Secular rise and pro-cyclical variation in markups: Evidence from US grocery stores*

Bulat Gafarov, Tengda Gong, Jens Hilscher

June 13, 2023

Abstract

This paper documents substantial time variation in price elasticities of demand and therefore markups. We propose a two-step procedure to identify time-varying markups. Using the scanner data of US grocery stores from 2001 to 2020 we first estimate elasticities at the market-good-year level. We then efficiently aggregate these data by year to estimate a common trend and cyclical variation in elasticities and impute markups from there. We estimate (i) a secular increase in U.S. grocery store markups of 3.9% per year over the sample period and (ii) an average 13.6% cyclical decline at times of aggregate demand contractions. Our results imply pro-cyclical changes in markups. Across markets, elasticities vary with market-wide factors that we expect to influence preferences and market structure—real GDP, unemployment and market concentration.

*All authors are from the Department of Agricultural and Resource Economics at UC Davis. Contacts: bgafarov@ucdavis.edu, tdgong@ucdavis.edu, jhilscher@ucdavis.edu. We would like to thank Saroj Bhattarai, Jim Chalfant, Oscar Jorda, and Sanjay Singh for helpful comments and Hovhannes Mnatsakanyan for research assistance. Research for this paper was supported by a grant from the Giannini Foundation of Agricultural Economics and the USDA National Institute of Food and Agriculture, Hatch project S-1072: U.S. Agricultural Trade and Policy in a Dynamic Global Market Environment.

1 Introduction

Competitive markets are one of the main reasons that consumers are protected from companies taking advantage of their market power. Free entry into markets ensures that the adverse effects of excess profits resulting from oversized markups are limited. However, in a recent book, Philippon (2019) finds that there are many sectors in the U.S. with high markups. Such high markups are detrimental to consumer welfare, but they can also cause reductions in investment and productivity growth. Based on a production side model, De Loecker et al. (2020) document that, since the 1980s, there has been a striking increase in markups in the U.S. for upstream production.

A necessary condition for such markups is downward-sloping demand curves. The relevant measure is the own-price elasticity of demand—the extent to which demand decreases when price increases. When consumers are sensitive to price changes, elasticities will be high. However, if elasticities are low, firms can charge higher markups because consumers are slower to react to price changes. The own-price elasticity of demand that we can estimate using retail scanner data is the one that a specific store faces. This elasticity will be determined by both characteristics of the average consumer and consumers' alternative choice sets and ability to switch—consumers can buy other goods in the same store or purchase goods in other stores; that outside option can also affect the elasticity.[1] An analysis of markups, market power and consumer welfare, following Lerner (1934) and Elzinga and Mills (2011), therefore starts with an estimation of the own-price elasticity of demand.

It is well-documented that there are differences in elasticities across goods and that they vary across markets (DellaVigna and Gentzkow, 2019), due to, among other factors, income effects. However, what previous work has assumed is that elasticities are constant over time (DellaVigna and Gentzkow, 2019) or that there are low-frequency changes (De Loecker et al., 2020). However, one might expect that, for example, the great recession and the global pandemic affected consumer preferences and their behavior in response to price changes.

In this paper we take as a starting point the possibility that elasticities may change over time. We assume that there may be a secular trend in elasticities but also allow for higher-frequency cyclical variation. We use the food retail sector in the U.S. as a setting to study the determinants of elasticities over time and across markets. There are several benefits of studying this sector. It is a large sector, for many households expenditure on food represents a significant fraction of discretionary spending (Cox and Harris-Lagoudakis, 2022) and therefore changes in food markups affect a large share of the population. Moreover, recent evidence suggests that consumers form their overall inflation expectations based on grocery bills (D'Acunto et al., 2021). The food retail sector also covers a wide variety of geographical

[1] Consumer characteristics will, among other factors, depend on their preferences, income, and time allocated to shopping. Separately, the demographics of actual shoppers may change over time.

locations, and there are many different goods, most of which consumers purchase frequently.

We propose a two-step procedure to estimate time-varying elasticities and markups. In the first step, we use the well-known Hausman (1996) IV strategy to estimate own-price elasticities of demand.[2] We choose 26 large markets in the U.S. 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. For each market, we estimate good-specific own-price elasticities of demand by year. We pool all the available items at the bar code (universal product code—UPC) level within each specific good category.[3] This approach leads to accurate market-good-year elasticities.

We use weekly data and include week fixed effects to control for demand shocks. That same fixed effect also captures demand shocks that may result from substitution from other products due to changes, for example, in their prices. What we do not capture and do not want to capture is within good substitution. Our idea is that there is an average price for items that are members of a specific good group and that we capture the average elasticity of products in that good group using our estimation strategy.

The scanner data we use comes from two sources—IRI (2001-2012) and Nielsen (2006-2020). Compared to the IRI data set, Nielsen covers both more goods and more stores in each market. For both data sets, there are sufficient observations to estimate demand elasticities for each market-good-year pair and the estimation strategy produces precise and realistic results. Specifically, we find that (i) the Hausman IV strategy works—a test for weak instruments rejects the null for about 94% of elasticity estimations; (ii) estimates are precise under strong IV—more than 95% of their standard errors are below 0.35 and more than 99% of their t-statistics are above 1.96; and (iii) estimates are reasonable—for only approximately 5% of the estimates we can reject the hypothesis of the elasticity lying above one at the 5% significance level.[4] We find large variation across markets, goods, and, importantly, over time, validating our initial assumption of time-varying elasticities.[5]

In the second step of the estimation, we then pool elasticity estimates across markets and goods to isolate variation in elasticities over time—both lower-frequency trend and higher-frequency cyclical variation. We find a pronounced downward trend in elasticities. Using the

[2]The main other methods to estimate elasticities use demand control variables with high-frequency data (Levin et al., 2017; Brand, 2021) or impose covariance restrictions on supply and demand shocks (Döpper et al., 2022; 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 (2016).

[3]Recent studies by Hitsch et al. (2019) and Chernozhukov et al. (2019) highlighted the importance of pooling elasticity estimates at the UPC level to a good-category level to reduce noise. To regularize the UPC-level price elasticity, the former paper proposed a Bayesian-hierarchical approach and the latter used a ridge regression approach. Since we are not interested in the UPC-level estimates, but in the category-level trends, we directly pool all UPCs within a category when estimating the category-level elasticity.

[4]This may be due to statistical errors given the expected 5% of false positives at this hypothesis testing.

[5]De Loecker and Scott (2016) also finds time-varying own-price elasticities for the beer market. Their paper combines production data and retail sector data.

standard transformation (e.g., DellaVigna and Gentzkow, 2019), we convert our estimates of locally-linear demands to implied markups. The downward trend in elasticities thus implied an upward trend in markups. Using 2001 as the base year, our results imply that average markups—across all markets and goods—went up 45% by 2019 and 100% by 2020. The slow-moving trend in markups has also been identified by other studies measuring low-frequency movements in markups. Philippon (2019) considers all industries in the U.S. De Loecker et al. (2020), using a production side approach, measures markups every five years. They also find a large increase in markups between 1997 and 2012, consistent with our findings. Neither study considers higher-frequency cyclical variation, though.

The second time series pattern we identify is an important effect of the business cycle, specifically large shocks to aggregate demand. It has been a long debate in macroeconomics whether markups are counter-cyclical (as predicted by sticky price models) or pro-cyclical (see a discussion of this literature in Nekarda and Ramey (2020)).[6] Since we estimate elasticities year by year, our approach allows us to capture cyclical variation. Our data extends from 2001 to 2020 and includes two substantial contractions—the 2001 dot-com recession and the great recession of 2008, as well as a contraction of demand due to the tightening monetary policy in 2017. Our data set also includes the recession corresponding to the global pandemic year of 2020. However, as a result of aggressive monetary and fiscal stimuli and potential changes in preferences for online shopping, some aggregate demand measures increased during the pandemic.[7] We find that, during contractions of aggregate demand, elasticities increase, while they decrease when aggregate demand expands. Markups are therefore pro-cyclical.

We calculate standard errors of the time variation in elasticities by using weighted least squares to precisely weight market-good-year elasticity estimates from the first step of the estimation. We find that the increase in average elasticities in 2002, 2009 and 2018 as well as the decline in elasticity during the pandemic in 2020 are all individually statistically significant. Taken together, our findings thus document both trend and cyclical variation in elasticities across markets and good categories. Two closely related papers also consider annual elasticities. Brand (2021) considers a subset of nine food categories and uses a methodology that results in low precision. In contrast, we include the bulk of categories sold in the food retail sector. Döpper et al. (2022) use the Berry et al. (1995) method to estimate US-level pooled estimates of elasticities for 133 categories available in the Nielsen scanner data, including many non-food categories. They also end up with fairly high standard errors; drawing conclusions from year-by-year variation in elasticities is therefore difficult. Neither paper uses an IV approach to estimate elasticities. As a result, both may underestimate elasticities and overestimate markups.

[6] A macroeconomic indicator is called pro-cyclical if it moves in the same direction as the GDP gap between the actual GDP and the corresponding trend level.

[7] See consumers' shifting to online shopping in Harris-Lagoudakis (2023). The risk of getting COVID may also make in-store consumers less price sensitive as they may stop shopping around to reduce that risk.

We next proceed to explain the time variation in market-good-year elasticities. To ensure that our results are not driven by market or good-specific effects, we demean elasticities at the market-good level. Having already identified the common cyclical and trend variation in elasticities, we include time fixed effects in the regression. We therefore estimate effects of market-specific factors from time-varying cross-sectional heterogeneity in elasticities.

When explaining variation in elasticities we use measures capturing both factors affecting demand and those affecting market structure. We find three effects. First, we identify a negative effect of household income, measured by per capita GDP. That is, when people's income declines, for example because of a more severe recession than experienced by other markets, demand elasticity goes up. Aguiar et al. (2013) also find that consumers become more price sensitive in response to income losses, for example, by shopping around more.

The second effect is that population increase leads to lower elasticity. An increase in population might be a sign of improved economic performance or anticipated income growth, potentially making consumers less price sensitive. However, higher market concentrations (fewer stores per capita) do not lower elasticities, which is consistent with Dong et al. (2023).

Third, we find that higher housing prices result in higher elasticities. If we use housing prices as a proxy for rent, then a higher house price may result in more price-conscious consumers (Stroebel and Vavra, 2019).

We find that these effects, plus a few other market-level factors (unemployment and dependency ratio), can explain a large share of the common time variation in elasticities and therefore markups. The two time effects that the model misses are the decrease in markups after the dot-com bust and the sharp increase in markups during the pandemic. The latter is not surprising given the large shifts in shopping behavior during the lockdowns in 2020.[8] Our results suggest that in the food retail sector, markups are driven to a large extent by the growing income of customers rather than by the concentration of firms.

Our paper is related to several strands of the literature. Recent studies use the Berry et al. (1995) method (BLP) to calculate elasticities, including Brand (2021) and Döpper et al. (2022) as well as MacKay and Miller (2023). In contrast, our focus is narrower than BLP, a comprehensive structural approach that explicitly incorporates product and consumer characteristics and allows for counterfactual analyses in addition to the recovery of elasticities. Since our aim is to estimate demand elasticities with respect to the own price, we can use a panel regression model with fixed effects. Controlling for demand shocks from substitute goods using fixed effects makes elasticity estimations on a larger scale more tractable.

As mentioned above, the overall pattern of increased markups has been studied in the existing literature. In addition to Philippon (2019) and De Loecker et al. (2020), De Loecker and Scott (2016) study the beer industry and show that alternative methods such as BLP and production side methods give similar results for beer.

[8] The former might be due to the fact that only the relatively smaller IRI sample covers earlier years.

The remainder of the paper is organized as follows. Section 2 discusses the retail scanner data that we use, outlines some of the data choices we make, and presents summary statistics. Section 3 presents our panel-IV regression model and discusses the associated results of market-good-year elasticity estimates. This section also pools estimates by year to show the time variation in elasticities and markups. Section 4 analyzes market-specific factors driving the dynamics in elasticities. Section 5 discusses the implications of these results and the merits of our panel-IV approach. Section 6 concludes the paper.

