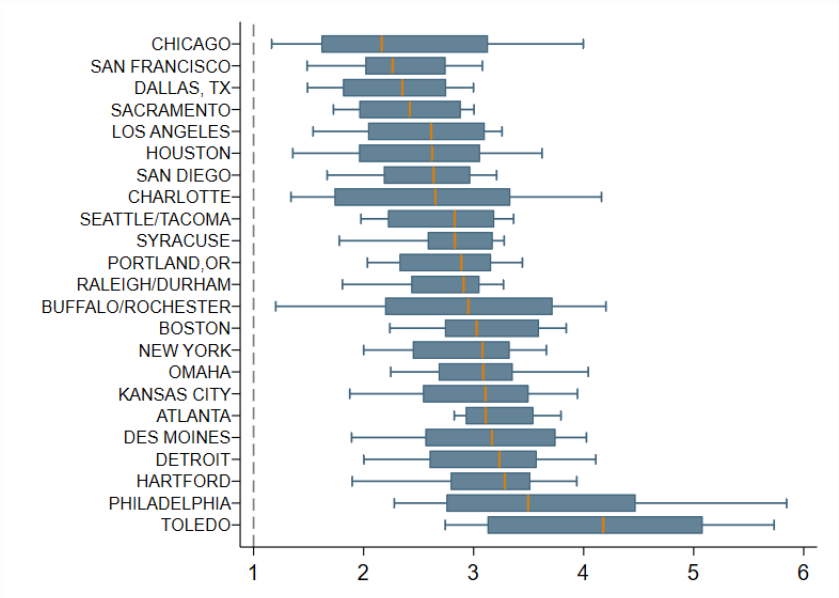


Figure 4: IRI Demand Elasticity Estimates by Markets in 2010.

Note: This boxplot shows the elasticity distribution across 16 food categories for 23 markets in the 2010 IRI sample. We order markets by their median elasticity estimates colored in dark orange. South Carolina and Peoria/Springfield markets have limited elasticity estimates in this particular sample so that we do not show their elasticity distributions here.

Figures 3 and 4 show that IRI demand elasticity estimates generally range from one to five across food categories and markets in 2010. Within each food category (market), these elasticity estimates also have sizable differences across markets (food categories). NielsenIQ demand elasticity estimates have similar features, as shown in Figures B.1, B.2, and B.3 in the Appendix. We will control for these notable level differences among market-category-year specific demand elasticity estimates through market- and category-specific fixed effects in our factor analyses for elasticity and markup dynamics in Section 4.

3 Unconditional cyclical and trend variation

Having estimated elasticities that vary across good categories, markets and over time we are now in a position to analyze time variation. In this section, we consider unconditional cyclicality, i.e. the extent to which elasticities vary with the business cycle. In the next section we then decompose this variation and attribute it to various drivers; at that point we are able to measure conditional cyclicality.

We consider time variation by estimating the economy-wide average elasticity; we then analyze trend and cyclical components. We begin by summarizing changes in the elasticity estimates over time. We pool the IRI and NielsenIQ market-category-year estimates.[19] Each year, we then calculate the median, 25th and 75th percentiles of the full distribution of estimated market-category elasticities. Figure 5 reports these three time series. There are three patterns that we immediately notice. First, there is a pronounced downward trend in demand elasticities over the sample period. Median elasticities decline from 2.52 (peak in 2002) to 1.44 (in 2022).

Second, at times of contraction in aggregate demand, elasticities increase, implying counter-cyclical elasticities.[20] For ease of identifying important aggregate demand shocks, specifically recessions in 2001, 2008-2009, the monetary contractions of 2017 and 2022, and the COVID-19 pandemic and associated recession in 2019-2020, we add vertical lines to Figure 5 to denote the first year of these events. We note that the Tax Cuts and Jobs Act (TCJA) in 2017 and expansionary fiscal and monetary policy during the COVID pandemic imply that we need to control for various factors driving variation in elasticities and markups. We perform such analysis in Section 4 when we study conditional cyclicality.

We check that the unconditional cyclical and trend variation is present both across good

[19] Due to differences in store attributes and good coverage, the estimates of elasticities across the two data sets are not directly comparable. Indeed, we find that the average elasticity estimates during the overlapping part of the sample is not exactly the same. The IRI data set is much smaller in scope (fewer goods, stores, and years); we therefore shift IRI elasticity estimates by a constant to match that estimated in the NielsenIQ dataset. One reason for the difference may be the fact that a NielsenIQ food category has relatively more substitutable products at the lower UPC level, as shown in Table B.3.

[20] We note that the increase in elasticities in 2008-2009 is present both in the IRI and NielsenIQ data.

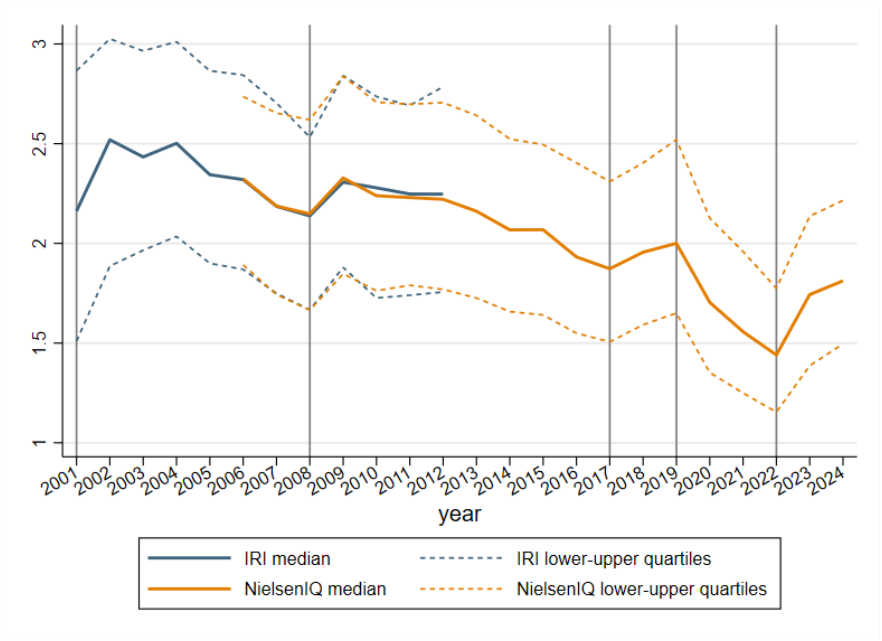


Figure 5: Demand Elasticity Estimates over Time.

Note: For each year, we report median, upper, and lower quartiles of the pooled market-category distribution of elasticities. The five vertical gray lines refer to important aggregate demand shocks—the recessions in 2001, 2008-2009, the monetary contraction in 2017 and 2022, and the COVID-19 pandemic and associated recession in 2019-2020. As explained in the main text, due to differences in store attributes and good coverage between the IRI and NielsenIQ samples, we shift IRI demand elasticity estimates down by a constant so that they have the same median elasticity estimate as NielsenIQ's during the overlapping period of 2006-2012.

categories and markets (see the examples shown by Figures B.5 and B.6). The pattern is thus not driven by potential time-varying changes in the composition of the sample. This motivates us to pool the IRI and NielsenIQ elasticity estimates into one sample for our main analysis below.

Third, we note that the dynamic pattern of elasticities is present not only in the median elasticity time series but is also reflected in the quartiles of the entire distribution.[21]

In addition to a visual inspection of the estimates in Figure 5 we decompose these changes into a secular trend component and a higher frequency cyclical component. The former part captures long-run changes in consumers' preferences and shopping behavior as well as market structure. The latter part captures the fluctuations of consumers' purchasing power over the business cycle and in response to demand shocks.

We run WLS of the market-category-year demand elasticity estimates in the pooled sample on year dummies and a linear year trend, using within-year market-category revenue shares as weights. We use market-category-specific fixed effects cm_{jr} to control for level differences in demand elasticity estimates across markets r, categories j, and samples. We pool

[21] Recently De Loecker et al. (2020) argue that the upward trend in markups across US industries is driven, to a large extent, by the upper tails of the markups distribution. The results in Figure 5 suggest declining elasticities and therefore increasing markups across all quantiles of the distribution.

market-category-year-specific elasticity in the IRI and NielsenIQ samples and estimate:

$$\hat{e}_{rjt} = cm_{jr} + \alpha T + \sum_{t=2001}^{2024} \theta_t D_t + \epsilon_{rjt}. \tag{12}$$

For normalization purposes, we drop year dummies for 2005 and 2016, the two reference years with $\theta_{2005} = \theta_{2016} = 0$, such that the sum of the cyclical variation around the linear trend is statistically indistinguishable from zero.

Table 2: Trend and Cyclical Variation in Elasticity and Implied Markup.

	elasticity[a]	gross markup, $\ln(1+\mu)$[b]
Linear trend		
average annual change, 2001-2024	-0.030***	0.014***
	(0.006)	(0.003)
Aggregate demand contractions		
from 2001 to 2002	0.380***	-0.204***
	(0.072)	(0.039)
from 2008 to 2009	0.119***	-0.042***
	(0.024)	(0.009)
from 2017 to 2018	0.087***	-0.047***
	(0.018)	(0.010)
from 2022 to 2023	0.266***	-0.297***
	(0.026)	(0.029)
COVID period annual changes (detrended)		
from 2019 to 2020	-0.268***	0.120***
	(0.050)	(0.022)
from 2020 to 2021	-0.107**	0.072**
	(0.053)	(0.036)
from 2021 to 2022	-0.079***	0.069***
	(0.020)	(0.018)

Note: [a]This table reports the estimates of linear trend and cyclical components as well as COVID-specific yearly changes in elasticities and markups during our sample period. [b]Using the chain rule, we can calculate annual gross markup, $\ln(1+\mu)$, effects as elasticity effect times $-1/(e_t(e_t-1))$. We revenue-weight annual average elasticity estimates e_t when computing markups effects. Standard errors, clustered at the market, food category, and year levels (three-way clustering), are listed in the parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2 presents estimates for annual changes in detrended average elasticities. For example, elasticities increased by 0.38 from 2001 to 2002, during the recession associated with the dot-com bubble bursting, 0.12 after the 2008 financial crisis, and 0.27 in response to the 2022-23 monetary tightening. Given these economically large changes, we next discuss their implications for markups.

Table 2 reports the results from the trend-cycle decomposition of changes in annual gross markups based on the trend-cycle estimation for annual elasticities. We convert annual

effects of elasticities into gross markups using the chain rule and measure annual changes in $ln(1+\mu)$ where $1+\mu$ represents gross markup (price divided by marginal cost). The average trend increase in gross markups was equal to 1.4% per year from 2001 to 2024. The impacts of negative aggregate demand shocks on markups vary between -29.7% (for 2022-2023) and -4.2% (for 2008-2009).

Our trend and cycle decomposition also allows us to measure the unconditional detrended effect on elasticities of the COVID years. There was a drastic reduction in elasticity (and a corresponding jump in markup) in the first year of the COVID-19 pandemic, 2020. Gross markups increased by 12.0% relative to the long-run trend. Over the next two years, there were slightly smaller increases of 7.2% and 6.9% in 2020/21 and 2021/22 respectively (Table 2). These increases were likely a short-run phenomenon driven by strong fiscal and monetary stimuli. People's concerns over getting COVID during multi-stop shopping may have made them less price sensitive and thus raised stores' markups. Indeed, the subsequent reduction in markups in response to the monetary tightening of 2022, equal to 29.7%, was of similar magnitude to the combined change over the COVID years.

4 Estimation of conditional markup dynamics

4.1 Cross-sectional variation in conditional cyclical patterns

Results in the previous section show that elasticities and markups exhibit substantial secular and unconditional cyclical variation. In particular, the evidence suggests that average markups react to aggregate demand shocks. It is possible that markups are affected by various channels, which we explore in this section. Specifically, we consider demographic, macro (income and wealth effects), competitive landscape, and real interest rate. We therefore present evidence of conditional dynamics to complement the evidence on unconditional dynamics presented in the previous section.

Since our data is at the market-category-year level, it reflects regional variation. Different markets may follow heterogeneous patterns, and so our panel data set allows us to estimate the contributions of various underlying channels. In turn, we can measure cyclicality conditional on changes in various factors.