2 Retail scanner data

We use the retail scanner data from the IRI Marketing Dataset (Bronnenberg et al., 2008) and NielsenIQ Datasets to estimate own-price elasticities of demand for food across grocery stores.[9] The IRI retail scanner data covers 12 years, from 2001 to 2012. The Nielsen scanner data spans 15 years, from 2006 to 2020. They both have weekly transaction records of products sold by retail stores located in physical markets across the U.S. In the following, we describe key features of these rich data that are relevant to our empirical estimation strategy.

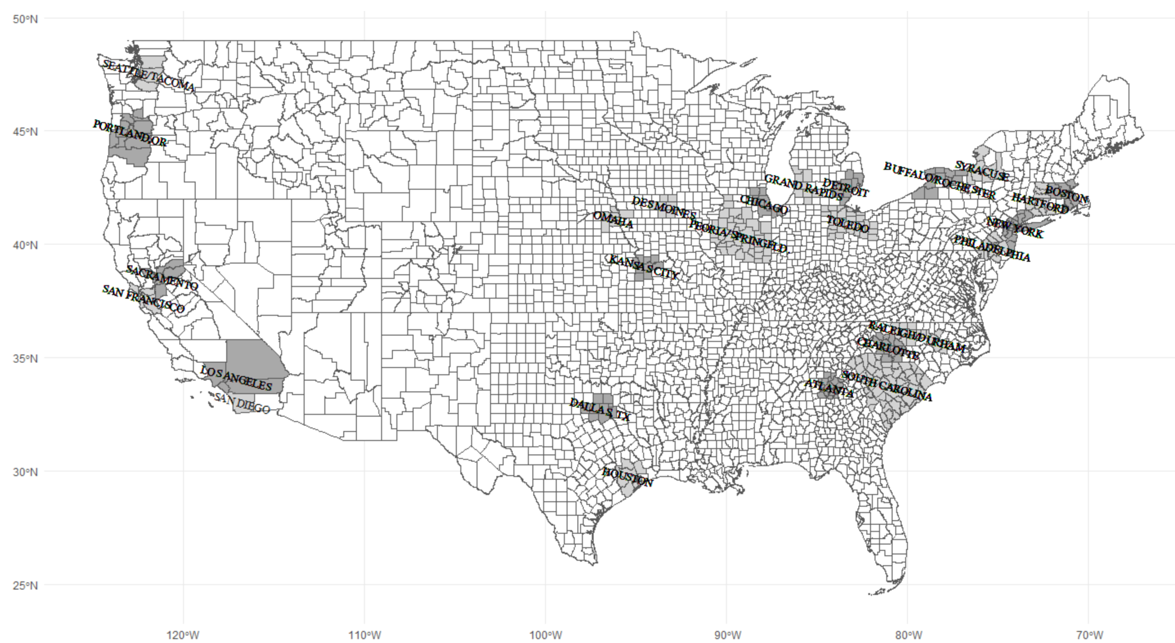


Figure 1. The 12 Pairs of 26 Markets in the US Mainland.

Note: Each pair of neighboring markets are colored in gray and light gray. See detailed pairs in Table 1.

[9]The IRI data was purchased by the authors; the NielsenIQ data is obtained from the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Table 1. The 12 Pairs of 26 Markets.

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

Note: [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 26 markets in total for demand elasticity estimations.

Markets We group markets, defined in the IRI scanner data, into neighboring pairs.[10] 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). As shown above in Figure 1 and Table 1, we strategically select 12 pairs of 26 markets that are spread out across major US regions.[11] Table 2 below shows that these markets have 125.5 million residents in 2010, over 40% of the 2010 US resident population.[12] Also, they contribute to around 70% of observations in the IRI scanner data.[13] Hence, these selected markets are representative of both population and data.

[10]Each IRI market consists of multiple adjacent counties. We use the unique federal county code for each Nielsen grocery store, FIPS, to pin down the IRI markets to which they belong.

[11]We have 50 IRI markets in total. The other 24 markets, which are not used in this paper, are mostly small in terms of the number of reported stores or not easy to find neighboring markets for pairing.

[12]The 2010 Census shows that the resident population of the United States was about 308.7 million.

[13]They also contribute to a large but smaller proportion of observations in the Nielsen scanner data.

Table 2. Summary Statistics of Markets.

market name	population in 2010 (million)	IRI (2001-2012) obs per year (million)	stores per year	Nielsen (2006-2020) obs per year (million)	stores per year
DES MOINES	0.7	1.1	9.1	19.9	31.0
OMAHA	1.1	1.6	14.3	27.0	44.8
SYRACUSE	1.2	2.2	22.3	8.5	17.5
GRAND RAPIDS	1.7	1.7	15.0	6.6	7.0[a]
KANSAS CITY	2.0	2.9	21.8	18.8	23.6
PEORIA/SPRINGFLD.	2.0	1.5	14.6	25.3	42.5
TOLEDO	2.0	1.6	14.0	30.7	44.2
BUFFALO/ROCHESTER	2.5	2.4	21.6	14.9	32.4
CHARLOTTE	2.6	3.1	35.5	109.9	227.9
SACRAMENTO	2.8	2.7	26.9	53.8	98.9
SAN DIEGO	3.1	3.1	29.8	74.4	129.8
PORTLAND,OR	3.2	3.3	33.2	91.4	157.0
HARTFORD	3.2	3.4	28.7	27.8	51.2
RALEIGH/DURHAM	3.3	3.8	41.2	133.6	267.3
SEATTLE/TACOMA	3.4	4.1	42.3	123.4	211.1
DETROIT	4.8	2.8	25.4	79.9	109.3
ATLANTA	4.9	3.6	31.6	100.4	154.4
SOUTH CAROLINA	5.1	4.6	60.0	140.5	301.3
BOSTON	5.5	5.2	41.1	119.6	187.1
HOUSTON	5.9	3.7	37.8	106.8	170.9
SAN FRANCISCO	6.1	3.6	39.6	126.9	229.1
DALLAS, TX	6.2	4.6	51.0	130.0	238.3
PHILADELPHIA	6.5	5.4	47.3	118.0	206.9
CHICAGO	9.0	4.8	41.7	178.8	272.0
LOS ANGELES	17.1	9.2	94.8	326.9	610.1
NEW YORK	19.5	10.5	101.5	232.2	441.8
total	125.5	96.5	941.7	2426.1	4307.5

Note: [a]Part of the reason for the limited number of stores per year in GRAND RAPIDS is that 5 out of 39 counties in this market have no store records in the Nielsen scanner data.

Table 3. Summary Statistics of Food Categories.

	IRI categories[a] mean	p10	p90	Nielsen categories[b] mean	p10	p90
obs per year (million)	6.0	1.6	12.5	40.4	6.9	86.9
No. of UPCs per year	2096.8	304.2	4784.5	4412.3	619.0	11164.5
No. of UPCs per year-market	540.6	121.7	1059.5	1052.1	179.4	2324.7

Note: [a]IRI has 16 food categories. [b]Nielsen has 60 food categories. See the list of IRI and Nielsen food categories in Table A.1.

Food categories We estimate food demand elasticities at the category level rather than at the lower UPC level. As shown above in Table 3, different markets sell quite different goods in terms of UPC.[14] But they always sell common some goods of the same categories. We take the 16 food categories defined in the IRI sample as given. In the Nielsen sample, we regard the 60 food groups as the categories of interest.[15] See the complete list of food categories in Table A.1 in Appendix A.

Stores We keep all stores in our sample as they have sufficient observations each year as shown below in Table 4. Also, most stores have over 40 weeks of sales transactions per year recorded in the data. Moreover, each store sells almost all categories of food groceries every year. For each food category, they have weekly sales records of various food products at the UPC level. All these features allow us to use flexible fixed effects to control demand shocks when estimating price elasticities of demand (see details in Section 3.1).

Table 4. Summary Statistics of Stores.

	IRI stores			Nielsen stores		
	mean	p10	p90	mean	p10	p90
estimated revenue per year (million)[a]	29.5	13.3	50.3	19.0	6.7	33.9
obs per year (thousand)	90.4	46.0	134.4	536.0	283.9	787.2
No. of weeks per year	43.8	31.0	51.8	50.1	45.9	52.1[b]
No. of categories per year	16.0	16.0	16.0	58.8[c]	59.0	60.0
No. of UPCs per year-category	202.0	141.3	266.2	327.9	210.4	454.7

Note: [a]All store revenues are measured in the 2015 US dollar. IRI directly provides annual estimates of store revenues while Nielsen does not. We estimate annual revenues for each Nielsen store by aggregating its reported revenues across all products in the data. However, some stores do not report sales for some weeks in a year. So, our revenue estimates for Nielsen stores should be taken as a lower bound of their actual annual revenues. [b]A Nielsen week starts on Sunday and ends on Saturday. Hence, some Nielsen stores have sales records of 53 weeks in years like 2011 and 2016. [c]About 1.3% of Nielsen stores have sales records of less than 36 food categories, which drives down the average number of categories across stores.

3 Demand elasticity estimation

3.1 A panel IV-regression approach

We estimate demand elasticities at the market-good-year level. Specifically, for each market m, product category c, and year t, we run regression (1) below to obtain market-good-year specific own-price elasticities of demand in the IRI and Nielsen samples, respectively:

$$log(q_{v,s,w}) = -e_{m,c,t}log(p_{v,s,w}) + upc_v + store_s + week_w + \varepsilon_{v,s,w}, \tag{1}$$

[14]Nevertheless, goods sold in neighboring markets are mostly the same. This allows us to instrument good prices in a market of interest by those in neighboring market(s) when estimating demand elasticities.

[15]For comparability, we do not take the lower product module in the Nielsen sample as the product category since it has far fewer UPCs than a product category in the IRI sample does.

where $q_{v,s,w}$ and $p_{v,s,w}$ denote the quantity and (average) price of the product variety v (identified by the product UPC within each category c) sold by store s in week w; upc_v, $store_s$, and $week_w$ are fixed effects; and $\varepsilon_{v,s,w}$ is the error term.[16] The coefficient of interest $e_{m,c,t}$ is the average own-price elasticity of demand for product category c facing stores in market m and year t. This estimation approach allows for heterogeneity in demand elasticities across markets, goods (product categories) and, importantly, over time.

The UPC and store fixed effects above control for slow changing (approximately time-invariant within a year) factors like consumers' preferences over specific products and stores. The time fixed effects, on the other hand, control for weekly demand shocks like holiday needs.[17] They also absorb the impacts of weekly price changes among substitute and complementary products. The error term $\varepsilon_{v,s,w}$ is clustered at store and week levels in two ways to allow for arbitrary correlations caused by the other unobservables within each store and week that are not captured by fixed effects.

It is widely recognized that product price is endogenous as it is determined together with product quantity through the equalization of product demand and supply in the market equilibrium. The fixed effects above, however, can not capture the store-product-week variant demand shocks that may simultaneously drive the quantity and price of a product sold by a store in a week. To tackle this endogeneity problem, we follow Hausman (1996) and 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).[18] Note that product UPCs are manufacturer-specific, i.e., a product with the same UPC sold in the paired markets comes from the same manufacturer, which delivers the relevance of our price instrument. As shown below in Section 3.2, this Hausman-type price IV is strong for most of our demand elasticity estimates.

3.2 Estimates of market-good-year elasticities

When implementing our estimation strategy outlined above, we find the proposed Hausman-type IV for most weekly product prices at the UPC level across markets and years. This is due to the fact that products sold in neighboring markets are almost the same each year. In total, we have about 94% and 93% of price observations successfully matched with their IVs in the IRI and Nielsen samples, respectively.[19] These market-good-year specific IV regressions

[16] Following the literature, we impute average product prices from their revenues and quantities.

[17] The time fixed effects also control for weekly supply shocks like the seasonality in agricultural production.

[18] Our implementation of Hausman IV strategy is different from DellaVigna and Gentzkow (2019) who used average price for a given UPC across the entire nationwide store chain. We believe, that by focusing on geographically close markets we better capture common local cost shocks.

[19] In the IRI sample, 1.09 billion out of 1.16 billion price observations have their paired IVs. In the Nielsen sample, 33.9 billion out of 36.4 billion price observations have their paired IVs.

deliver us 27,531 raw demand elasticity estimates.[20] About 94% of them are obtained under the strong price IV with a similar percentage within the IRI or Nielsen sample.[21]

As shown above in Section 3.1, in this paper we use the term demand elasticity as a shorthand for negative demand elasticity for convenience. Among the strong demand elasticity estimates, less than 1% are negative possibly due to the unavoidable estimation bias in a typical IV regression. We drop these noisy estimates and those that are not obtained under the strong price IV for all the following analyses. Additionally, we trim off lower and upper 1% of the remaining elasticity estimates each year to get rid of the extreme values that might contaminate our key findings about elasticity dynamics. After cleaning, we have about 95% and 90% of raw demand elasticity estimates left in the IRI and Nielsen samples, respectively. Our final pooled sample has 25,073 demand elasticity estimates with 19% in the IRI sample and 81% in the Nielsen sample.[22]

Table 5 below shows that the average demand elasticity estimate in the IRI sample is higher than that in the Nielsen sample. This is largely due to the downward trend in demand elasticities, as documented below. The IRI sample ends in 2012 while the Nielsen sample ends in 2020. The fact that a Nielsen food category has relatively more substitutable products at the lower UPC level, as shown above in Tables 3 and 4, may also contribute to their differences in levels. Leung and Li (2022) shows that product varieties increase over time, making consumers more likely to have one-stop shopping and thus less price sensitive.

Table 5. Summary Statistics of Cleaned Demand Elasticities.

	IRI sample[a]			Nielsen sample[b]		
	mean	p10	p90	mean	p10	p90
point estimate	2.46	1.70	3.25	1.65	0.94[c]	2.39
standard error	0.16	0.08	0.29	0.12	0.05	0.24
t statistics	18.64	8.58	30.42	19.32	6.57	34.4
obs in an estimation (thousands)	225.6	37.3	492.8	1616.7	111.9	3975.7
Cragg-Donald F statistics	176.9	18.9	439.9	1168.5	47.7	2193.3

Note: [a]The IRI sample of 2001-2012 has 4,735 cleaned demand elasticity estimates. [b]The Nielsen sample of 2006-2020 has 20,338 cleaned demand elasticity estimates. We clean raw demand elasticity estimates by dropping those that have unreasonable values and those that are not obtained under the strong price IV. See detailed descriptions in the text above. [c]The Nielsen sample has notably more demand elasticity estimates below 1 than the IRI sample as the former has price-inelastic goods like baby food, ice, and eggs. The downward trend in demand elasticities over time and others, as explained in the text above, also contribute to this. In total, however, only about 5% of our demand elasticity estimates are significantly below 1 at the 5% statistical significance level.

[20]In the IRI sample (2001-2012), our elasticity estimates are balanced in time across market-good pairs. In the Nielsen sample (2006-2020), about 5% of market-good pairs have elasticity estimates of less than 15 years. Many of them are concentrated in the market of GRAND RAPIDS and in the good category of YEAST.

[21]We follow the traditional rule of thumb for a linear IV regression with one endogenous variable (Staiger and Stock, 1997), setting 10 as the minimum Cragg-Donald F statistics required for a strong IV.

[22]Some market-good pairs have no elasticity estimates of some year(s) in the IRI or Nielsen sample.

Importantly, our demand elasticities are precisely estimated. Their standard errors are small such that t statistics are almost all above 1.96 (the traditional 5% significance threshold). This is because we have sufficient observations and strong price IVs in nearly every elasticity regression, as reported above in Table 5. Moreover, our demand elasticity estimates are spread almost evenly across all markets, goods, and years. These nice features give us confidence in later regression analyses that use these precise and representative demand elasticities as the dependent variable.

Our demand elasticity estimates also exhibit notable heterogeneity across markets and goods, which is in line with the empirical literature (DellaVigna and Gentzkow, 2019). For example, Figures 2 and 3 below show that IRI demand elasticity estimates generally range from 1 to 4 across food categories and markets in 2010. Within each food category (market), these elasticity estimates also have sizable differences across markets (food categories). Nielsen demand elasticity estimates have similar features, as shown by Figures A.1, A.2, and A.3 in the Appendix. We will tease out these notable level differences among market-good-year specific demand elasticity estimates through demeaning when we study their time variation later on.

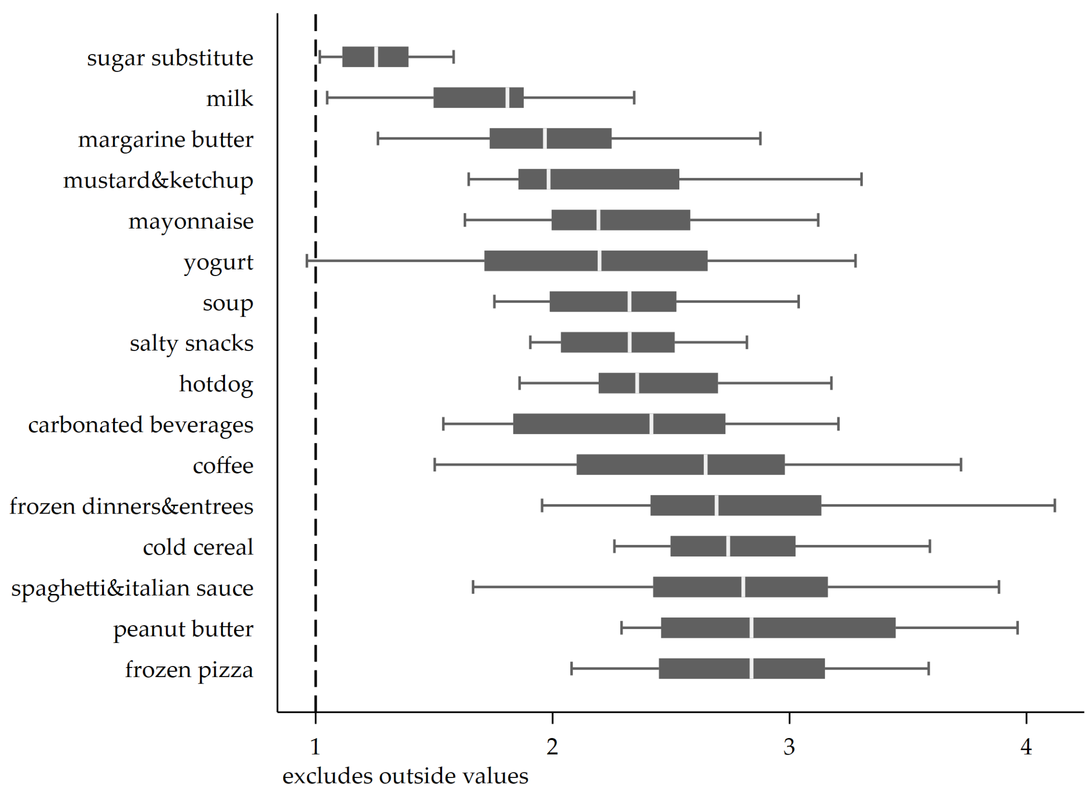


Figure 2. IRI Demand Elasticity Estimates by Food Categories in 2010.

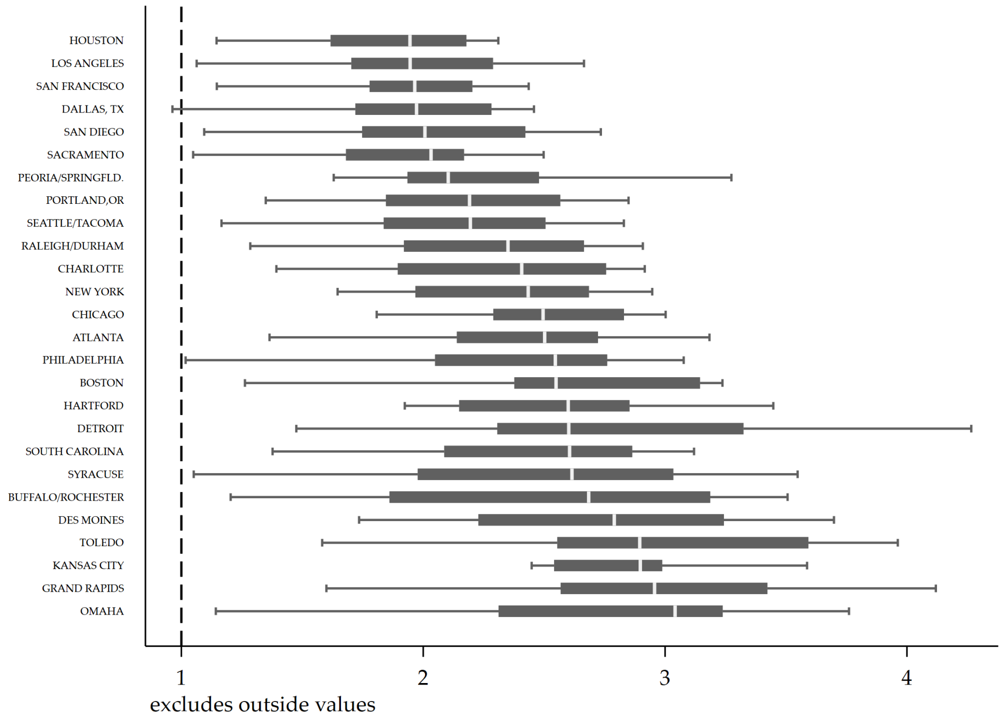


Figure 3. IRI Demand Elasticity Estimates by Markets in 2010.

3.3 Cyclical and trend variation of elasticities

We now proceed to the second step of our estimation. In order to identify time variation in average elasticities we efficiently pool the IRI and Nielsen market-good-year estimates. We next analyze the patterns in the dynamics of elasticities across goods and markets. Then, we discuss markup implications for the food retail sector in Section 3.4. In Section 4, we study macroeconomic and demographic factors that drive demand elasticity dynamics.

We begin by summarizing changes in the elasticity estimates over time. Each year, we calculate the mean, tenth and ninetieth percentiles of the full distribution of estimated market-good elasticities. Figure 4 below reports these three time series.[23] There are three patterns that we immediately notice. First, there is a pronounced downward trend in demand elasticities over the sample period. Mean elasticities decline from 2.16 (peak in 2002) to 1.31.

Second, at times of contraction in aggregate demand, elasticities increase, implying counter-cyclical elasticities.[24] Section 4.3 below investigates the associated implication for pro-cyclical markups in detail. For ease of identifying aggregate demand contractions, we add shaded

[23]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 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 the IRI elasticity estimates by a constant to match that estimated in the Nielsen data set.

[24]We note that the increase in elasticities in 2008-2009 is present both in the IRI and Nielsen data. See Figure A.4 for the cyclical variation of demand elasticities in both IRI and Nielsen samples.

bars in Figure 4 below.

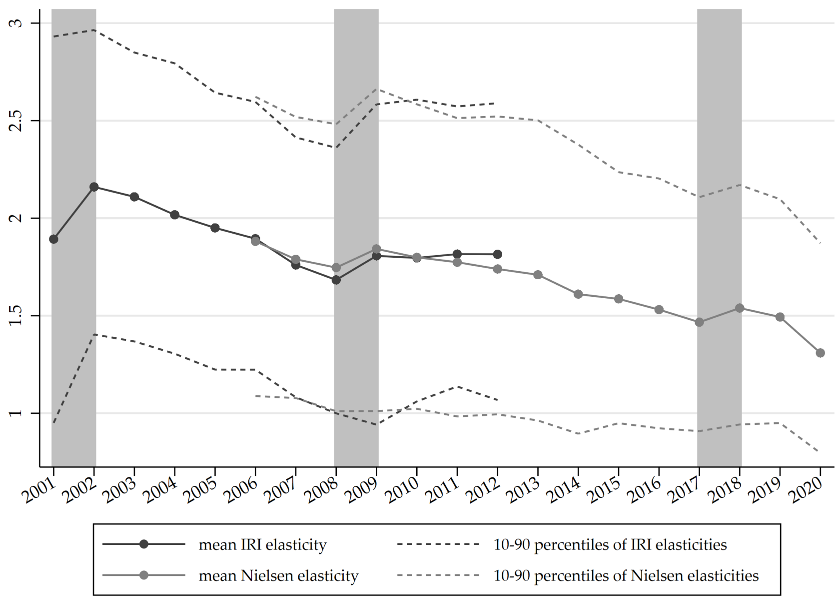


Figure 4. Demand Elasticity Estimates over Time.