There have been several channels discussed in the literature that transmit the business cycles into price elasticities and markups. In particular, there is evidence that wealth affects price elasticities (Stroebel and Vavra, 2019). Elasticities and therefore markups may also vary along with income changes during business cycles (Aguiar et al., 2013). Alternatively, markups may grow as a result of increased market concentration (Philippon, 2019) and market structure may evolve together with the business cycles (Rotemberg and Saloner, 1986).

The panel structure of our dataset gives us the opportunity to study underlying channels

of markup dynamics. We next investigate cross-sectional variation in aggregate demand measures and indicators of market concentration.

4.2 Partial adjustment model

We next estimate a dynamic panel model. It is well known that demand functions may gradually adjust to external factors due to, for example, habit formation. We therefore use the partial adjustment dynamic panel model of Balestra and Nerlove (1966) and Houthakker et al. (1974) to estimate the short and long-run effects of various measures on weighted average elasticities and therefore harmonic average markups. Harmonic markups (6) are a function of revenue-weighted demand elasticities. We therefore follow a two-step procedure to estimate the effects of explanatory variables on markups. We firstly estimate a model to explain variation in elasticities across markets and over time, and then convert these effects to changes in markups.[22]

We pool market-category-year-specific elasticities in the IRI and NielsenIQ samples and run a dynamic panel:

$$\hat{e}_{rjt} = \rho\hat{e}_{rj(t-1)} + \sum_{\ell=0}^{L} \left(\alpha_\ell R_{t-\ell} + M_{r(t-\ell)}\beta_\ell + C_{rj(t-\ell)}\gamma_\ell + D_t^{20-23}V_{r(t-\ell)}\theta_\ell\right) + c_j + m_r + D_t^{20-23}m_r + \epsilon_{rjt}. \tag{13}$$

We include several macroeconomic, demographic, and grocery market factors to explore channels that affect demand elasticity and markup dynamics: R_t is the economy-wide real interest rate, M_{rt} are market-specific factors (real GDP per capita, real housing prices, retailer HHI, population), C_{rjt} are food-category-specific variable (the average number of UPCs per category across stores within each market). We also add category and market fixed effects, c_j and m_r.[23] Our sample period includes the COVID years. During this period, real GDP per capita may not have been a good measure of purchasing power for groceries. Due to changes in shopping behavior and the increase of food-at-home consumption, the grocery sector was exposed to a large positive and persistent demand shock. To isolate the effect of this sectoral shock, we introduce a dummy variable $D_t^{20-23} = 1\{2020 \leq t \leq 2023\}$ and its interactions with real GDP per capita, $V_{r(t-\ell)}$.

[22] We recall that for some market-category-year observations, the persistence of deviations between actual and optimal markups, ξ, may be sufficiently large so that actual markups are measured with some noise even at the annual level. When estimated annual average elasticities are close to one, it is likely that the corresponding actual markup is different from the optimal markup. As discussed in Section 2.3, the harmonic average markup dampens the effects of such outliers and presents a more stable estimate of actual (harmonic) markups. Since elasticities are inverse markups (11), we therefore estimate the effects of explanatory variables on markups using corresponding elasticities. We note that, as reported in Table B.9, our main results are robust to converting elasticities to markups and then estimating the dynamic panel model.

[23] IRI and NielsenIQ have different product category names, which are captured by our category fixed effects. Also, Table B.7 shows that our regression results remain almost unchanged when a linear-time trend is included for robustness.

It is important to mention that both average elasticity (and therefore harmonic markups) as well as the explanatory factors are persistent (see Appendix Figures C.1 and C.2). This may cause inconsistent estimators due to the spurious correlation effect in a static model. Assuming that the persistent variables have at most one unit root, we can restore consistency by including both lagged dependent and explanatory variables into the model (Sims et al., 1990). We set $L = 1$ so that the lagged dependent and explanatory variables are included.

As pointed out earlier, the partial adjustment model allows us to estimate both the short- and long-run effects of different variables. The short-run effects are equal to $(\alpha_0, \beta_0, \gamma_0)$ and capture contemporaneous effects of exogenous factors. The long-run effects are equal to $(\sum_{\ell=0}^{L} \alpha_\ell/(1-\rho), \sum_{\ell=0}^{L} \beta_\ell/(1-\rho), \sum_{\ell=0}^{L} \gamma_\ell/(1-\rho))$. They capture the predicted changes in the dependent variable in response to permanent shifts in the explanatory variables.

Another feature of this setup is that we can perform a historical decomposition in which we analyze the effect of variation in different explanatory variables on the dependent variable. For example, what was the effect over time of the real interest rate on markup dynamics? We will perform this exercise after discussing the explanatory variables.

4.3 Data on explanatory variables

Before proceeding, we briefly discuss the explanatory variables. We consider four types of changes in local grocery demand: (i) slow-moving demographic factors, (ii) changes in the competitive landscape of the local market, (iii) cyclical demand shocks driven by business cycle variation, and (iv) effect of monetary policy. All of these factors may affect local demand, and, as such, have an effect on grocery store price elasticities. The data is gathered from publicly available sources. Appendix C lists data sources and describes the construction of some explanatory variables. Here, we provide a brief discussion of their potential impacts on demand elasticities.

(i) Population: Local population growth may reflect consumer's outlook on future economic development in each market. A better economic outlook may give households more confidence in their future income growth and make them less price sensitive. Population growth may also boost the economy and thus increase income in the near future, again potentially resulting in lower price sensitivity. We note that the population data show that, after the COVID-19 pandemic, some states, like California, saw a decline in population, while others, like Texas, saw an increase.

(ii.a) Average No. of UPCs per product category across stores: We calculate this measure at the market-category-year level based on the IRI and NielsenIQ scanner datasets. For each category and year, we take a weighted average of UPC numbers across stores within each market using store-category-specific revenue shares as weights. Intuitively, more UPCs in a product category mean more product varieties for consumers to choose from, which tends

to make consumers more price sensitive. However, new products may be better tailored to consumers with heterogeneous tastes, which tends to make consumers less sensitive to price changes. In theory, these two effects may offset each other (Neiman and Vavra, 2023;Grieco et al., 2024). We will use the data to analyze which effect dominates.

(ii.b) Herfindahl-Hirschman index: HHI is a standard way to measure market concentration (e.g., Neiman and Vavra, 2023; Grieco et al., 2024). Following the literature, we construct a retailer-HHI for each market during each year using NielsenIQ's consumer panel dataset.[24] For the years of 2001-2003 that NielsenIQ's consumer panel dataset does not cover, we use the 2004 retailer HHI for each market. In principle, a higher HHI means a higher market concentration so that consumers have more restricted retailer choices, resulting in lower price elasticities of demand.

(iii.a) Real GDP per capita: We use this as a measure of real income. Increases in household income may result in lower price sensitivity.

(iii.b) Cumulative percentage changes in real housing prices: For homeowners, rising housing prices can mean higher wealth and therefore make them less price sensitive (Stroebel and Vavra, 2019). However, Stroebel and Vavra (2019) also find that, for renters, higher housing prices can mean higher rent and therefore make them more sensitive to grocery prices. Which effect dominates is an empirical question.

(iv) Real interest rate: We use the one-year real interest rate from FRED.[25] The interest rate is the key determinant of current consumption spending decisions in modern macroeconomic models (Woodford, 2003). Higher interest rates incentivize more savings at the expense of current consumption. Moreover, in the short run, the Fed can adjust the real interest rate using standard monetary policy tools.

In Figures C.1 and C.2 in the Appendix, we show the variations of all these factors. Importantly, with the exception of the real interest rate, all other factors have both market-specific and time variation. In the partial adjustment model estimation we exploit this statistical power to identify the impacts of these factors on demand elasticities and therefore markups across markets.[26]

4.4 Estimation results

We use WLS to estimate the model in equation (13). For each year, we use market-category-specific revenue shares as weights within the IRI or NielsenIQ sample. For the overlapping period of 2006-2012, we adjust IRI and NielsenIQ sample weights using their observation

[24]Specifically, we use household-year-specific projection weights, provided by NielsenIQ, to calculate the retailer-specific market shares of each year, together with household-retailer-year-specific expenditures. We do not calculate HHI's for the IRI sample as it does not have the required data.

[25]See: https://fred.stlouisfed.org/series/REAINTRATREARAT1YE.

[26]Owyang et al. (2005) and Baumeister et al. (2022) find substantial differences in the timing of business cycles across the U.S.

shares within each year. As discussed in Section 2.4, it is important to use revenue share weights to capture the average effects of the factors on the markups of a typical consumer basket. Indeed, food groceries are very heterogeneous and the revenue shares reflect relative spending on different categories.

Although our dynamic panel is relatively long ($T = 24$), the regression coefficient estimates are still subject to $O(1/T)$ bias, as pointed out by Nickell (1981).[27] However, recent work by Klosin (2024) shows that such a dynamic bias is typically much smaller than the bias from alternative estimation methods such as the static panel model. We use robust three-way cluster standard errors (Cameron et al., 2011) at the market-category-year level to allow for common components in regression errors within a market, food category, and year.

Table 3 presents the results. We find that macroeconomic, demographic, and grocery market factors can explain a large share of the trend and cyclical variation of implied markups. Column (1) presents the main estimation results, which correspond to short-run effects of the explanatory variables on elasticities. Column (2) reports the long-run effects using the conversion discussed above. The adjusted R^2 is a whole-regression R^2 that takes into account degrees of freedom and reflects the variation explained by fixed effects. The within R^2 measures the explanatory power of our factor variables (excluding fixed effects). Of the 39% within R^2, 16% is due to the explanatory variables apart from the lagged dependent variable. This pattern highlights the fact that it is important to model markups in a dynamic model, while also indicating that the set of included factors explain a large share of the variation. In Columns (3) and (4) we then convert the marginal effects of the various channels on markups.

We find that population does not have a statistically significant effect on markups, possibly because it tends to be slow-moving and may be correlated with other variables, e.g. real GDP. In the short run we find a positive point estimate; in the long run the effect is zero. Grocery-market-specific factors also affect markups. In particular, the number of varieties in a specific market has a negative effect on markups, though it is not statistically significant. More varieties in a food category may lead to lower markups. The concentration of stores in a market, for example, if a particular chain has a larger market share, is associated with higher markups. The effect is statistically significant in both the short and long run.

In the group of macro effects, real per capita GDP and housing prices have strong short-run effects on markups. Real GDP per capita measures the contribution of income on markups, while housing prices measure the effect of wealth. Consistent with this interpretation, the effect of housing prices has a statistically significant long-run effect, while the effect of real GDP is not statistically significant in the long run. We note that consumption and wealth are cointegrated in the long run through the effect of permanent income on

[27] A typical standard error of an estimator in a time series application is $O(1/\sqrt{T})$. So, the second order bias in our application is comparable in scale to a standard error in a time series regression with $T = 24^2 = 576$ observations, which is common in modern empirical practice.