Note: In the figure above, the shaded periods refer to the two substantial demand contractions—the 2001 dot-com recession and the 2008 great recession—and one monetary-policy-induced contraction in 2017. As explained below in the text, due to differences in store attributes and good coverage between the two samples, we shift IRI demand elasticity estimates down by a constant. In this way, we can focus on the time variation.

We check that the cyclical and trend variation is present both across goods and by markets (see the examples shown by Figures A.5 and A.6 in Appendix A). 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 Nielsen elasticity estimates into one sample for all the analyses below.

Third, we note that the dynamic pattern of elasticities is not only present in the mean elasticity but also reflected by the entire distribution. We can clearly see this by examining the tenth and ninetieth percentiles of the elasticity distribution, which are also reported in Figure 4 above.[25]

We next fit a linear trend to our elasticity estimates and uncover the cyclical variation around it. Specifically, we regress market-good-year elasticities on a linear trend and year dummies (two reference years dropped) with market-good-specific fixed effects.[26] We efficiently extract the trend and cyclical variation by using standard errors of elasticity estimates as weights, concerning heterogeneous precisions among these elasticity estimates (see

[25]Recently De Loecker et al. (2020) argues 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 4 suggest declining elasticities and therefore increasing markups across all quantiles of the distribution.

[26]We use market-good-specific fixed effects to control the level differences in elasticity estimates across markets, categories, and samples. We drop year dummies for 2006 and 2015, the two reference years, such that the cyclical variation around the linear trend has a statistically zero sum.

Table 5). The coefficients associated with the linear trend and year dummies will deliver us the size of the trend and cyclical variation embedded in our elasticity estimates.

We find a substantial and highly statistically significant downward trend in demand elasticities of 3.5 percentage points per year, and a sizable increase of, on average, 16.3 percentage points during times of aggregate demand contractions.[27] Figure 5 below provides a graphical representation of their trend and cyclical variation over time. In particular, Table 6 in Section 3.4 below shows that cyclical changes in demand elasticities are exceptionally large right after negative aggregate demand shocks. For instance, elasticities increase by 3 times the downward trend after the 2008 financial crisis. Given these economically meaningful changes over time, we discuss their implications for markups below in Section 3.4.

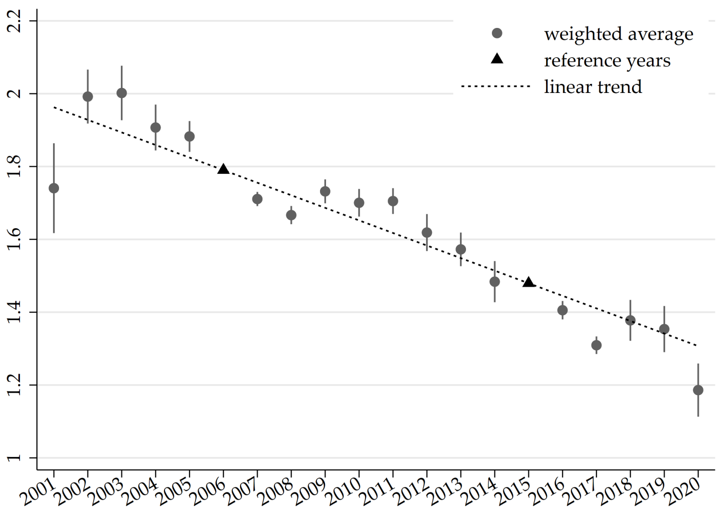


Figure 5. Trend and Cyclical Variation of Demand Elasticity.

Note: This figure shows the trend and cyclical variation of demand elasticity estimates in the pooled sample. The spikes present the 95% confidence intervals of the weighted-average annual demand elasticities relative to the linear trend. To extract the trend and cyclical parts, we run OLS of demand elasticity estimates in the pooled sample on year dummies and a linear year trend, using the estimated standard errors of demand elasticity estimates as weights. Also, we use market-good-specific fixed effects to control the level differences in demand elasticity estimates across markets, goods, and samples. In addition, we drop year dummies for 2006 and 2015, the two reference years, such that the cyclical variation around the linear trend has a statistically zero sum. Finally, we cluster standard errors at the market level.

3.4 Implications for markups

Consider a profit maximizing monopolist that optimizes profits for each product separately. The first order condition for the profit optimization implies

$$\mu = \frac{p}{c} = \frac{e(p)}{e(p)-1}, \tag{2}$$

[27]Table 6 gives the estimation information together with the implied markup changes in Section 3.4.

where $e(p)$ is the elasticity of demand with respect to the own price p at the optimum and c is marginal cost. Assuming that all stores behave as local monopolists, this formula predicts markup μ (price over marginal cost).[28] This assumption implies that each store, at every given moment, adjusts product prices to maximize profits given time varying costs.

By implication, our elasticity estimates measure the local average demand elasticity at a point close to the optimal price. When the demand curve is sufficiently smooth, within a year we can treat price movements as being along a linear demand curve (in logs). Then, our markup estimates will be valid even if the true demand curve is non-linear in logs.[29]

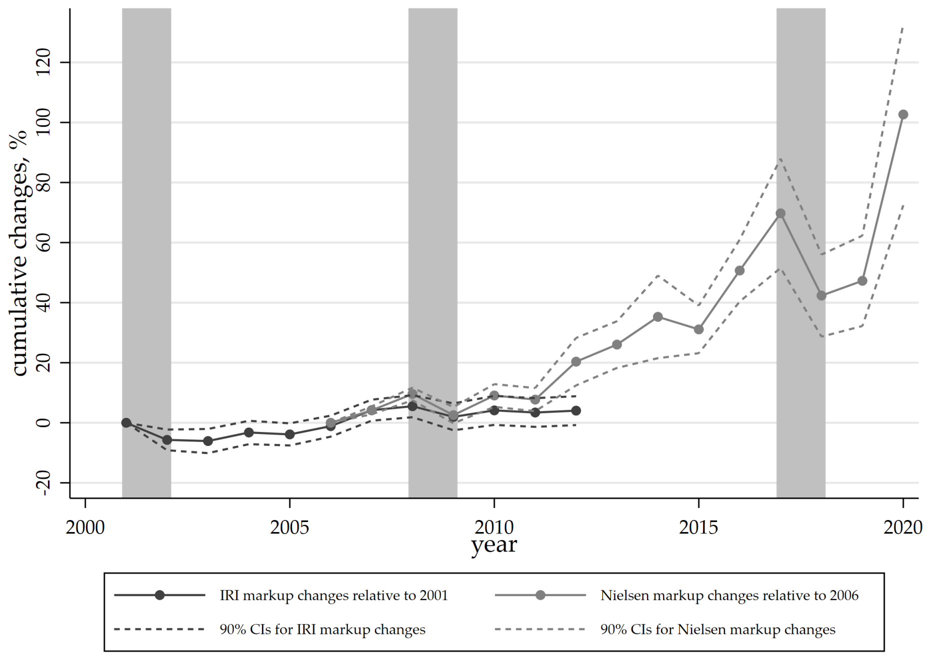


Figure 6. Cumulative Changes in Implied Markups.

Note: For each sample, either IRI or Nielsen, we first run demand elasticity estimates on year dummies, using the estimated standard errors of these demand elasticity estimates as weights, to obtain the weighted average annual demand elasticities and their variance-covariance matrix. Then, we calculate markups and their percentage changes relative to the initial year, 2001 in the IRI sample or 2006 in the Nielsen sample. Also, we use the standard delta method, outlined in Appendix C.1, to obtain the standard errors of these cumulative markup changes and their 90% confidence intervals based on the variance-covariance estimates of annual demand elasticities. The shaded periods refer to the two substantial contractions—the 2001 dot-com recession and the 2008 great recession—and one tightening-monetary-policy-induced contraction in 2017.

Markup is widely used to measure market power in the empirical literature (Basu, 2019). It is also closely related to the conventional Lerner's index (Lerner, 1934), another standard measurement of market power.[30] The main advantage of the approach in which we impute

[28]The original equation is due to Cournot (1838). For a complete recent discussion see Section 8.5.1 in Gallego et al. (2019).

[29]DellaVigna and Gentzkow (2019) provide evidence that stores do not set prices according to the store-level demand elasticities. Their paper, however, assumes that demand elasticities are fixed over time (i.e. demand function is globally linear in logs), while we find that elasticities vary a lot from year to year.

[30]Lerner's index is defined as the difference between product price and its marginal cost over the marginal

markups from elasticities is that it does not require knowledge of actual marginal costs or the competitive structure of the industry.

Ideally, we could impute markups from elasticity estimates for each good sold in each market and year. However, a small share of our elasticity estimates are below one, although most of them are statistically indistinguishable from one. For these elasticity estimates, the markup formula in (2) is inapplicable. On the other hand, the average elasticity estimate in each year is well above one. Given our research interest in the general markup dynamics, we choose to impute annual markups from average elasticity estimates within each year.

Figure 6 above reports time variation in annual markups. It is clear that markups have been growing in the last two decades. The cumulative change over the sample period is close to 100%. However, the annual growth rate has not been constant. There are noticeable and statistically significant decreases in markups in times of negative aggregate demand shocks (years 2002, 2009, and 2018).

Table 6. Trend and Cyclical Variation in Elasticity and Implied Markup.

	elasticity[a]	markup[b]
Trend		
average annual change, 2001-2020	-0.035***	3.9%***
	(0.004)	(0.7%)
Cyclical changes		
from 2001 to 2002	0.286***	-15.1%***
	(0.052)	(2.8%)
from 2008 to 2009	0.100***	-8.0%***
	(0.013)	(1.0%)
from 2017 to 2018	0.103***	-17.8%***
	(0.025)	(4.3%)

Note: [a]To extract the linear trend and cyclical variation of elasticity, we run OLS of demand elasticity estimates in the pooled sample on year dummies and a linear year trend, using the estimated standard errors of demand elasticity estimates as weights. Standard errors are clustered at the market level. Also, we use market-good-specific fixed effects to control the level differences in demand elasticity estimates across markets, goods, and samples. In addition, we drop year dummies for 2006 and 2015 such that the cyclical variation around the linear trend has a statistically zero sum. The average cyclical variation is around 0.064 per year, which is not reported here. [b]We first impute the (nonlinear) markup trend implied by the linear trend of elasticity, following equation (2) above in the text. We set the base level of the estimated trend and cyclical variation in elasticity to that of the year 2006 in the Nielsen sample as the Nielsen sample covers more goods and years as well as more data. Also, 2006 is the year when the IRI sample starts to overlap with the Nielsen sample in time. Then, we calculate their average annual percentage change and its standard errors using a conservative method outlined in Appendix C.2. Likewise, we calculate annual percentage changes in markups relative to their trend, which are implied by the cyclical variation of elasticity. Standard errors are in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

cost of the product, namely $\frac{p-c}{p}$ or equivalently $\frac{1}{e(p)}$ based on our syntax. As you can see, it delivers the same information about the market power of a seller as markup.

In addition to the nonparametric estimates of the changes in annual markups, it is instructive to 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 business cycles.

Table 6 above provides a trend-cycle decomposition of changes in annual markups based on the trend-cycle estimation for annual elasticities provided in Section 3.3. The average trend growth in markups has been equal to 3.9% per year over 2001-2020. The impacts of negative aggregate demand shocks vary between −17.8% (for 2017-2018) and −8.0% (for 2008-2009).[31] In light of the trend-cycle decomposition of annual elasticities in Figure 5, the drastic reduction in elasticity (and a corresponding jump in markup) in the first year of the COVID-19 pandemic, 2020, is likely a short-run phenomenon driven by strong fiscal and monetary stimuli. People's concerns over getting COVID during multi-stop shopping may make them less price sensitive and thus raise stores' markups.

4 Factors driving the dynamics in the elasticities

Results in the previous section show that elasticities and markups exhibit substantial secular and cyclical variation. There have been several mechanisms discussed in the literature that can drive price elasticities and markups. In particular, there is suggestive evidence that, with higher wealth, people become less sensitive to price changes (Stroebel and Vavra, 2019). Alternatively, markups may grow as a result of increased market concentration (Philippon, 2019). In this section, we will investigate the contribution of various factors, specifically average income, wealth, and market concentrations, in explaining the observed dynamics in the elasticities and thus the corresponding markup measures.

4.1 Explanatory variables

We use measures of regional macroeconomic and demographic factors that may drive the time variation of local demand elasticities based on data gathered from publicly available sources.[32] Appendix B shows data sources and describes the construction of the market-year-specific explanatory variables listed below. Here, we give a brief discussion about their potential impacts on demand elasticities.

[31] Note that these cyclical changes are computed relative to the trend. That is, when calculating cyclical changes in annual markups, we subtract the estimated trend component from the raw changes in annual markups. Nevertheless, these raw changes in annual markups are also statistically significant right after negative aggregate demand shocks, which is not reported here but partly visible in Figure 6.

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

Real GDP per capita: Higher GDP means higher income. An increase in household income may make them less price sensitive.

Unemployment rate: Households experiencing unemployment are more likely to shop around for cheaper prices due to income loss.

Economic dependency ratio: We use the economic dependency ratio to supplement the unemployment rate as the latter does not include the changing population that is not in the labor force but needs to be fed. Intuitively, the higher the feeding burden is, the more likely a household will be budget-constrained and thus more sensitive to price changes.

Cumulative changes in real housing prices: For homeowners, rising housing prices may mean higher wealth and thus make them less price sensitive (Stroebel and Vavra, 2019). However, Stroebel and Vavra (2019) also noticed that for renters, higher housing prices may mean higher rent burdens and thus make them more price sensitive. We rely on the data to find out which effect dominates or if they simply cancel out each other.

Grocery establishments per capita: A higher number of establishments per capita means less costly for households to shop around for cheaper prices. Hence, they may become more price sensitive. However, more stores in a local grocery market may also mean less concentration and thus more competition (Philippon, 2019).[33] If this is the case, then households may have less need to shop around for cheaper prices. We rely on the data to find out which effect dominates or if they simply cancel out each other.

Population: Local population growth may reflect people's outlook on future economic development in each market. A better economic outlook generally gives households confidence in their future income growth and thus makes them less price sensitive today.

In Figure A.7, we show the cross-sectional and time variations of these six market-level factors. They not only differ substantially across markets within a year but also have notable changes over time. Importantly, their non-negligible market-specific time variation gives us the statistical power to identify the impacts of these macroeconomic and demographic factors on food demand elasticities across markets (see details below).

4.2 Weighted least squares approach

The second part of our empirical work is to investigate if the macroeconomic and demographic factors proposed above drive demand elasticity dynamics. To this end, we pool demand elasticity estimates in the IRI and Nielsen samples as before, and run regression (3) below:

$$\tilde{\hat{e}}_{m,c,t} = \tilde{X}'_{m,t}\beta + year_t + \epsilon_{m,c,t}, \tag{3}$$

[33] The IRI scanner data does not have the information about which retail each store in a local market belongs to, although the Nielsen scanner data has that information. Hence, we do not use our scanner data to calculate market concentration measures at the retail level for each market.

where $\tilde{\hat{e}}_{m,c,t}$ and $\tilde{X}_{m,t}$ are demeaned demand elasticity estimates and factors, respectively, while $year_t$ stands for year fixed effects.

Within each sample, either IRI or Nielsen, we demean demand elasticity estimates $\hat{e}_{m,c,t}$ by subtracting them from their market-good-specific means. When calculating these means, we treat the same geographic market in IRI and Nielsen data sets as two separate markets given their differences in store attributes and good coverage. Similarly, we treat common good categories like milk from the two data sets as two separate categories. By demeaning, we only use the time-variant part of demand elasticity estimates $\tilde{\hat{e}}_{m,c,t}$ as the variable to be explained in regression (3). Likewise, we subtract factors $X_{m,t}$ from their market-specific means and solely use their time-variant part $\tilde{X}_{m,t}$ as explanatory variables.

We use year fixed effects to control nationwide factors that uniformly affect demand elasticities across all local markets. This ensures that our identification of the key coefficients β comes from the local dynamics of the proposed highly-persistent market factors as it eliminates their spurious correlations with demand elasticities. As shown above, there is non-negligible market-specific time variation in both these factors and demand elasticity estimates, giving us the statistical power to identify β.

In the final step, we then use the estimated standard errors of demand elasticity estimates as weights when we implement regression (3). This will improve our estimation precision given the notable dispersion in these standard error estimates (see Table 5). For the standard errors in regression (3), we cluster them at the market level to allow for arbitrary time dependence in the error term within each market.

4.3 Results based on cross-sectional variation

In this section, we study whether the six macroeconomic and demographic factors, proposed in Section 4.1, can explain the documented trend and cyclical variation of demeaned market-good-year demand elasticity estimates.

Table 7 below reports the results from estimating equation (3). We use year fixed effects to control for any nationwide factors that affect demand elasticity dynamics across all local markets. This gives us confidence that the regression coefficients of the proposed market-specific factors are reliable.[34] Results in columns [2]-[7] show that *real GDP per capita* has the most explanatory power as an individual factor, measured by $adj.R^2$, and it is the only one that is statistically significant. Its negative sign means that demand elasticity decreases along with economic growth. This makes sense as higher income relaxes households' consumption budgets and thus makes them less price sensitive.

[34]Nationwide factors can simultaneously affect these market-specific factors and local demand elasticities. The year fixed effects eliminate such spurious correlations.

Table 7. Factor Regression Results.

	dependent variable: *demand elasticity estimate*[b]							
explanatory variables[a]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
real GDP per capita		-0.72***						-0.85***
		(0.21)						(0.16)
unemployment rate			1.77					1.49
			(1.18)					(1.25)
cum. change in real housing price				-0.01				0.40***
				(0.12)				(0.13)
economic dependency ratio					0.56			0.35
					(0.37)			(0.38)
population						-0.54		-1.20**
						(0.55)		(0.56)
grocery establishments per capita							-0.03	-0.02
							(0.23)	(0.17)
year fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
$adj.R^2$	0.317	0.324	0.319	0.317	0.322	0.320	0.317	0.337
N^c	25062	25062	25062	25062	25062	25062	25062	25062

Note: [a]See Appendix B for the detailed descriptions of these market-specific macroeconomic and demographic factors. We subtract their market means before putting them into regressions. [b]Demand elasticity estimates, however, are at the market-good level. We subtract their market-good-specific means before putting them into regressions. Hence, all the regression results above only capture the relationships between the proposed factors and demand elasticities over time. See the main text above in Section 4.2 for detailed descriptions of our estimation approach. [c]A small number, 11 out of 25,073, raw demand elasticity estimates were dropped as they do not have well-estimated standard errors. All the models above are estimated by OLS using the estimated standard errors of demand elasticity estimates as weights. Standard errors, clustered at the market level, are listed in the parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Combining all six factors notably improves the explanatory power, as shown by the regression result [8]. The $adj.R^2$ increases to 0.337, i.e., the six factors can explain about 34% of demand elasticity dynamics together with nationwide factors. Importantly, *cumulative change in real housing price* and *population* now become statistically significant and have expected signs. Rising housing prices may mean more rent burdens, which may constrain households' budgets and thus make them more price sensitive. Population growth indicates good economic development, which may give people confidence in future income growth and thus make them less price sensitive.

The coefficient signs of the other three factors are also reasonable given their potential impacts discussed in Section 4.1. For example, the positive coefficient sign of *unemployment rate* echoes the intuition that households with more unemployed people tend to be more price sensitive due to income loss. Although they are not statistically significant, we cannot conclude that these factors do not contribute to demand elasticity dynamics. Because nationwide factors may largely absorb their effects through year fixed effects in the regressions.[35]

[35]One example of nationwide factors that contributed to the decline in elasticity is the increasing share of online shopping (Döpper et al., 2022; Harris-Lagoudakis, 2023). Döpper et al. (2022) find that this factor

Likewise, we cannot say that the six proposed factors as a whole can only explain a small part of demand elasticity dynamics, even though the $adj.R^2$ only has a moderate increase after adding them into the regression model [1] that only has year fixed effects.

To see how well the six proposed factors explain the documented trend and cyclical variation of demand elasticities, we turn to a graphical analysis. Using the coefficients obtained in result [8] of Table 7 (but not year fixed effects), we first calculate annual changes in the average factor-fitted demand elasticity. We then compare fitted values with annual changes in the average estimated demand elasticity. Figure 7 below plots the two resulting series. We can see that cumulative changes in the factor-fitted demand elasticity relative to the base year 2001 are well aligned with the pattern of the cumulative changes in the estimated demand elasticity. This means that the six proposed factors can largely explain the nationwide dynamics in demand elasticities.

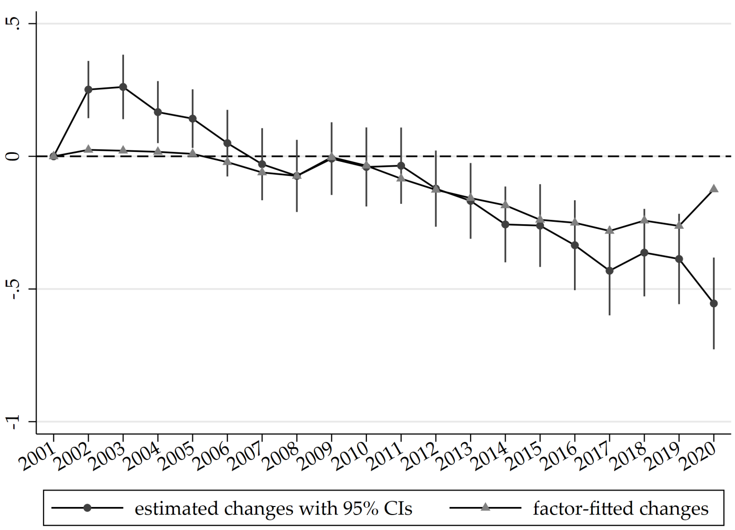


Figure 7. Cumulative Changes in Annual Demand Elasticity.

Note: We estimate these cumulative changes in annual demand elasticity as follows: (i) Run OLS of demand elasticity estimates in the pooled sample on year dummies (the year dummy for 2001 dropped) with market-good fixed effects, using the estimated standard errors of demand elasticity estimates as weights. This step delivers us the estimated changes in annual demand elasticity relative to the base year 2001 with 95% confidence intervals. (ii) Take simple averages of the six market-specific macroeconomic and demographic factors in each year. Then, calculate the factor-fitted changes in annual demand elasticity relative to the base year 2001 using their estimated coefficients in the regression result [8] of Table 7 above.

We also note the two time series are not perfectly aligned. This is especially true in the first part of the sample, where estimates are based only on the IRI data.[36] The big divide in 2020, however, may be caused by people's concerns about the higher risk of getting COVID through multiple-stop shopping relative to one-stop shopping. In other words, households

only explains about 1% of the observed time variation though.

[36] The fact that the Nielsen sample contributes to more than 80% of observations in the pooled sample may also cause these gaps mechanically as the Nielsen sample does not cover early years.

are less likely to shop around for cheaper prices during the pandemic. On the other hand, government relief policies like stimulus checks, which help offset households' income losses during the pandemic, can also contribute to the unexpected drop in demand elasticities.

The analyses in this section—both regression-based and graphical—indicate that regional and macroeconomic factors can explain a large share of the time variation in demand elasticities and thus implied markups.