Table 3: Factor Regression Results for Estimated Elasticities and Implied Markup Effects.

explanatory variables[a]	dependent variable: *elasticity*			
	elasticity effects		gross markup effects[b] (*ln*)	
	(1)	(2)	(3)	(4)
	short-term	long-run[c]	short-term	long-run[c]
real interest rate	4.43***	8.84***	-2.10***	-4.20***
	(1.45)	(2.55)	(0.69)	(1.21)
real GDP per capita	-1.16***	-0.89	0.55***	0.42
	(0.31)	(0.60)	(0.15)	(0.28)
real housing price	-0.37***	-0.53**	0.17***	0.25**
	(0.09)	(0.23)	(0.04)	(0.11)
retailer HHI	-0.39**	-0.52*	0.19**	0.25*
	(0.16)	(0.29)	(0.08)	(0.14)
# UPCs per category	0.19	0.16	-0.09	-0.08
	(0.14)	(0.10)	(0.06)	(0.05)
population	-0.13	0.00	0.06	-0.00
	(0.40)	(0.36)	(0.19)	(0.17)
COVID controls[d]	YES			
one lag of all variables	YES			
fixed effects of markets and goods	YES			
$adj.R^2$	0.66			
within R^2	0.39			
N	27,526			

Note: [a]Explanatory variables are computed in natural log except for real interest rate and cumulative percent change in real housing price (both in decimal points). See details in the main text and Appendix C. [b]Using the chain rule, we can calculate annual gross markup, $\ln(1+\mu)$, effects as elasticity effect times $-1/(e_t(e_t-1))$. We revenue-weight annual average elasticity estimates e_t when computing markups effects. [c]Given the distributed lag model, we impute the long-run effects for each factor and estimate standard errors using the Delta method. [d]Pandemic controls for 2020-2023 have two parts: one captures heterogeneous level shifts across markets and the other captures interactions with real GDP per capita. All the models above are estimated by OLS using market-category-specific revenue shares as weights within each year. Standard errors, clustered at the market, food category, and year levels, are listed in the parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

consumption (Friedman, 1957). Therefore, we expect the effects of real GDP and income to measured less precisely in the long run (the standard error of real GDP approximately doubles from short to long run). In the next section, when we perform a historical decomposition of the variation in markups we group variables together and compute the overall effect of these variables. One reason for this approach is the within-group correlation across variables. We note that the effect of house prices on markups indicates that the wealth effect of homeowners may dominate the effect of potentially higher rents.[28]

The real interest rate has a negative and statistically significant effect on markups. This effect is consistent with the standard consumption and savings decision where a higher real rate results in lower demand today. Our results indicate that markups decline when the real interest rate increases. This pattern is consistent with conditionally procyclical markups. We note that this effect is present not only in the short-run but also in the long-run. If we assume that prices are not sticky over longer horizons, our result suggests a direct effect of monetary policy on grocery store prices. This result is relevant for the literature on the transmission mechanism of monetary policy.[29]

The actual economic impact of any specific variable on markups depends both on the estimated coefficient but also on the variation of the explanatory variable. In the next section we therefore perform a historical decomposition. Finally, we find evidence of a structural break in the effect of per capita GDP during the COVID-19 pandemic and its immediate aftermath.[30]

Robustness. We perform robustness checks. We include a linear time trend in the regression as an additional control. The linear time trend is not statistically significant and the main results are robust to this change (see Table B.7). In our main specification, we do not include the unemployment rate due to its high correlation and resulting potential collinearity with real GDP per capita. We obtain similar results qualitatively when we replace real GDP per capita with unemployment rate (Table B.8). As expected, higher unemployment leads to higher elasticities and thereby lower markups. Finally, we re-estimate the model by first converting to markups and then estimating the dynamic panel (Table B.9). These results represent estimates of effects on average markups rather than harmonic markups. We find that the results are robust.

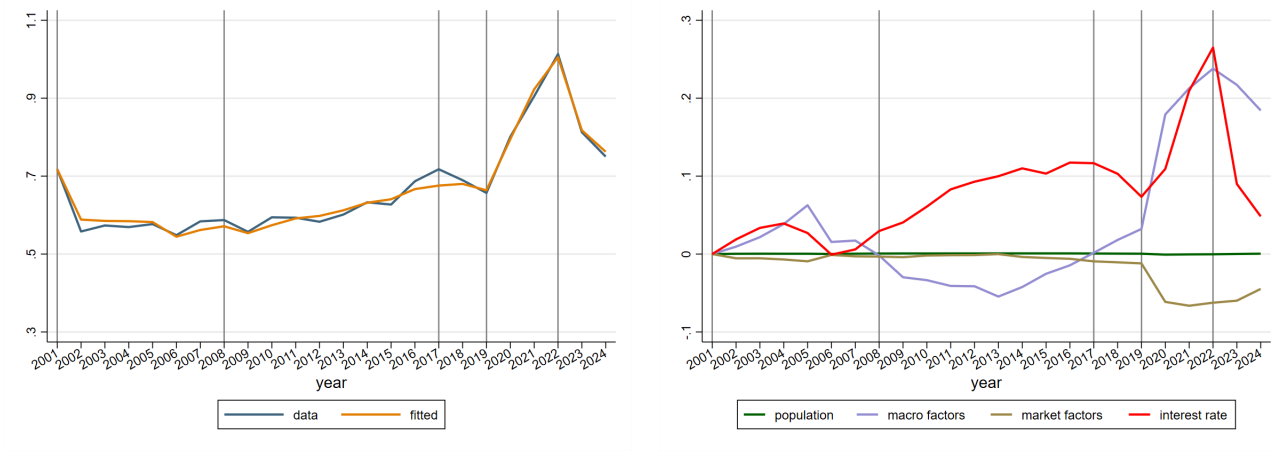
[28] We note that Stroebel and Vavra (2019) find a related wealth-effect mechanism. Their argument is based on the assumption that food retail costs (e.g., store rents) do not react to changes in housing prices, thus the positive effect of higher housing prices on grocery prices from homeowners leads to higher retail markups.

[29] See, for example Eggertsson et al. (2021).

[30] The CDC declared the end of the pandemic in 2023.

4.5 Channels of cyclical variation in markups

As discussed in Section 3 there is evidence of unconditional procyclical markups. We now use the estimation results from Table 3 to study the channels affecting cyclical variation in markups, and thereby shed light on conditional cyclicality. We perform a historical decomposition to study the contribution of different groups of factors on overall and market-specific markup dynamics. The purpose of this exercise is to measure the effect of different channels—demographic, market-specific, macroeconomic (income and wealth), and real interest rate—on the harmonic average markups in various markets while accounting for transition dynamics.



(a) Annual gross markup in *ln*

(b) Factor contribution decomposition

Figure 6: Annual Gross Markups and Factor Contributions.

Note: We first obtain predicted market-category-year-specific elasticities based on the estimated regression model (13). We then convert these predicted elasticities into annual gross markups, as in Table 2. To calculate factor contributions, we first set all factors at their initial values to obtain predicted annual gross markups under endogenous elasticity dynamics, which are used as the baseline annual gross markups. We then calculate the cumulative contributions of each group of factors by taking the cumulative differences between the factor-predicted annual gross markups and the baseline annual gross markups. We group our six factors into four groups: (1) population; (2) real GDP per capita and cumulative changes in real housing prices combined as macro factors together with pandemic effects; (3) the number of UPCs per product category and retailer HHI combined as grocery market factors; and (4) real interest rate.

For each market, we calculate the difference in predicted values of harmonic markups under scenarios in which a given factor remains constant at the initial level and one where the factor follows the historical trajectory. We use results from the specification in Table 3 to calculate dynamics of average elasticities and then convert those to harmonic markups. We revenue-weight across categories within a market and then across markets. Figure 6 Panel (a) reports overall markups, across all of the markets in our sample, together with the markups predicted by the variation in the explanatory variables. The figure shows that there is a close fit of unconditional markup dynamics.

There are both slow-moving and higher-frequency channels affecting markups. We note a slow upward trend in markups due to population and a slow downward trend due to grocery market factors. This implies that changes in the competitive landscape put downward pressure on markups over the sample period and suggests that the grocery retail sector became more competitive. We revisit this implication in Section 5.

We find that higher frequency cyclical changes are mainly driven by macro factors and the real interest rate as well as transitional dynamics. For example, in 2008, there is a negative effect of macro factors on markups, and in 2017 higher real interest rates put downward pressure on markups. Overall, macro factors reduced markups from 2009 to 2013, a period of somewhat sluggish growth after the Great Recession. In contrast, accommodative monetary policy had a supportive effect on markups during the period. From 2019 to 2020 markups increased, which was at least partly due to COVID-specific developments captured by the macro factors.

Aggregate markups are averages across markets and categories. In our panel data set all variables, except the real interest rate, vary at least at the market level. That is, real GDP growth is different in LA and Chicago. This setup therefore allows us to explore the ability of our model to explain regional variation in markups.

We next consider the largest markets from the four macro-regions of the country – NYC (East), Chicago (Midwest), Dallas (South) and LA (West). Figure 7 shows overall actual and fitted markups as well as factor-specific historical decompositions.

The top row of Figure 7 plots average actual and fitted harmonic average markups. The level of fit for the individual markets is similar as what we saw for the overall average. This implies that the unexplained variation is primarily due to differences across categories.

The bottom row of Figure 7 shows the contribution of the various channels for Chicago, New York, Dallas, and Los Angeles. These plots therefore reflect conditional markup dynamics by showing effects of different channels. We find that population dynamics generally have a very small impact on markups, except for Dallas. The effect of grocery market factors is similar across markets. The largest driver of differences across markets are the macro factors – per capita GDP growth, housing prices, and the effects of the COVID pandemic. The largest post-Great-Recession impact was in LA, which also saw the highest recent increase in markups due to macro factors. Chicago, which experienced only small changes in macro factors, saw the lowest overall increase in markups during the sample period.

We note that, in the figures, the real interest rate has a slightly different reported effect across markets. This is because of compositional effects in the unbalanced panel. Therefore, the average effects of common factors are not exactly aligned across markets.

The results imply that our model does a good job explaining variation in average markups over time across markets. We also note that the real interest rate plays an important role in explaining variation in markups, a point which we discuss in the next section.

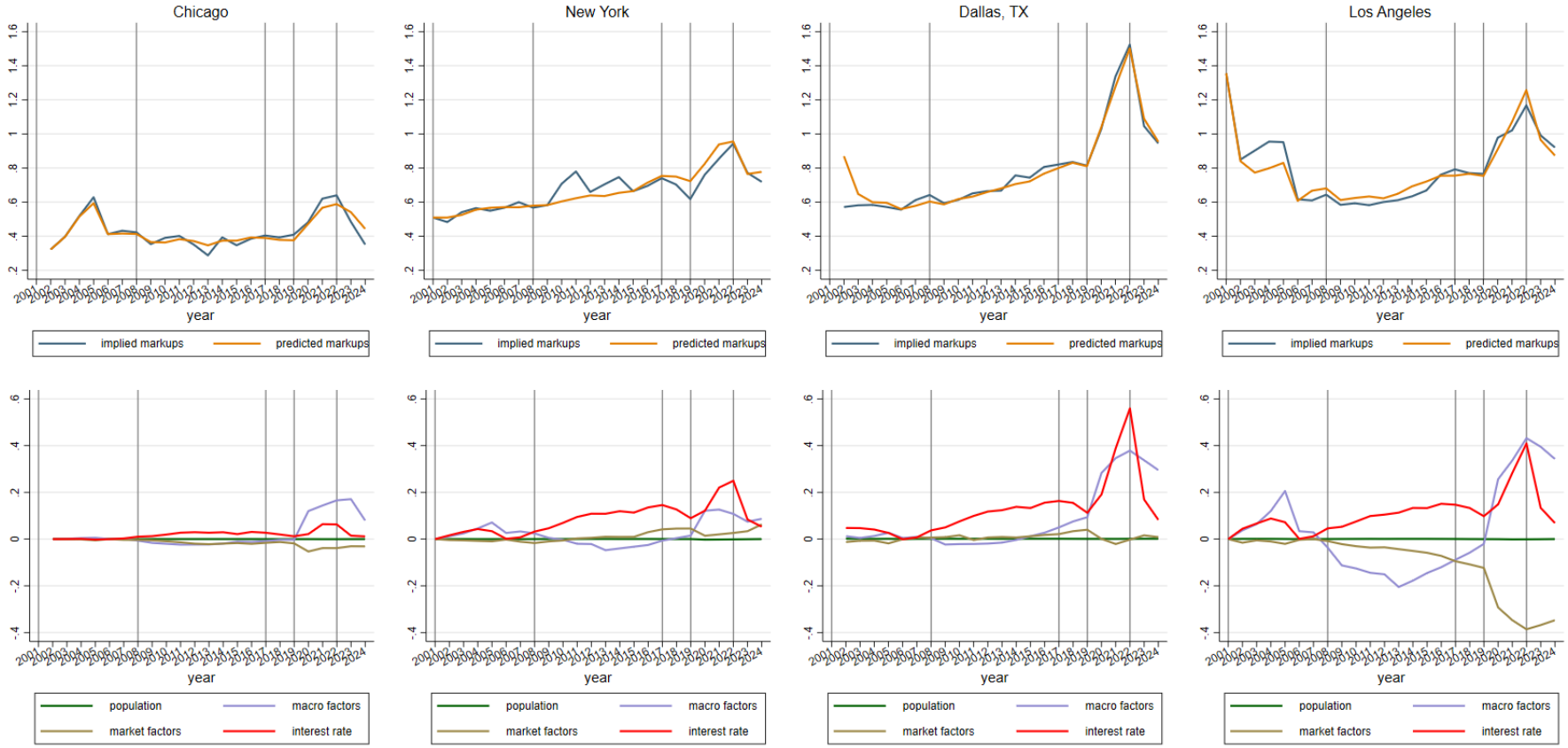


Figure 7: Revenue-weighted Annual Gross Markups of the Four Largest Regional Markets.

Notes: We calculate and decompose the predicted annual gross markups as in Figure 6. Chicago does not have elasticity estimates in 2001 and thereby no annual gross markup in that year. For Chicago, we also use the average annual gross markups of 2002 and 2004 for 2003 in which it has limited elasticity estimates.

5 Discussion

5.1 Evidence of procyclical markups

We document procyclical movement in markups. This finding is related to a growing empirical literature on procyclical markups as summarized in Nekarda and Ramey (2020). The existing literature predominantly uses the production function approach which can be prone to specification errors (Bond et al., 2021). Complementing earlier work, our paper provides nonparametric demand-side evidence on the procyclicality of optimal markups (the markups in the absence of sticky prices).

Based on the results from our dynamic panel estimation, we find effects of income and wealth shocks and the interest rate. In our estimation approach we can distinguish between these two effects using cross-sectional variation in demand shocks. As shown in Table 2, the monetary tightening in 2017 coincided with an average decrease of 4.7% in gross markups, while the sector-specific demand shock from COVID and concurrent government spending stimulus in 2020 was associated with an average increase of 12% in gross markups. These effects are unconditional responses. The conditional responses in Table 3 reflect different effects potentially offsetting one another. For example, the negative aggregate demand shock in 2008 resulted in only a small change in markups. Our estimates imply that loose monetary policy counteracted low aggregate demand. To be specific, if we combine population, grocery market and macro factors from 2007 to 2011, the cumulative effect on markups would have been -4.7%. Adding in the interest factor instead resulted in a cumulative change of +3.0% (see Figure 6). Our results therefore imply an important role of monetary policy in stabilizing prices through their effect on gross markups.

The patterns in grocery-market factors are consistent with earlier theoretical models of countercyclical markups (Rotemberg and Saloner, 1986; Rotemberg and Woodford, 1999). We find that competition tends to increase during expansions, e.g. the recovery from the great recession and the expansion following the 2020 recession. However, our results imply that this effect is not strong enough to counteract the direct positive effect of aggregate demand and monetary policy on markups.

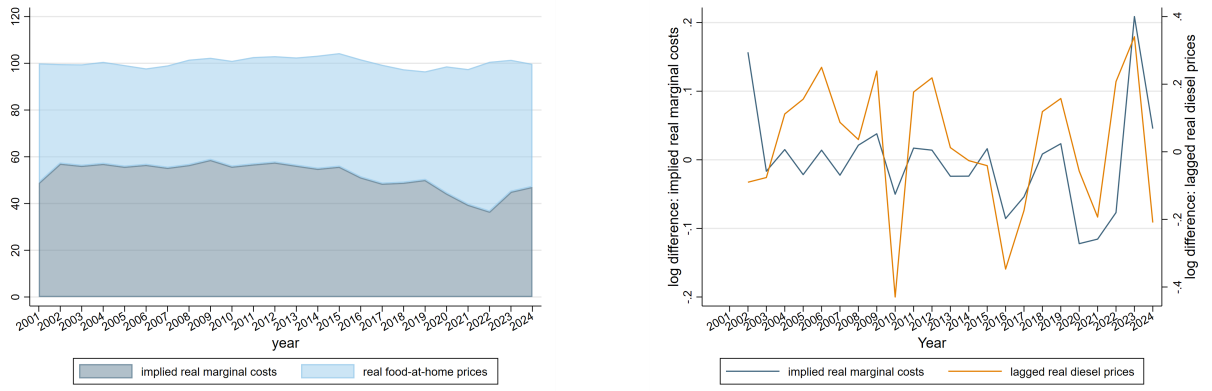
5.2 Implication for movements in marginal costs

Our results can be used to explore the dynamics of marginal costs by combining observed food-at-home prices with our new estimates of time-varying markups. Each year, we deflate the food-at-home CPI using our average markup estimate and the overall CPI. Panel (a) of Figure 8 plots food at home prices deflated using the overall CPI and the resulting time series of real marginal costs. We note that for food at home prices changed little over the sample

period.[31]

In contrast, there was quite a bit of variation in real marginal costs. We note a decline in real marginal costs of food from 2019 to 2022 and an increase from 2022 to 2023. We expect variation in marginal costs to be driven by variation on input prices. Specifically, it is possible that the decline and subsequent increase in marginal costs were related to fuel prices. Diesel prices represent a large share of marginal costs for some food categories such as milk (Gafarov et al., 2025) and the timing of the reduction and subsequent increase in marginal costs is aligned with changes in diesel prices. Due to evidence of a lagged pass-through from diesel prices to marginal costs and therefore food prices, we compare changes in real marginal costs and lagged changes in real diesel prices. The correlation coefficient is equal to 0.44 for our sample period. Panel (b) of Figure 8 plots log diesel prices and log real marginal costs. We note a strong comovement especially during the second half of the sample.

Our findings are consistent with the literature on oil shocks and inflation. Coibion and Gorodnichenko (2015) observed that oil prices explain approximately one half of the variation in consumer inflation expectations. One possible channel is that diesel prices are a major contributor to the variation in marginal costs of groceries, and grocery prices drive consumer expectations of inflation (D'Acunto et al., 2021).



(a) Prices and Marginal Costs of Food-at-Home

(b) Changes in Diesel Prices and Marginal Costs

Figure 8: Prices and Marginal Costs of Food at Home v.s. Diesel Retail Prices.

Note: We obtain food-at-home CPI and overall CPI from FRED. We obtain No.2 Diesel Retail Prices from EIA. For Panel (a), we deflate food-at-home CPI by the overall CPI to obtain real prices for food at home. We also normalize them to the base year of 2001. We then impute real marginal costs for food at home using our estimated annual gross markups, following equation (3). For Panel (b), we deflate diesel prices before taking log difference; we also take log difference for real marginal costs for food at home.

[31]Over our sample period the CPI subindex for food at home (food groceries) increased by 76.8% while the overall CPI increased by 77.2%. The overall CPI index refers to the "Consumer Price Index for all urban consumers: all items in U.S., city average" and the CPI subindex for food is the "Consumer Price Index for all urban consumers: food at home in U.S., city average." We obtain both data from FRED.

6 Conclusion

In this paper, we make two main contributions to the literature on markup estimation and the cyclicality of markups. First, we estimate revenue-weighted demand elasticities in a panel-IV regression model that implements the widely-used Hausman (1996) IV strategy. This allows us to compute a panel of harmonic markup indices. Our new estimates of markups are non-parametric in the sense that we do not assume any single particular parametric model of consumer demand and are based on a local average derivative of the demand function in the neighborhood of the market equilibrium (Imbens and Angrist, 1994). Our particular implementation only assumes common wholesale suppliers for each specific pair of geographically close markets. We work directly with weekly sales data which results in strong price IVs.

Second, we use our new estimates to explore dynamics of markups and the channels through which aggregate demand affects them. We provide direct micro-level evidence of a slow-moving upward trend in US markups of 1.4% per year from 2001 to 2024. Moreover, our markup measure exhibits economically large and statistically significant drops at times of contractions in aggregate demand, implying unconditional procyclical variation in economy-wide markups. The patterns in food markups that we find are representative of the whole economy since we cover markets in the Western, Mid-Western, Eastern, and Southern US.

We use a dynamic adjustment model to estimate the partial effects of aggregate demand on markups corresponding to various channels. We then conduct a structural decomposition to attribute changes in markups across regions and over time to various factors. This exercise allows us to study cyclicality of markups conditional on various channels. Our structural attribution also allows us to answer the question we asked at the beginning: Does policy intervention supporting aggregate demand increase competition? We find that increases in income and wealth together with accommodative monetary policy have strong positive effects on markups. Changes in the competitive landscape affect markups countercyclically—when competition increases during economic expansions, markups decline. Overall, as reflected in the unconditional dynamics, the effects of income and wealth dominate.

References

Abadie, A. (2003): "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, 113, 231–263.

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2020): "Sampling-based versus design-based uncertainty in regression analysis," *Econometrica*, 88, 265–296.

Afrouzi, H. and L. Caloi (2024): "Endogenous firm competition and the cyclicality of markups," *Review of Economics and Statistics*, 106, 1638–1654.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013): "Time use during the great recession," *American Economic Review*, 103, 1664–1696.

Anderson, E., B. A. Malin, E. Nakamura, D. Simester, and J. Steinsson (2017): "Informational rigidities and the stickiness of temporary sales," *Journal of Monetary Economics*, 90, 64–83.

Anderson, E., S. Rebelo, and A. Wong (2023): "Markups across space and time," *Unpublished Manuscript*.

Anderson, E. T. and E. J. Fox (2019): "How price promotions work: A review of practice and theory," *Handbook of the Economics of Marketing*, 1, 497–552.

Angrist, J. D., K. Graddy, and G. W. Imbens (2000): "The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish," *The Review of Economic Studies*, 67, 499–527.

Bachmann, R. and E. R. Sims (2012): "Confidence and the transmission of government spending shocks," *Journal of Monetary Economics*, 59, 235–249.

Backus, M., C. Conlon, and M. Sinkinson (2021): "Common ownership and competition in the ready-to-eat cereal industry," Tech. rep., National Bureau of Economic Research.

Balestra, P. and M. Nerlove (1966): "Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas," *Econometrica: Journal of the econometric society*, 585–612.

Balk, B. M. (2012): *Price and quantity index numbers: models for measuring aggregate change and difference*, Cambridge University Press.

Baumeister, C., D. Leiva-León, and E. Sims (2022): "Tracking weekly state-level economic conditions," *Review of Economics and Statistics*, 1–45.

Bernanke, B. S., M. Gertler, M. Watson, C. A. Sims, and B. M. Friedman (1997): "Systematic monetary policy and the effects of oil price shocks," *Brookings papers on economic activity*, 1997, 91–157.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890.

Bogin, A., W. Doerner, and W. Larson (2019): "Local house price dynamics: New indices and stylized facts," *Real Estate Economics*, 47, 365–398.

Bond, S., A. Hashemi, G. Kaplan, and P. Zoch (2021): "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data," *Journal of Monetary Economics*, 121, 1–14.

Bonhomme, S., T. Lamadon, and E. Manresa (2022): "Discretizing unobserved heterogeneity," *Econometrica*, 90, 625–643.

Bonhomme, S. and E. Manresa (2015): "Grouped patterns of heterogeneity in panel data," *Econometrica*, 83, 1147–1184.

Brand, J. (2021): "Differences in differentiation: Rising variety and markups in retail food stores," *Available at SSRN 3712513*.

Broer, T., N.-J. Harbo Hansen, P. Krusell, and E. Öberg (2020): "The new Keynesian transmission mechanism: A heterogeneous-agent perspective," *The Review of Economic Studies*, 87, 77–101.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela (2008): "Database paper—The IRI marketing data set," *Marketing Science*, 27, 745–748.

Bugni, F. A., I. A. Canay, and S. McBride (2023): "Decomposition and interpretation of treatment effects in settings with delayed outcomes," *arXiv preprint arXiv:2302.11505*.

Burdett, K. and K. L. Judd (1983): "Equilibrium price dispersion," *Econometrica: Journal of the Econometric Society*, 955–969.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011): "Robust inference with multiway clustering," *Journal of Business & Economic Statistics*, 29, 238–249.

Chetverikov, D. and E. Manresa (2022): "Spectral and post-spectral estimators for grouped panel data models," *arXiv preprint arXiv:2212.13324*.

Coggeshall, F. (1886): "The arithmetic, geometric, and harmonic means," *The Quarterly Journal of Economics*, 1, 83–86.

Coibion, O. and Y. Gorodnichenko (2015): "Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation," *American Economic Journal: Macroeconomics*, 7, 197–232.

Consumer Reports (2025): "Instacart AI pricing experiment inflating grocery bills," *Consumer Reports*, online article.

Cournot, A.-A. (1838): *Recherches sur les principes mathématiques de la théorie des richesses par Augustin Cournot*, chez L. Hachette.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2021): "Exposure to grocery prices and inflation expectations," *Journal of Political Economy*, 129, 1615–1639.

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

De Loecker, J. and P. T. Scott (2022): "Markup Estimation using Production and Demand Data: An Application to the US Brewing Industry," Accepted at the Review of Economic Studies (2025).

De Ridder, M., B. Grassi, and G. Morzenti (2024): "The hitchhiker's guide to markup estimation: assessing estimates from financial data," *Work. Pap., London Sch. Econ./Bocconi Univ./Analysis Group, UK/Milan, Italy/Paris, France*.