5 Discussion

5.1 Implications of trend variation

According to our estimates, the average markup growth from 2001 to 2019 is about 47%. This number is comparable with other studies, e.g., De Loecker et al. (2020), Philippon (2019), and Brand (2021). At first glance, this markup increase seems quite large. Interestingly, the CPI subindex for food at home (food groceries) has increased by 39% over the same period while the overall CPI index has increased by 46%.[37] The most likely explanation for this parallel pattern is that the nominal marginal cost of food supply has almost remained unchanged, i.e., the real marginal cost has declined during this period. Indeed, Döpper et al. (2022) note that this is the reason that larger markups did not reach consumers in the form of higher real food prices.

It is possible that the lower real marginal cost was the result of large fixed-cost investments and efficiency increases in the food industry (as suggested, for example in Watson and Winfree (2022)). Such investments may have resulted in large improvements in logistics and food waste management. Suggestive evidence in support of this interpretation is that labor productivity in the food retail industry increased by 36% from 2001 to 2019.[38] We leave the exploration of these patterns for future research.

We also note that our measure of market concentration does not have a significant impact on price elasticities and therefore markups. It is possible that, since we include time fixed effects, our regression model is unable to capture changes in markups across the whole country driven by changes in market structure. The common fixed effect would simply capture this overall change. We have checked that market concentration increases steadily over time.[39] However, we cannot run a panel regression model without time fixed effects since there may be other changes that have caused increases in markups and which we might spuriously

[37] 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.

[38] The labor productivity among US food and beverage stores has increased from 84.9 (an index) in 2001 to 115.4 in 2019, according to the FRED data (https://fred.stlouisfed.org/series/IPUHN445L000000000).

[39] Dong et al. (2023) come to the same conclusion based on a wider set of indicators of market concentration.

attribute to the trend in market concentration. Instead, what we find are important effects of income measures. These effects are estimated off of differential changes in the cross section over time, and so the effects can be identified by our model.

5.2 Economic significance and implications of cyclical variation

Our results also directly show that macroeconomic conditions have a sizable impact on markups, making them pro-cyclical. Monetary policy and aggregate demand shocks work by making consumers more or less sensitive to price changes. The monetary tightening in 2017 coincided with a decrease of 18% in markups, while the government spending stimulus in 2020, among others, was associated with an increase of 50% in markups.

There has been growing empirical literature on pro-cyclical markups (Nekarda and Ramey, 2020). Our paper provides micro-level evidence on the pro-cyclicality of natural markups (the markups in the absence of sticky prices), which is largely driven by income effects.

A related wealth-effect mechanism has been found by Stroebel and Vavra (2019). 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. However, our results suggest that higher housing prices may also mean higher rent burdens for renters, which tends to make them more price sensitive and thus lower retail markups.

5.3 Benefits of our panel-IV approach to estimating elasticities

Our strategy of pairing geographically close markets worked very well with our Hausman-type price IV. As shown in Section 3.2, most instruments are strong and elasticity estimates are precise. Results are very different from OLS, possibly due to the severe endogeneity problem of the price variable in the elasticity estimation.[40] Other studies have used the Berry et al. (1995) methodology (BLP) to estimate an internally consistent structural model of consumer demand. BLP has been successful at incorporating product and consumer characteristics and allowing for counterfactual and welfare analyses in addition to recovering the price elasticities of all related items in a product market. However, for the markup estimation, we only need the own-price elasticity of demand, which can be directly and efficiently estimated in a reduced-form approach like our panel-IV regression model.

In our panel-IV regression model, demand shocks from competing items within a good category are largely absorbed by the time fixed effects. Moreover, our model has separate fixed effects to control for consumer preferences over stores of different retailers and items of different brands. In addition, we account for differential demand changes among markets,

[40] We have a figure, showing that the OLS elasticity estimates are notably smaller in magnitude relative to our IV estimates (downward bias), which is available upon request.

which may arise from local economic and demographic dynamics, by conducting elasticity estimations on a market-yearly basis. Put everything together, our panel-IV approach can control for almost the same demand shifters as BLP does but without imposing additional structural assumptions that are not directly relevant to the parameter of interest (the own-price elasticity of demand). As a result, our elasticity and markup estimates are similar to those based on the structural BLP model like Döpper et al. (2022). But our estimates have smaller standard errors possibly because we do not consolidate items within a good category and allow for high-frequency weekly price changes. This enables us to statistically detect cyclical variation in elasticities and therefore markups.

Finally, there has been a growing body of literature on the estimation of demand elasticities in the presence of many items within a good category. Importantly, it has been documented that one cannot estimate elasticities at the UPC level precisely without imposing regularization or shrinkage towards the category-level average (Hitsch et al., 2019; Chernozhukov et al., 2019). As a natural simplification of these complicated impositions, our category-level estimation approach also allows us to precisely estimate elasticities.

In sum, our panel-IV approach is both economically flexible and statistically efficient in estimating disaggregated price elasticities at scale. This enables us to precisely uncover meaningful economy-wide variation in elasticity and therefore markup from these disaggregated elasticity estimates using weighted least squares in the second step of our estimation.

6 Conclusion

In this paper we provide direct micro-level evidence of a substantial downward trend in demand elasticities and an associated upward trend in markups in the U.S. from 2001 to 2020. Moreover, our markup measure exhibits economically large and statistically significant drops at times of contractions in aggregate demand, implying pro-cyclical variation in markups.

Our findings are non-parametric in the sense that we do not assume any single particular parametric model of consumer demand. Instead, we approximate local linear demands in panel-IV regressions based on a widely-used Hausman IV strategy. Our particular implementation of this strategy only assumes common wholesale suppliers for each specific pair of geographically close markets. Unlike some others, we work directly with weekly sales data which results in strong IVs. The methodology used in this paper can also be applied to study markups in other consumer product industries where high-frequency micro-level sales and price data are available.

The trends in food markups that we find are representative of the whole U.S. economy since we cover markets in the Western, Mid-Western, Eastern, and Southern US. We also find common trends across a wide variety of food categories. The food markup behavior could also be seen as representative of consumer good market behavior in general. Nevertheless, the trend in elasticity that drives the trend in markup is likely to flatten in the near future as the average elasticity is getting too close to one.

References

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

Basu, S. (2019): "Are price-cost markups rising in the United States? A discussion of the evidence," *Journal of Economic Perspectives*, 33, 3–22.

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

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.

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

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

Chernozhukov, V., J. A. Hausman, and W. K. Newey (2019): "Demand analysis with many prices," Tech. rep., National Bureau of Economic Research.

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

Cox, C. and K. Harris-Lagoudakis (2022): "SNAP eligible products and behavioral demand," *Available at SSRN 4146549*.

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 (2016): "Estimating market power: Evidence from the US brewing industry," Tech. rep., National Bureau of Economic Research.

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

Dong, X., J. V. Balagtas, and A. T. Byrne (2023): "A closer look at the relationship between concentration, prices, and market power in food retail—A monopolistic competition and differentiated products approach," *Applied Economic Perspectives and Policy.*

Döpper, H., A. MacKay, N. Miller, and J. Stiebale (2022): "Rising markups and the role of consumer preferences," *Harvard Business School Strategy Unit Working Paper*.

Elzinga, K. G. and D. E. Mills (2011): "The Lerner index of monopoly power: Origins and uses," *American Economic Review*, 101, 558–564.

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

Harris-Lagoudakis, K. (2023): "The effect of online shopping channels on brand choice, product exploration and price elasticities," *International Journal of Industrial Organization*, 87, 102918.

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.

Hitsch, G. J., A. HortaÃ§su, and X. Lin (2019): "Prices and promotions in US retail markets: Evidence from big data." Tech. rep., National Bureau of Economic Research.

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

Leung, J. and Z. Li (2022): "Rising retail concentration: Superstar firms and household demand," *Available at SSRN 3981995.*

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*.

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.

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

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

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.

Watson, P. and J. Winfree (2022): "Should we use antitrust policies on big agriculture?" *Applied Economic Perspectives and Policy*, 44, 1313–1326.

A Additional tables and figures

Table A.1. The List of Food Categories.

IRI food (16)	Nielsen non-dry food (24)[a]	Nielsen 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	CRACKERS
soup	ICE CREAM, NOVELTIES	DESSERTS, GELATINS, SYRUP
spaghetti&italian sauce	JUICES, DRINKS-FROZEN	FLOUR
sugar substitute	**MILK**	FRUIT - CANNED
yogurt	PACKAGED MEATS-DELI	FRUIT - DRIED
	PIZZA/SNACKS/HORS DOEURVES-FRZN	JAMS, JELLIES, SPREADS
	PREPARED FOODS-FROZEN	JUICE, DRINKS - CANNED, BOTTLED
	PUDDING, DESSERTS-DAIRY	NUTS
	SNACKS, SPREADS, DIPS-DAIRY	PACKAGED MILK AND MODIFIERS
	UNPREP MEAT/POULTRY/SEAFOOD-FRZN	PASTA
	VEGETABLES-FROZEN	PICKLES, OLIVES, AND RELISH
	YEAST	PREPARED FOOD-DRY MIXES
	YOGURT	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]Nielsen non-dry food categories span five departments including dairy, deli, fresh produce, frozen foods, and packaged meat. [b]Nielsen dry food refers to food categories in the dry grocery department. The bold food categories are the ones that IRI and Nielsen share the same names, although Nielsen have more varieties or larger numbers of UPCs within each of these five common food categories.

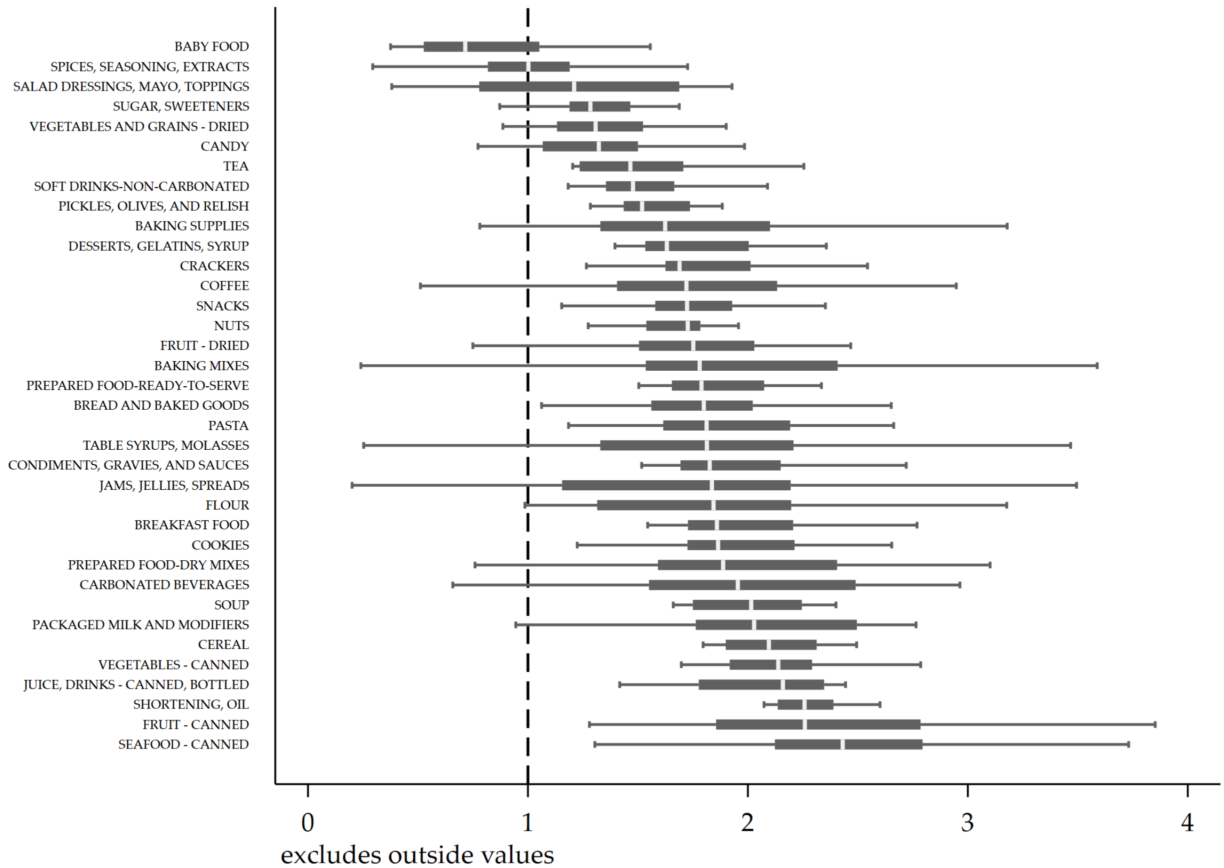


Figure A.1. Nielsen Demand Elasticity Estimates by Dry Food Categories in 2010.