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

Döpper, H., A. MacKay, N. H. Miller, and J. Stiebale (2025): "Rising Markups and the Role of Consumer Preferences," *Journal of Political Economy*, 133, 2462–2505.

Eggertsson, G. B., J. A. Robbins, and E. G. Wold (2021): "Kaldor and Piketty's facts: The rise of monopoly power in the United States," *Journal of Monetary Economics*, 124, S19–S38.

Eichenbaum, M., N. Jaimovich, and S. Rebelo (2011): "Reference Prices, Costs, and Nominal Rigidities," *American Economic Review*, 101, 234–262.

Friedman, M. (1957): "Introduction to" A theory of the consumption function"," in *A theory of the consumption function*, princeton university press, 1–6.

Gaarder, I., L. H. de Frahan, M. Mogstad, A. Torgovitsky, and O. Volpe (2024): "Linear Supply and Demand in Heterogeneous Markets," *Unpublished paper*.

Gafarov, B., T. Gong, and J. Hilscher (2024): "Do organic products have higher markups?" *UC Davis working paper*.

Gafarov, B., J. Hasbany, and J. Hilscher (2025): "Diesel price passthrough into California grocery store milk prices," *Unpublished manuscript, UC Davis*.

Gallego, G., H. Topaloglu, et al. (2019): *Revenue management and pricing analytics*, vol. 209, Springer.

Gandhi, A., S. Navarro, and D. A. Rivers (2020): "On the identification of gross output production functions," *Journal of Political Economy*, 128, 2973–3016.

Grieco, P. L., C. Murry, and A. Yurukoglu (2024): "The evolution of market power in the us automobile industry," *The Quarterly Journal of Economics*, 139, 1201–1253.

Hausman, J. A. (1996): "Valuation of new goods under perfect and imperfect competition," in *The economics of new goods*, University of Chicago Press, 207–248.

Houthakker, H. S., P. K. Verleger, and D. P. Sheehan (1974): "Dynamic demand analyses for gasoline and residential electricity," *American Journal of Agricultural Economics*, 56, 412–418.

Imbens, G. W. and J. D. Angrist (1994): "Identification and estimation of local average treatment effects," *Econometrica: Journal of the Econometric Society*, 467–475.

Keynes, J. M. (1936): *The general theory of interest, employment and money*, London: MacMillan.

Klosin, S. (2024): "Dynamic biases of static panel data estimators," *arXiv preprint arXiv:2410.16112*.

Lerner, A. (1934): "The concept of monopoly and the measurement of monopoly power," *The Review of Economic Studies*, 1, 157–175.

Levin, L., M. S. Lewis, and F. A. Wolak (2017): "High frequency evidence on the demand for gasoline," *American Economic Journal: Economic Policy*, 9, 314–347.

MacKay, A. and N. Miller (2023): "Estimating models of supply and demand: Instruments and covariance restrictions," *Harvard Business School Strategy Unit Working Paper*.

Mankiw, N. G. and R. Reis (2002): "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve," *The Quarterly Journal of Economics*, 117, 1295–1328.

Munro, E., S. Wager, and K. Xu (2023): "Treatment effects in market equilibrium," *arXiv preprint arXiv:2109.11647*.

Neiman, B. and J. Vavra (2023): "The rise of niche consumption," *American Economic Journal: Macroeconomics*, 15, 224–264.

Nekarda, C. J. and V. A. Ramey (2020): "The cyclical behavior of the price-cost markup," *Journal of Money, Credit and Banking*, 52, 319–353.

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

Nickell, S. (1981): "Biases in dynamic models with fixed effects," *Econometrica: Journal of the econometric society*, 1417–1426.

Owyang, M. T., J. Piger, and H. J. Wall (2005): "Business cycle phases in US states," *Review of Economics and Statistics*, 87, 604–616.

Phelps, E. S. and S. G. Winter (1970): "Optimal price policy under atomistic competition," *Microeconomic foundations of employment and inflation theory*, 309–337.

Philippon, T. (2019): *The great reversal*, Harvard University Press.

Richards, T. J. and S. F. Hamilton (2015): "Variety pass-through: An examination of the ready-to-eat breakfast cereal market," *Review of Economics and Statistics*, 97, 166–180.

Rossi, P. E. (2014): "Even the rich can make themselves poor: A critical examination of IV methods in marketing applications," *Marketing Science*, 33, 655–672.

Rotemberg, J. J. and G. Saloner (1986): "A supergame-theoretic model of price wars during booms," *The American economic review*, 76, 390–407.

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

——— (1999): "The cyclical behavior of prices and costs," *Handbook of macroeconomics*, 1, 1051–1135.

Sims, C. A., J. H. Stock, and M. W. Watson (1990): "Inference in linear time series models with some unit roots," *Econometrica: Journal of the Econometric Society*, 113–144.

Staiger, D. and J. H. Stock (1997): "Instrumental variables regression with weak instruments," *Econometrica*, 65, 557–586.

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

Wager, S. and K. Xu (2021): "Experimenting in equilibrium," *Management Science*, 67, 6694–6715.

Wald, A. (1940): "The fitting of straight lines if both variables are subject to error," *The annals of mathematical statistics*, 11, 284–300.

Woodford, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

Working, E. J. (1927): "What do statistical 'demand curves' show?" *The Quarterly Journal of Economics*, 41, 212–235.

Appendix

A Details on IRI and NielsenIQ datasets

For comparability, we do not take the lower product module in the NielsenIQ sample as the product category, since it has far fewer UPCs than a product category in the IRI sample does. Instead, we take food groups as the product category for the NielsenIQ sample. Table B.3 shows that these two samples now have a comparable number of UPCs per food category in magnitude, although the NielseIQ sample has more UPCs per food category. We also drop four food categories or groups in the NielsenIQ sample: ICE, YEAST, GUM, and PET FOOD. For the first two, our paired markets have a limited number of common products at the UPC level, which would make our IV strategy infeasible if we included these two food categories in our estimation sample. For the last two food categories, we drop them as we focus on edible human food in this paper.

Regarding UPCs, we take them as given in the IRI and NielsenIQ retail scanner data sets. Both data sets have collapsed private-label UPCs across retailers to some extent. The assumption is that the core attributes are identical for these UPCs, although there may be some custom characteristics that differ depending on how the different products are marketed. Using these collapsed UPCs and retailer codes, we can construct retailer-specific private-label UPCs within each food category. However, we find that it will not only weaken the strength of our price IVs but also lower our elasticity estimates. This suggests that refining private-label UPCs will not help us recover reasonable price elasticities of demand. Hence, we stick to the collapsed private-label UPCs in our main results. Results based on retailer-specific private-label UPCs are available upon request.

When it comes to the food stores of paired markets listed in Table B.1, we focus on panel stores that have reported product transactions across all 12 years in the IRI retail scanner data set and across all 19 years in the NielsenIQ retail scanner data set. Both data sets have store entries and exits over time. In particular, NielsenIQ's coverage rate for food stores decreased from 53% in 2011 to 26% in 2017. We study time variation of annual markups among food stores that are covered in the data across all years. In this way, our results will not be mechanically driven by any changes in the store coverage of the data.

B Additional tables and figures

Table B.1 presents the specific pairs of the 25 markets shown on the US mainland map (Figure 1). Table B.2 reports the number of observations and stores across markets as well as population for the IRI and NielsenIQ samples. In total, each year, the IRI data include

close to 48 million observations from 429 stores, while the NielsenIQ data contain over 1700 million observations from 2884 stores. There are notable size differences across Markets. For example, the market with the most stores and observations, LOS ANGELES, had 12 times more population than Syracuse, one of the smallest markets, in 2010.

Table B.1: The 12 Pairs of 25 Markets.

region	market(s) A	market(s) B
EAST	BUFFALO/ROCHESTER[a]	SYRACUSE
	BOSTON	HARTFORD
	NEW YORK	PHILADELPHIA
SOUTH	CHARLOTTE	RALEIGH/DURHAM
	ATLANTA	SOUTH CAROLINA
	DALLAS, TX	HOUSTON
WEST	LOS ANGELES	SAN DIEGO
	SACRAMENTO	SAN FRANCISCO
	PORTLAND, OR	SEATTLE/TACOMA
MIDWEST	KANSAS CITY	DES MOINES + OMAHA[b]
	CHICAGO	PEORIA/SPRINGFLD.
	DETROIT	TOLEDO

Note: We take markets defined by IRI as given. Each IRI market consists of one or more adjacent counties. We use county-specific FIPS codes identify markets to which NielsenIQ stores belong. [a]The "/" means that the two areas belong to a single market. [b]The "+" means that the two areas belong to two markets but are combined as one by us when implementing our identification strategy. Specifically, when there is no single neighboring market with a sufficient number of stores, we pair a large market of interest with two smaller neighboring ones and utilize their average product prices to instrument the product prices of the large market of interest. However, when the two small markets are of interest, we treat them separately and use the product prices of the large market to instrument their product prices. Hence, we have 25 markets in total for demand elasticity estimations.

Table B.2: Summary Statistics of Markets.

market name	population in 2010 (million)	IRI (2001-2012) obs per year (million)	IRI (2001-2012) panel stores (count)	NielsenIQ (2006-2024) obs per year (million)	NielsenIQ (2006-2024) panel stores (count)
NEW YORK	19.5	4.9	41	139.4	244
LOS ANGELES	17.1	5.6	55	242.7	413
CHICAGO	9.0	1.9	16	124.8	178
PHILADELPHIA	6.5	3.0	24	58.1	108
DALLAS, TX	6.2	1.5	15	99.8	163
SAN FRANCISCO	6.1	2.4	23	100.3	171
HOUSTON	5.9	2.0	20	80.3	122
BOSTON	5.5	2.6	19	87.1	138
SOUTH CAROLINA	5.1	1.6	20	70.0	138
ATLANTA	4.9	2.2	17	82.4	114
DETROIT	4.8	0.9	7	58.7	79
SEATTLE/TACOMA	3.4	2.4	24	105.2	169
RALEIGH/DURHAM	3.3	1.8	19	121.6	236
HARTFORD	3.2	1.1	8	14.6	28
PORTLAND, OR	3.2	1.6	15	85.0	137
SAN DIEGO	3.1	1.8	16	44.4	72
SACRAMENTO	2.8	1.4	12	32.5	56
CHARLOTTE	2.6	1.3	13	90.4	176
BUFFALO/ROCHESTER	2.5	1.7	15	0.9	8
TOLEDO	2.0	0.9	7	21.6	29
PEORIA/SPRINGFLD.	2.0	0.6	5	17.9	27
KANSAS CITY	2.0	2.3	17	15.1	18
SYRACUSE	1.2	0.8	8	4.5	10
OMAHA	1.1	0.8	6	21.5	28
DES MOINES	0.7	0.9	7	15.7	22
total	123.8	47.8	429	1734.5	2884

Note: We report the store-UPC-week observations only for the food categories listed on Table B.6.

Table B.3: Summary Statistics of Food Categories.

	IRI categories[a]			NielsenIQ categories[b]		
	mean	p10	p90	mean	p10	p90
obs per year (million)	3.0	0.8	6.1	29.9	7.3	63.9
No. of UPCs per year	1887.2	283.8	4307.3	4114.8	599.8	10024.4
No. of UPCs per year-market	467.8	106.3	959.0	953.4	172.7	1963.5

Note: [a]IRI has 16 food categories. [b]NielsenIQ has 58 food categories. See the list below.

Table B.4: Summary Statistics of Stores.

	IRI stores			NielsenIQ stores		
	mean	p10	p90	mean	p10	p90
obs per year (thousand)	113.1	81.5	145.3	610.0	437.3	801.9
No. of weeks per year	50.9	48.8	52.1[a]	52.0	51.6	52.2[b]
No. of categories per year	16.0	16.0	16.0	58.0	58.0	58.0
No. of UPCs per year-category	222.3	171.8	272.2	372.5	270.1	477.1

Note: [a]A IRI week starts on Monday and ends on Sunday. [b]A NielsenIQ week starts on Sunday and ends on Saturday. Hence, both IRI and NielsenIQ stores have sales records of 53 weeks in some years.

At the store level, Table B.4 shows that almost all panel stores report transaction records for most weeks. The average reported weeks per year-store are 51 and 52 for the IRI and NielsenIQ samples, respectively. Moreover, stores sell products for all the food categories of interest in this paper. Hence, we have sufficient observations nearly for every store. The average obervations per year are 113 and 610 thousands for the IRI and NielsenIQ samples, respectively, at the store level. Stores sell fewer UPC products per category than markets.

For each market-category-year pair, we use store-UPC-specific quantity shares within a week to compute the quantity weighted average of log weekly store-UPC-specific prices (IV) across all stores in the paired market. We impute the weekly store-UPC-specific prices from the data of weekly store-UPC-specific quantities and revenues before taking log. After constructing the price IV, we use store-UPC-specific revenue shares within a year to estimate the revenue-weighted price elasticity of demand based on the IV regression (9). These revenue weights represent a typical food basket of each category purchased by consumers in each market and year. We apply the same weights to both stages of each IV regression.

Table B.5: Summary Statistics of the Equally-weighted Demand Elasticity Estimates.

	IRI sample[a]			NielsenIQ sample[b]		
	mean	p10	p90	mean	p10	p90
point estimate	2.46	1.64	3.31	1.62[c]	0.87	2.44
standard error	0.19	0.09	0.34	0.13	0.04	0.27
t statistics	16.37	7.24	27.11	18.44	5.95	33.43
obs in an estimation (thousands)	118.3	18.8	257.4	1132.9	57.3	2859.0
Cragg-Donald F statistics	131.6	14.8	274.8	1457.1	34.1	2822.3

Note: [a]The IRI sample has 4,009 elasticity estimates. [b]The NielsenIQ sample has 25,432 elasticity estimates. These samples include elasticity estimates that are obtained under the strong price IV and exclude those that have negative values. [c]The NielsenIQ sample has lower elasticity estimates than the IRI sample. The downward trend in demand elasticities over time largely contributes to the lower elasticity estimates in the NielsenIQ sample of more recent years. The NielsenIQ sample also has a small proportion of price-inelastic goods such as baby food, spices, seasoning, extracts and eggs.

For comparison, we report the following summary stats for the equally-weighted IV estimates. Table B.5 shows that the equally-weighted elasticity estimates are sizably lower not only in mean but also in percentiles, such as the 10 and 90 percentiles of the distribution, compared to the revenue-weighted estimates as summarized in Table 1. However, their standard errors are slightly smaller.

Table B.6: The List of Food Categories.

IRI food (16)	NielsenIQ non-dry food (22)[a]	NielsenIQ dry food (36)[b]
carbonated beverages	BAKED GOODS-FROZEN	BABY FOOD
coffee	BREAKFAST FOODS-FROZEN	BAKING MIXES
cold cereal	BUTTER AND MARGARINE	BAKING SUPPLIES
frozen dinners&entrees	CHEESE	BREAD AND BAKED GOODS
frozen pizza	COT CHEESE, SOUR CREAM, TOPPINGS	BREAKFAST FOOD
hotdog	DESSERTS/FRUITS/TOPPINGS-FROZEN	CANDY
margarine butter	DOUGH PRODUCTS	**CARBONATED BEVERAGES**
mayonnaise	DRESSINGS/SALADS/PREP FOODS-DELI	CEREAL
milk	EGGS	**COFFEE**
mustard&ketchup	FRESH MEAT	CONDIMENTS, GRAVIES, AND SAUCES
peanut butter	FRESH PRODUCE	COOKIES
salty snacks	ICE CREAM, NOVELTIES	CRACKERS
soup	JUICES, DRINKS-FROZEN	DESSERTS, GELATINS, SYRUP
spaghetti&italian sauce	**MILK**	FLOUR
sugar substitute	PACKAGED MEATS-DELI	FRUIT - CANNED
yogurt	PIZZA/SNACKS/HORS DOEURVES-FRZN	FRUIT - DRIED
	PREPARED FOODS-FROZEN	JAMS, JELLIES, SPREADS
	PUDDING, DESSERTS-DAIRY	JUICE, DRINKS - CANNED, BOTTLED
	SNACKS, SPREADS, DIPS-DAIRY	NUTS
	UNPREP MEAT/POULTRY/SEAFOOD-FRZN	PACKAGED MILK AND MODIFIERS
	VEGETABLES-FROZEN	PASTA
	YOGURT	PICKLES, OLIVES, AND RELISH
		PREPARED FOOD-DRY MIXES
		PREPARED FOOD-READY-TO-SERVE
		SALAD DRESSINGS, MAYO, TOPPINGS
		SEAFOOD - CANNED
		SHORTENING, OIL
		SNACKS
		SOFT DRINKS-NON-CARBONATED
		SOUP
		SPICES, SEASONING, EXTRACTS
		SUGAR, SWEETENERS
		TABLE SYRUPS, MOLASSES
		TEA
		VEGETABLES - CANNED
		VEGETABLES AND GRAINS - DRIED

Note: [a]NielsenIQ non-dry food categories span five departments including dairy, deli, fresh produce, frozen foods, and packaged meat. We exclude ICE and YEAST as our elasticity estimation records show that their UPC-level products are not commonly shared by neighboring markets such that there are missing elasticity estimates of these categories for many markets. These two categories have limited revenue shares within each store though. [b]NielsenIQ dry food refers to food categories in the dry grocery department; we exclude GUM and PET FOOD that are not of research interest in this paper. The bold food categories are the ones that IRI and NielsenIQ share the same names, although NielsenIQ have more varieties or larger numbers of UPCs within each of these five common food categories.

Table B.7: Factor Regression Results for Estimated Elasticities and Implied Markup Effects with a Linear Time Trend.

explanatory variables[a]	dependent variable: *elasticity*			
	elasticity effects		gross markup effects[b] (*ln*)	
	(1)	(2)	(3)	(4)
	short-term	long-run[c]	short-term	long-run[c]
real interest rate	4.50***	8.64***	-2.14***	-4.10***
	(1.42)	(2.78)	(0.68)	(1.32)
real GDP per capita	-1.07***	-0.59	0.51***	0.28
	(0.28)	(0.53)	(0.13)	(0.25)
real housing price	-0.37***	-0.55**	0.17***	0.26**
	(0.09)	(0.23)	(0.04)	(0.11)
retailer HHI	-0.39**	-0.51*	0.19**	0.24*
	(0.16)	(0.29)	(0.08)	(0.14)
# UPCs per category	0.19	0.17	-0.09	-0.08
	(0.14)	(0.11)	(0.07)	(0.05)
population	-0.11	0.05	0.05	-0.03
	(0.41)	(0.39)	(0.19)	(0.19)
linear time trend	-0.003			
	(0.004)			
COVID controls[d]	YES			
one lag of all variables	YES			
fixed effects of markets and goods	YES			
$adj.R^2$	0.66			
within R^2	0.39			
N	27,526			

Note: [a]Explanatory variables are computed in natural log except for real interest rate and cumulative percent change in real housing price (both in decimal points). See details in the main text and Appendix C. [b]Using the chain rule, we can calculate annual gross markup, $\ln(1+\mu)$, effects as elasticity effect times $-1/(e_t(e_t-1))$. We revenue-weight annual average elasticity estimates e_t when computing markups effects. [c]Given the distributed lag model, we impute the long-run effects for each factor and estimate standard errors using the Delta method. [d]Pandemic controls for 2020-2023 have two parts: one captures heterogeneous level shifts across markets and the other captures interactions with real GDP per capita. All the models above are estimated by OLS using market-category-specific revenue shares as weights within each year. Standard errors, clustered at the market, food category, and year levels, are listed in the parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.8: Factor Regression Results for Estimated Elasticities and Implied Markup Effects with Unemployment Rate.

explanatory variables[a]	dependent variable: *elasticity*			
	elasticity effects		gross markup effects[b] (*ln*)	
	(1)	(2)	(3)	(4)
	short-term	long-run[c]	short-term	long-run[c]
real interest rate	6.36***	10.29***	-3.01***	-4.86***
	(1.40)	(2.35)	(0.66)	(1.11)
unemployment rate	3.25***	3.27*	-1.54***	-1.55*
	(0.76)	(1.95)	(0.36)	(0.92)
real housing price	-0.31***	-0.64***	0.14***	0.30***
	(0.10)	(0.23)	(0.05)	(0.11)
retailer HHI	-0.36**	-0.46*	0.17**	0.22*
	(0.16)	(0.27)	(0.07)	(0.13)
# UPCs per category	0.19***	0.14	-0.09***	-0.06
	(0.06)	(0.10)	(0.03)	(0.05)
population	-0.30	-0.20	0.14	0.10
	(0.23)	(0.37)	(0.11)	(0.17)
COVID controls[d]	YES			
one lag of all variables	YES			
fixed effects of markets and goods	YES			
$adj.R^2$	0.66			
within R^2	0.39			
N	27,526			

Note: [a]Explanatory variables are computed in natural log except for real interest rate, unemployment rate and cumulative percent change in real housing price (both in decimal points). See details in the main text and Appendix C. [b]Using the chain rule, we can calculate annual gross markup, $\ln(1+\mu)$, effects as elasticity effect times $-1/(e_t(e_t-1))$. We revenue-weight annual average elasticity estimates e_t when computing markups effects. [c]Given the distributed lag model, we impute the long-run effects for each factor and estimate standard errors using the Delta method. [d]Pandemic controls for 2020-2023 have two parts: one captures heterogeneous level shifts across markets and the other captures interactions with unemployment rate. All the models above are estimated by OLS using market-category-specific revenue shares as weights within each year. Standard errors, clustered at the market, food category, and year levels, are listed in the parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Factor Regression Results for Market-category-year-specific Markups.

explanatory variables[a]	dependent variable: *ln(gross markup)*[b]	
	(1)	(2)
	short-term	long-run[c]
real interest rate	-4.08***	-4.93***
	(0.85)	(1.64)
real GDP per capita	0.70***	0.58**
	(0.16)	(0.23)
real housing price	0.29***	0.24***
	(0.07)	(0.09)
retailer HHI	0.07	0.03
	(0.09)	(0.06)
# UPCs per category	-0.12***	0.13***
	(0.03)	(0.05)
population	0.32	0.22
	(0.26)	(0.17)
COVID controls[d]	YES	
one lag of all variables	YES	
fixed effects of markets and goods	YES	
adj.R^2	0.63	
within R^2	0.40	
N	24,706[e]	

Note: [a]Explanatory variables are computed in natural log except for real interest rate and cumulative percent change in real housing price (both in decimal points). See details in the main text and Appendix C. [b]We calculate market-category-year specific markups based on equation (11); *gross markup* equals markup plus one by definition. [d]Pandemic controls for 2020-2023 have two parts: one captures heterogeneous level shifts across markets and the other captures interactions with real GDP per capita. [e]We trim off the upper and lower 1% of elasticity estimates in each year to avoid extreme values; we also drop the bottom 5% (in revenue share) of the rest elasticity estimates with the elasticity cutoff equal to 1.1. All the models above are estimated by OLS using market-category-specific revenue shares as weights within each year. Standard errors, clustered at the market, food category, and year levels, are listed in the parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

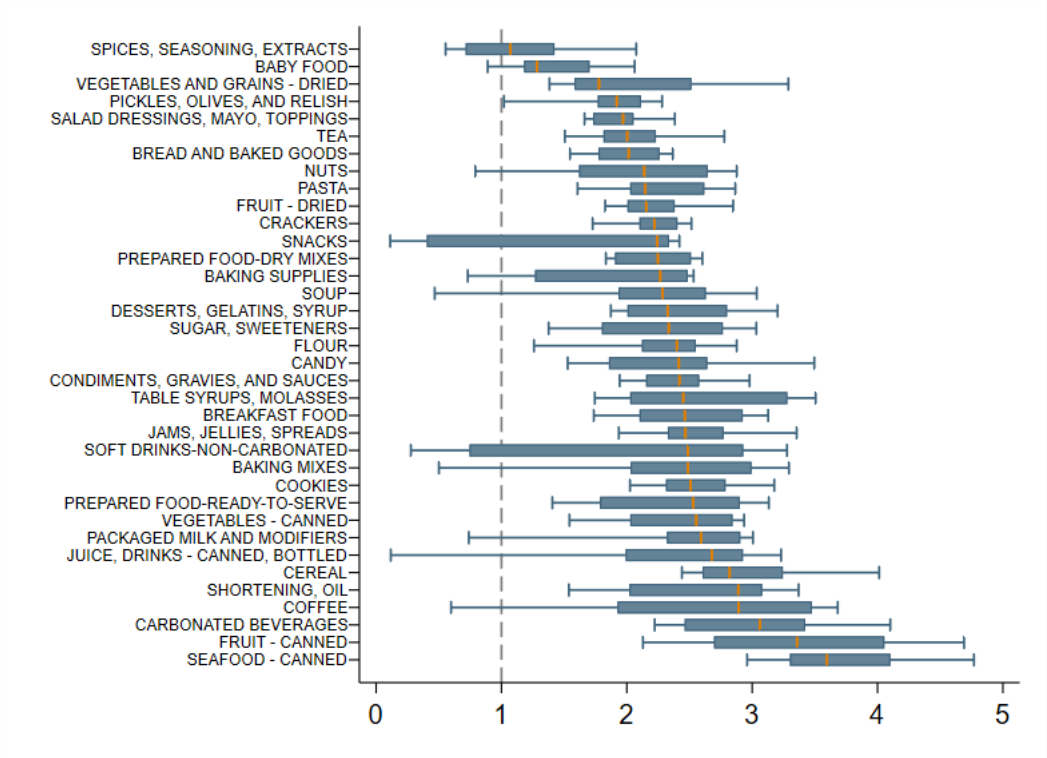


Figure B.1: NielsenIQ Demand Elasticity Estimates by Dry Food Categories in 2010.