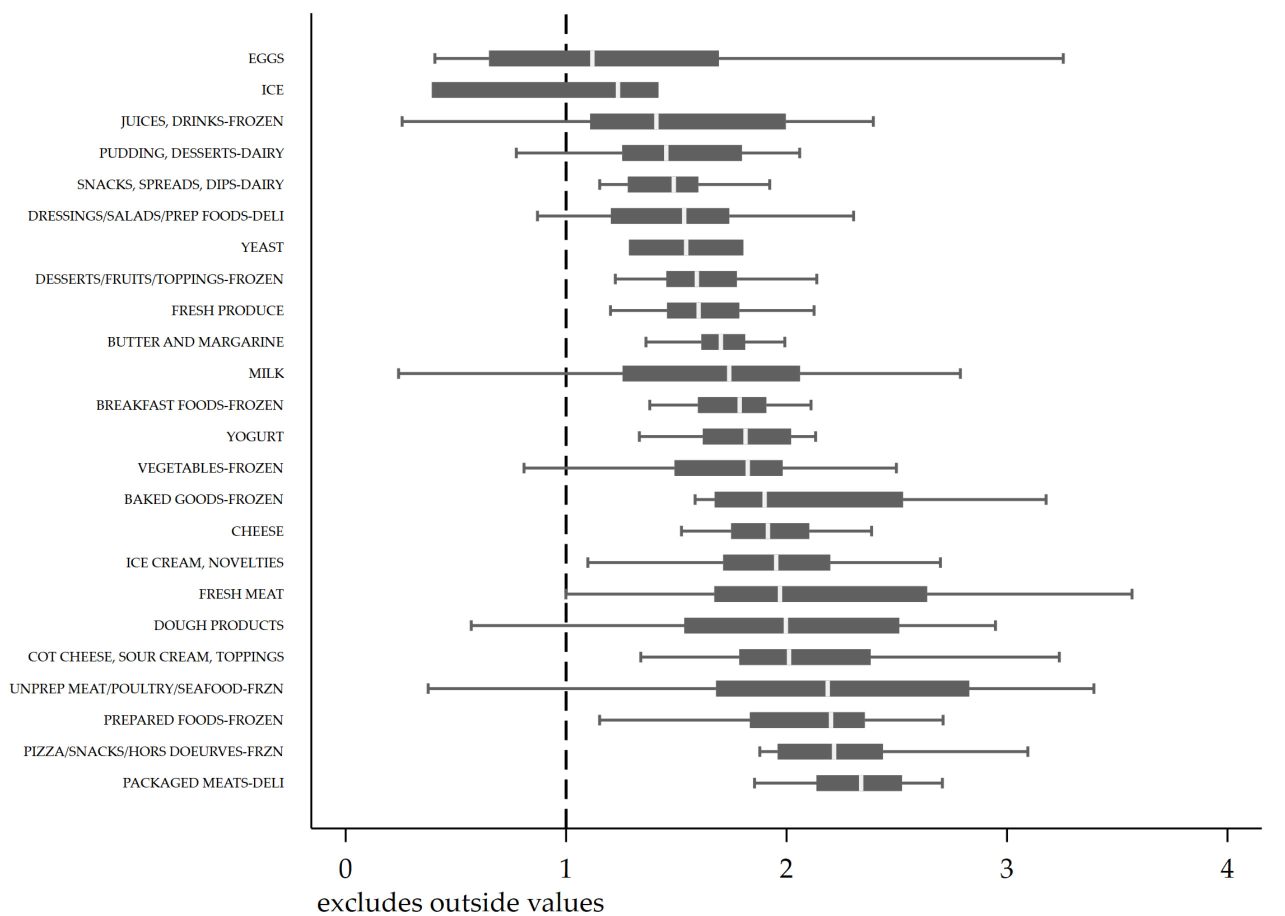


Figure A.2. Nielsen Demand Elasticity Estimates by Non-dry Food Categories in 2010.

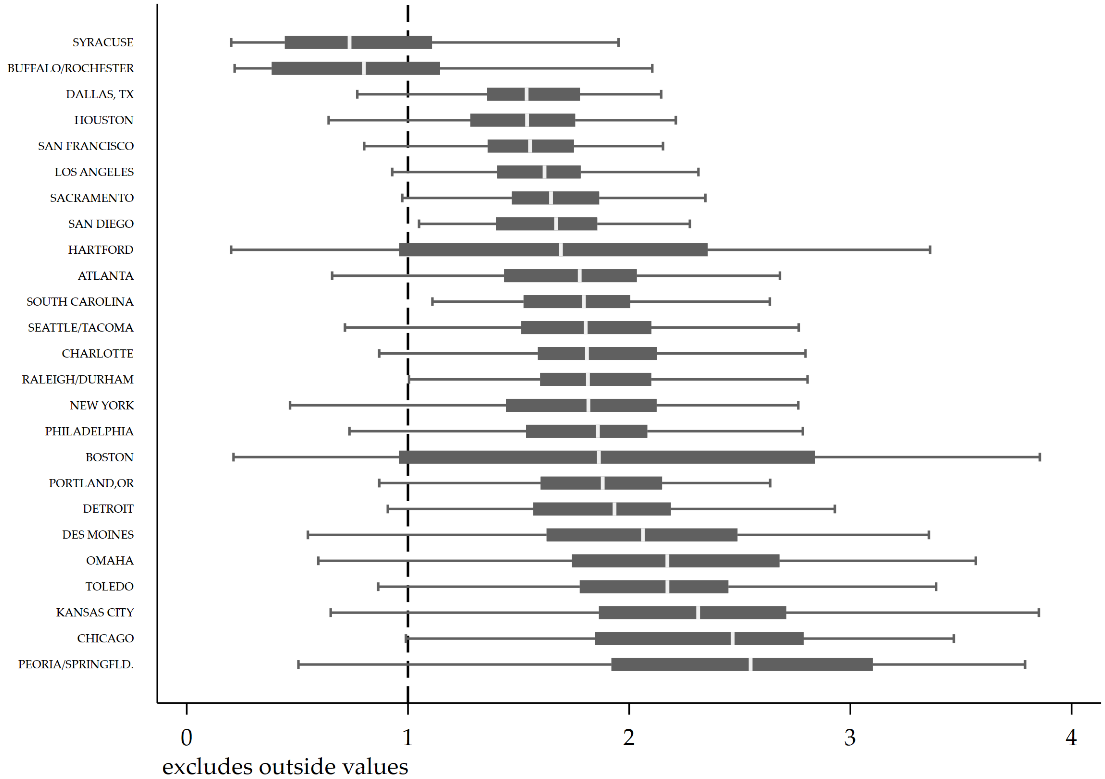


Figure A.3. Nielsen Demand Elasticity Estimates by Markets in 2010.

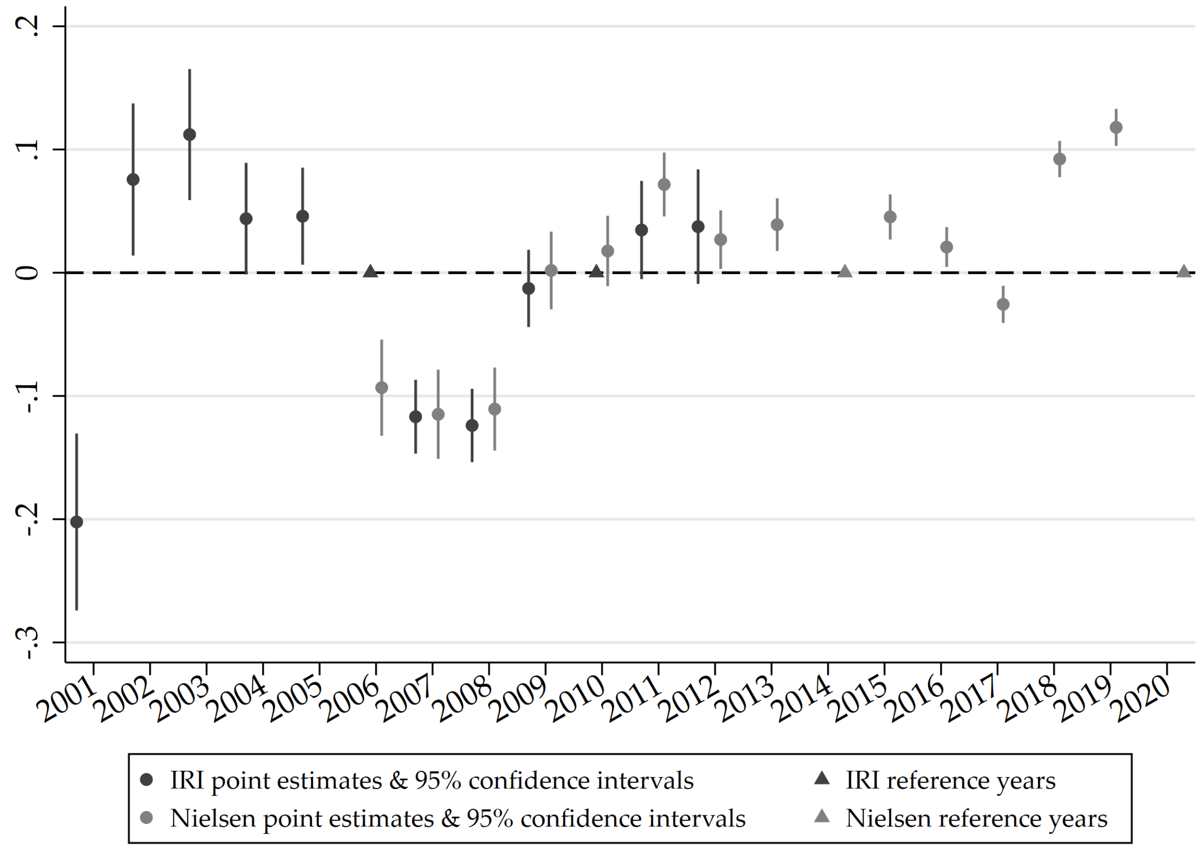


Figure A.4. Cyclical Variations of IRI and Nielsen Demand Elasticity Estimates.

Note: To obtain the cyclical variation in the IRI sample, we run demand elasticity estimates on dummies of years 2001-2005 and 2007-2011, and a linear year trend with market-good-specific fixed effects, using standard errors of raw elasticity estimates as weights. We drop year dummies for 2006 and 2010 such that the cyclical variation around the linear trend has a statistically zero sum. Using 2014 and 2020 as reference years, we run a similar regression to obtain the cyclical variation of demand elasticity estimates in the Nielsen sample. All the associated coefficients of year dummies and their 95% confidence intervals are plotted above.