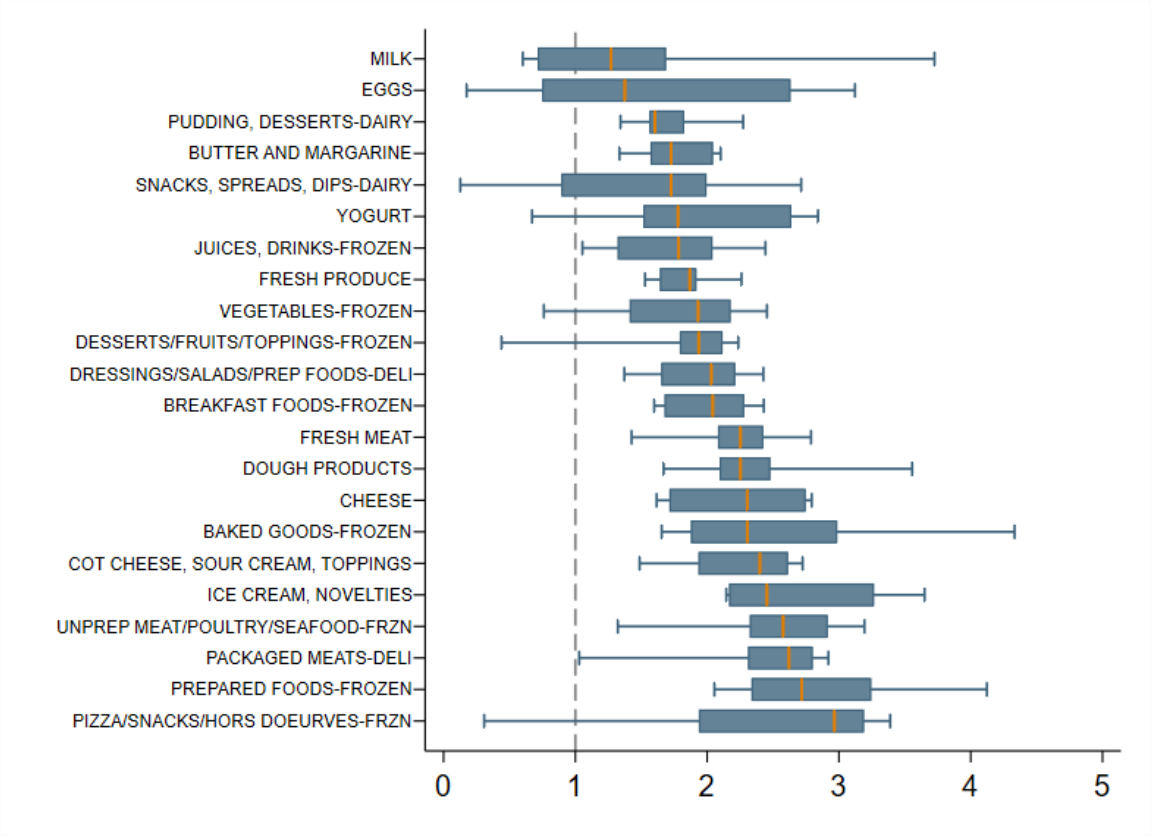


Figure B.2: NielsenIQ Demand Elasticity Estimates by Non-dry Food Categories in 2010.

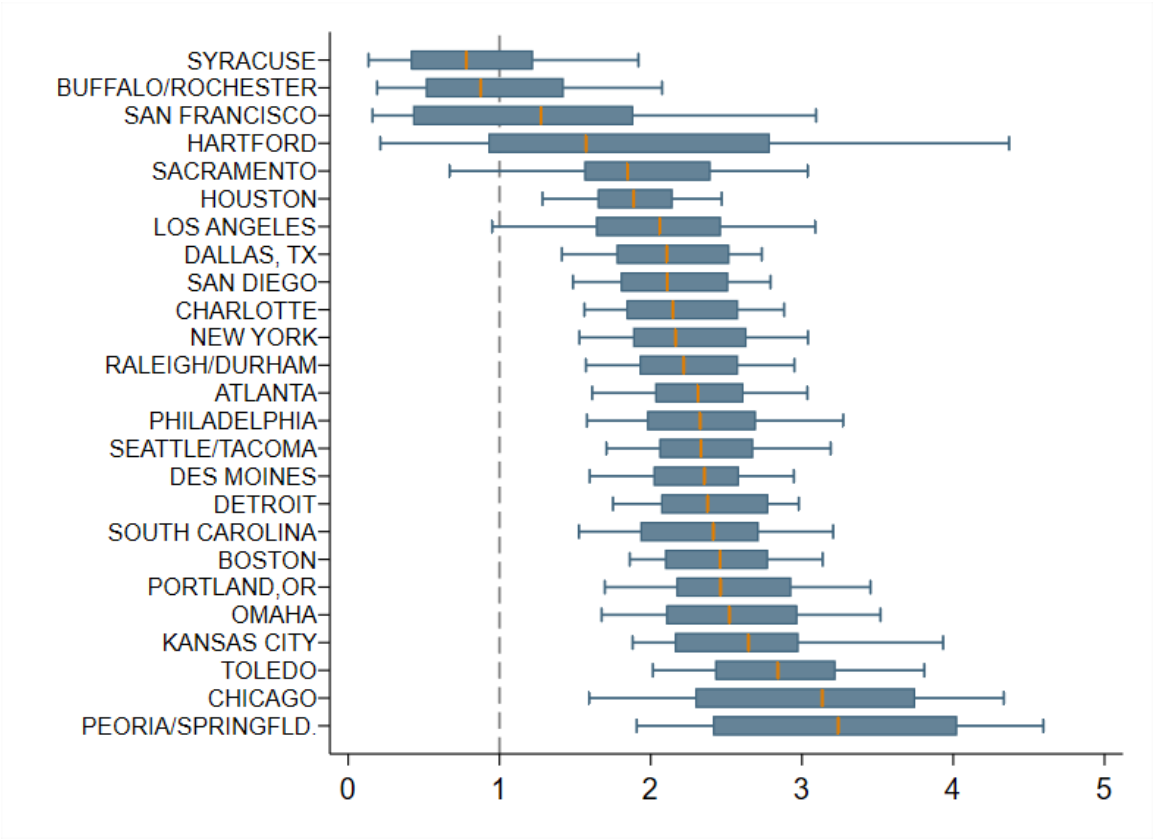


Figure B.3: NielsenIQ Demand Elasticity Estimates by Markets in 2010.

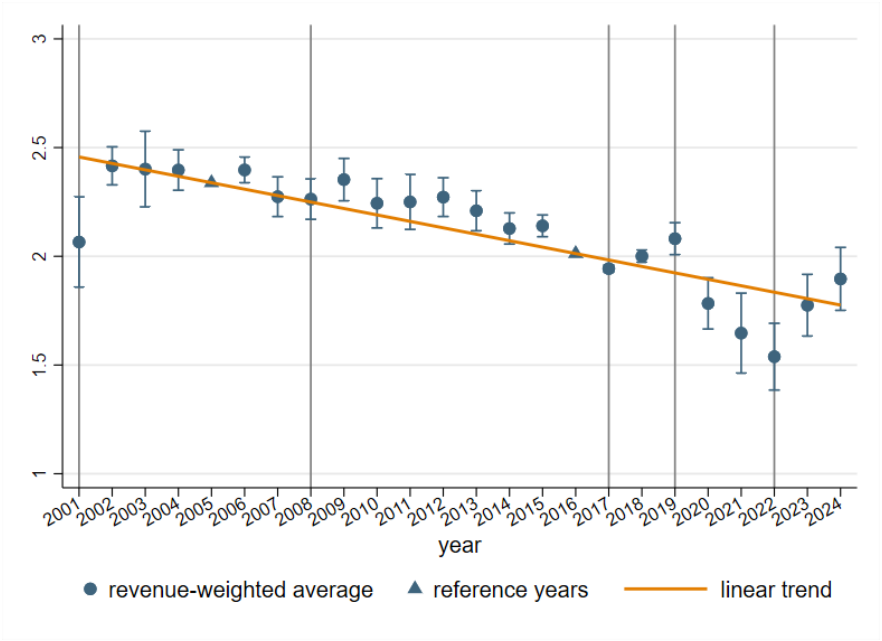


Figure B.4: Trend and Unconditional Cyclical Variation of Demand Elasticity Estimates.

Note: This figure shows trend and unconditional cyclical variation of demand elasticity estimates in the pooled IRI and NielsenIQ sample. Details of our estimation are discussed in the main text. The five vertical gray lines indicate aggregate demand shocks. The spikes show the 95% confidence intervals of the revenue-weighted average annual demand elasticities relative to the linear trend (2001-2024). Standard errors are clustered at the market, food category, and year levels (three-way clustering).

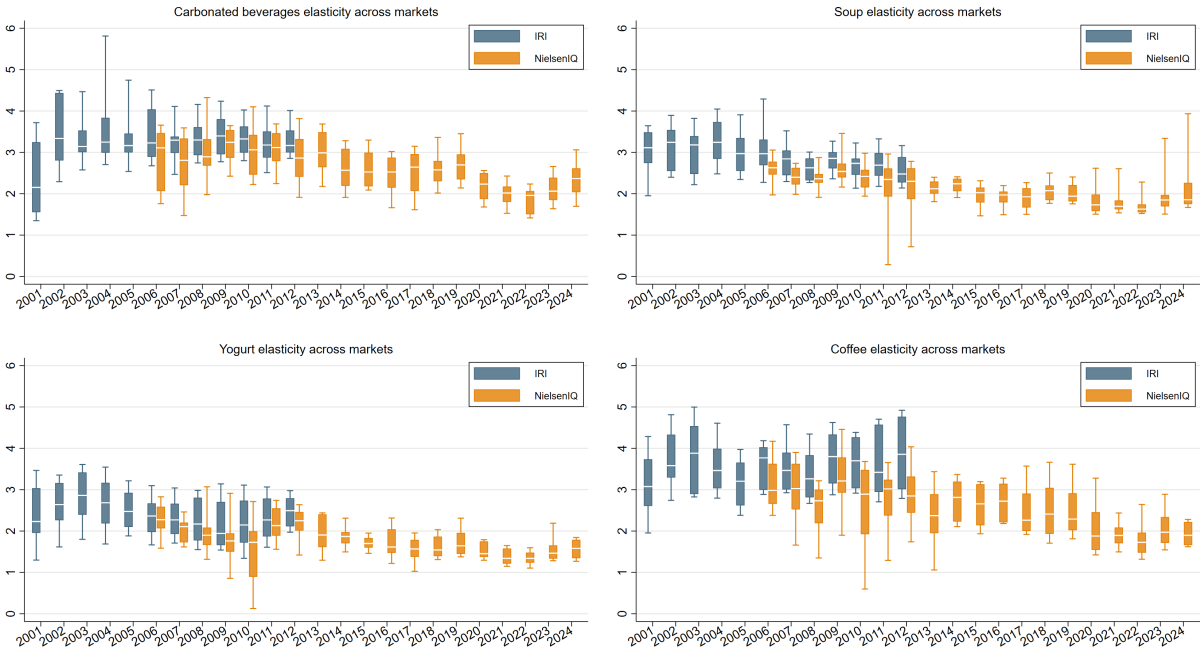


Figure B.5: Demand Elasticity Estimates of the Four Common Food Categories.

Note: IRI and NielsenIQ have different elasticity levels as they differ in store and UPC coverages in the data.

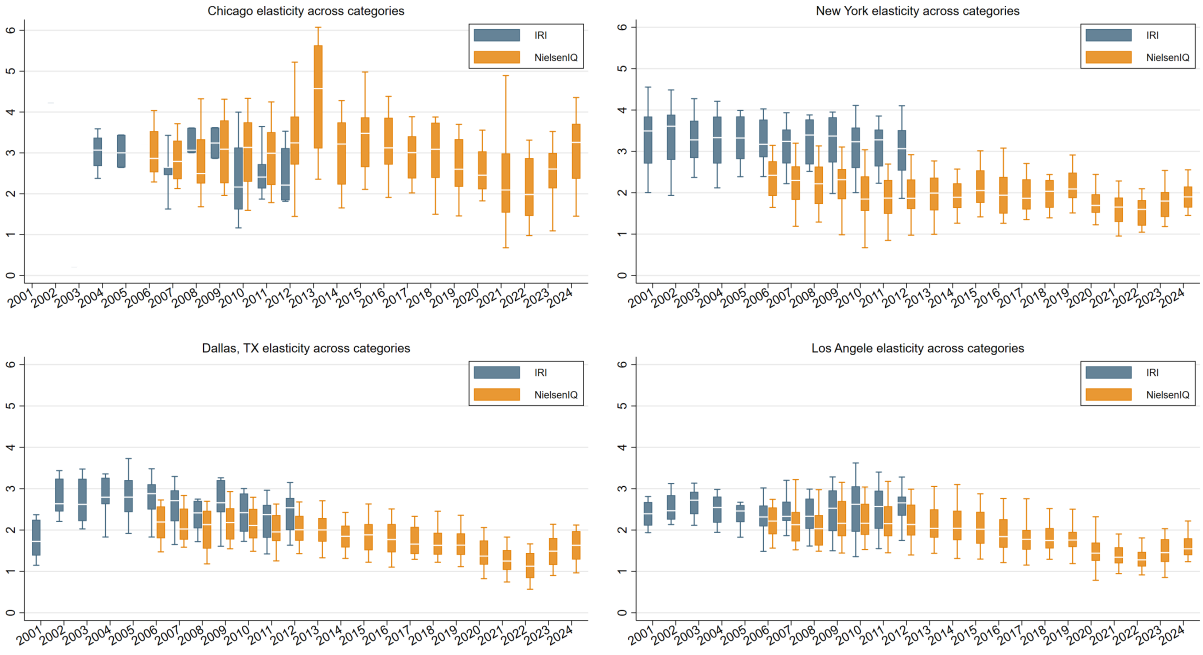


Figure B.6: Demand Elasticity Estimates in the Four Largest Regional Markets.

Note: There are limited elasticity estimates for Chicago in 2001-2003 and 2006 so that we do not have boxplots for those years in the figure.

C Data construction for market-year-specific factors

Several of our market-specific factor measures in the raw data, e.g., house prices, are nominal. We use the standard GDP deflator, measured at the national level, to take out the trend increases in prices. We reset its base year to 2001, the first year of our sample period 2001-2024.

Real GDP per capita We deflate the county-level nominal GDP to the base year 2001 using the GDP deflator. Then, we divide it by population to obtain the real GDP per capita for every county in each year. Finally, we take logarithm and then average using county-level populations as weights to obtain the market-level counterpart.

Unemployment rate The raw data is in percentage. We average them using county-level populations as weights to obtain the market-level measurement for each year.

Cumulative change in real housing price We start with the county-level housing price index that is calibrated using appraisal values and sales prices for mortgages bought or guaranteed by Fannie Mae and Freddie Mac (Bogin et al., 2019). We set the base year at 2001, the same base year for our GDP deflator, and deflate it using the GDP deflator. Then, we take differences relative to the base year for every county in each year, namely the county-level cumulative percentage changes relative to the base year. Finally, we average them using county-level populations as weights to obtain the market-level cumulative percentage change in real housing prices.

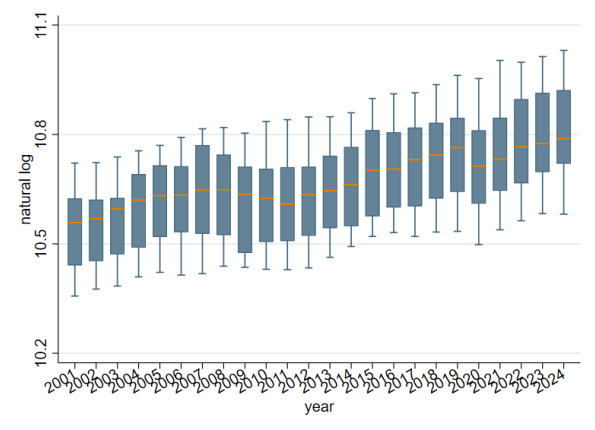
Population We aggregate annual county populations at the market level before taking logarithm.

Table C.1 shows the public sources of the original data while Figure C.1 gives their graphical representations. Other factors, including real interest rate, retailer-HHI, and UPCs per category, have been explained in Section 4.3. Figure C.2 shows their dynamics over the years.

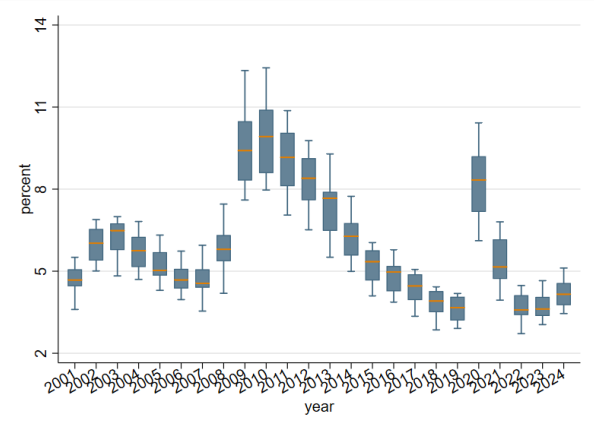
Table C.1: Public Sources of the Raw Data.

variable	level	source[a]
GDP deflator	national	Federal Reserve Bank of St. Louis
GDP	county	Federal Reserve Bank of St. Louis
housing price index	county	Federal Housing Finance Agency
unemployment rate	county	U.S. Bureau of Labor Statistics
population	county	U.S. Census Bureau

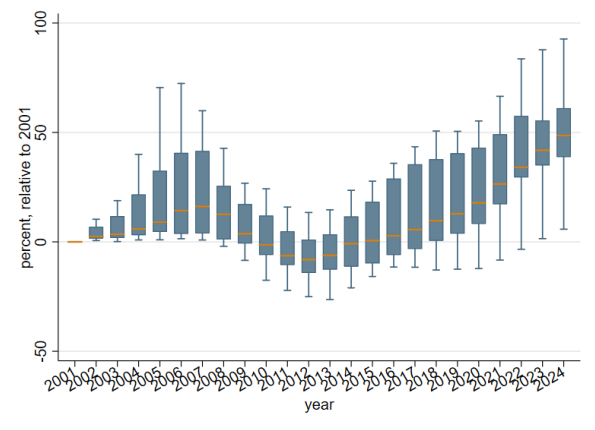
Note: [a]Here are website links below for all the data listed above.
(1) GDP deflator: "https://fred.stlouisfed.org/series/USAGDPDEFAISMEI".
(2) GDP: "https://fred.stlouisfed.org/release/tables?rid=397&eid=1054597".
(3) housing price index: "https://www.fhfa.gov/data/hpi/datasets?tab=annual-data".
(4) unemployment rate: "https://www.bls.gov/lau/tables.htmcntyaa".
(5) population: "https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates.html" for 2001-2019 & "https://www.census.gov/data/datasets/time-series/demo/popest/2020s-counties-total.html" for 2020 onward.



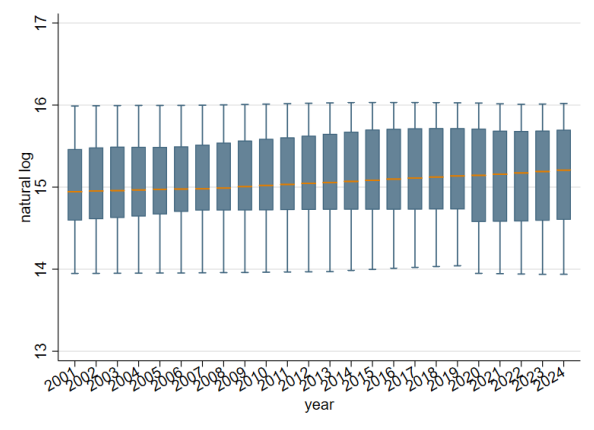
(a) real GDP per capita



(b) unemployment rate



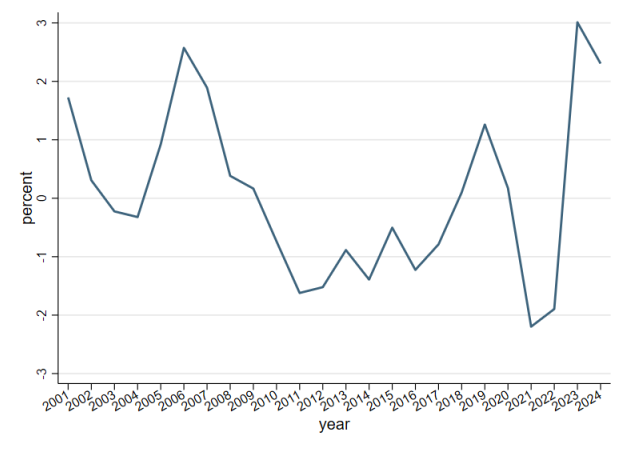
(c) cumulative change in real housing price



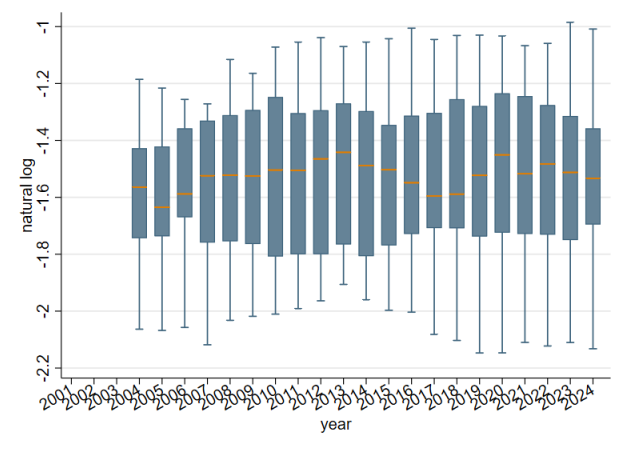
(d) population

Figure C.1: Cross-sectional and Time Variations of Market-specific Factors.

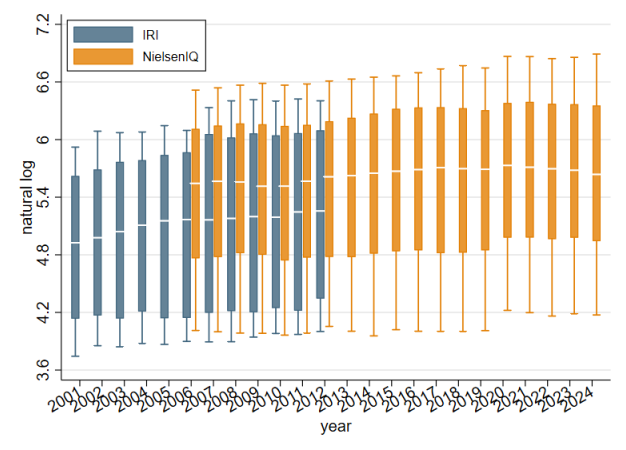
Note: See the text above for the constructions and descriptions of these market-level factors.



(a) One-year real interest rate



(b) Retailer-HHI



(c) UPCs per category

Figure C.2: Cross-sectional and Time Variations of Other Factors.

Note: The national one-year real interest rate is obtained from FRED. The market-level retailer-HHI is calculated from the NielsenIQ consumer panel that starts from 2004. We calculate the revenue-weighted average UPCs per category across stores from both IRI and NielsenIQ retail scanner data sets. This factor is at the market-category level. For the latter two factors, we show boxplots as in Figure C.1.