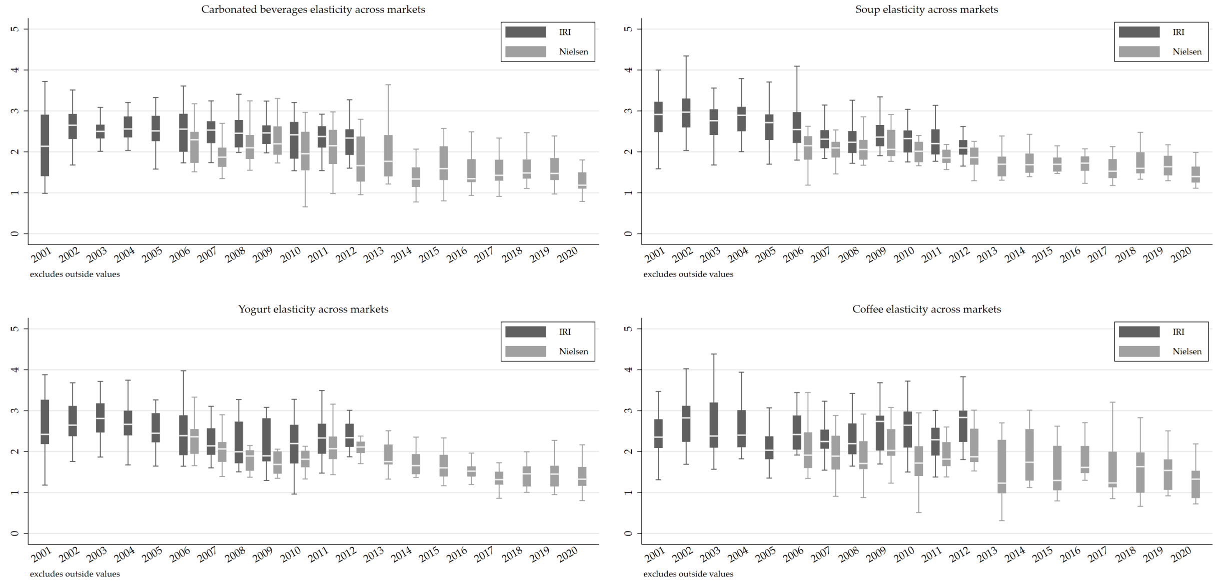


Figure A.5. Demand Elasticity Estimates of the Four Common Food Categories.

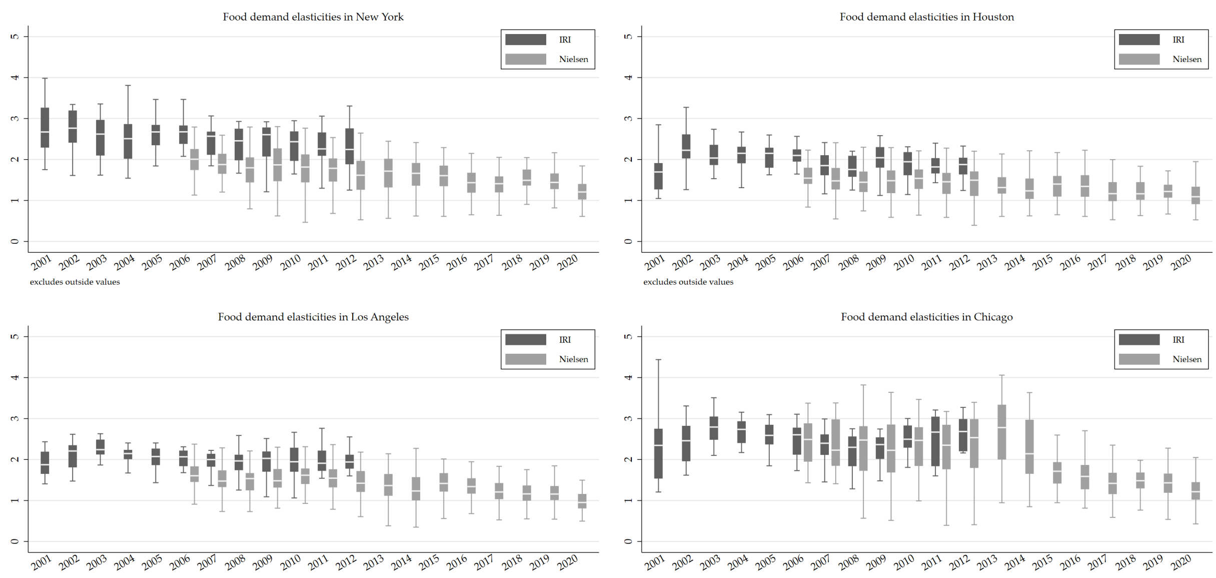


Figure A.6. Demand Elasticity Estimates in the Four Largest Regional Markets.

B Data construction for factor variables

Before describing the construction of actual factor variables, we note that several of our 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 2000, right before our sample period 2001-2020.

Population Population is measured at the county level. For each year, we aggregate data to the market level before taking ln.

Real GDP per capita We deflate the county-level nominal GDP to the base year 2000 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 ln and then average using county-level populations as weights to obtain the market-level counterpart.

Unemployment rate The raw data is in percentage. We transform it into a ln form using the formula $ln(1 + \textit{unemployment rate}/100)$. Then, we average them using county-level populations as weights to obtain the market-level measurement for each year.

Economic dependency ratio We use the formula (*total population - employed population*)/*employed population* for every county in each year. Then, we take ln before we average them into the market level using county-level populations as weights.

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 choose the version with the base year 2000 and deflate it by the GDP deflator. Then, we take ln-differences relative to the base year for every county in each year. Finally, we average them using county-level populations as weights to obtain the market-level cumulative change in real housing prices.

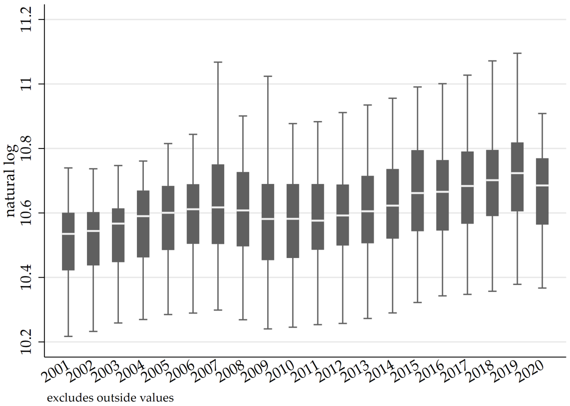
Grocery establishments per capita For each county, we divide the number of grocery stores, reported in the County Business Patterns from the Census Bureau, by its population and then times 10,000 to obtain the number of grocery stores per 10k residents across years. Then, we take ln before we average them into the market level using county-level populations as weights. This variable captures the availability of grocery stores.

Figure A.7 below gives their graphical representations, while Table A.2 below shows the public sources of the original data.

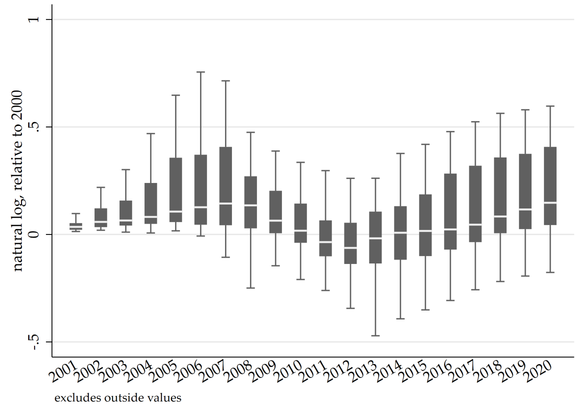
Table A.2. Public Sources of the Raw Data.

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
employed population	county	U.S. Bureau of Labor Statistics
population	county	U.S. Census Bureau
establishments of grocery stores	county	U.S. Census Bureau

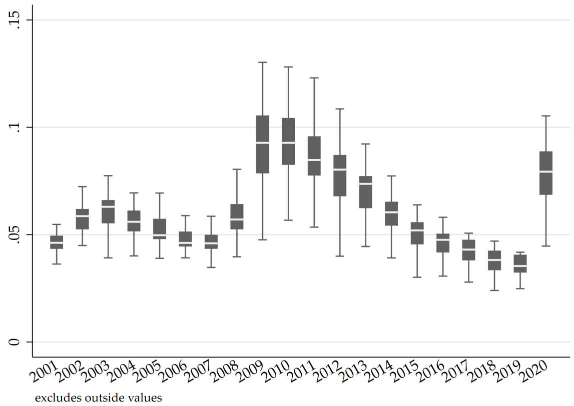
Note: [a]Here are website links below for all the data listed above.
(1) GDP deflator: "https://fred.stlouisfed.org/series/USAGDPDEFAISMEI".
(2) GDP: "https://fred.stlouisfed.org/release/tables?rid=397eid=1054597".
(3) housing price index: "https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx".
(4) unemployment rate & population employed: "https://www.bls.gov/lau/tables.htmcntyaa".
(5) population: "https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates.html".
(6) establishments of grocery stores: "https://www.census.gov/programs-surveys/cbp/data/datasets.html".



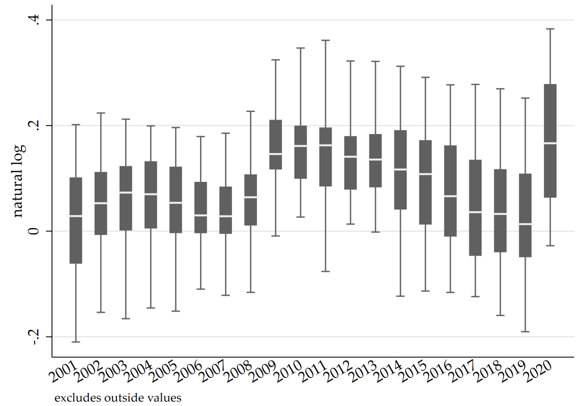
(a) real GDP per capita



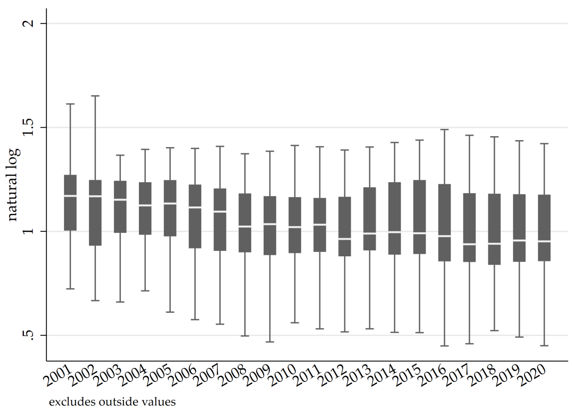
(b) cumulative change in real housing price



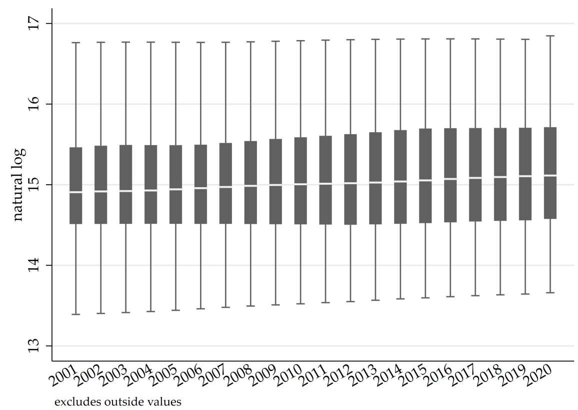
(c) unemployment rate



(d) economic dependency ratio



(e) No. of grocery stores per 10k residents



(f) population

Figure A.7. Cross-sectional and Time Variations of Market-specific Factors.

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

C Additional information about the methods

C.1 Cumulative changes in markups and their standard errors

First of all, we impute markup from demand elasticity using the standard formula below:

$$\mu_t = \frac{e_t}{e_t - 1}, \tag{4}$$

where μ, e, t denote markup, demand elasticity, and year, respectively.

Then, we calculate the cumulative percentage changes in markups by taking ln-differences relative to the base year, i.e.,

$$ln(\mu_t) - ln(\mu_{t_0}), \tag{5}$$

where the base year t_0 refers to 2001 in the IRI sample or 2006 in the Nielsen sample.

Finally, we use the delta method to compute standard errors for the cumulative markup changes. The specific formula is derived below.

Given that

$$\frac{dln(\mu_t)}{de_t} = \frac{1}{e_t} - \frac{1}{e_t - 1} = -\frac{1}{e_t(e_t - 1)}, \tag{6}$$

we have the variance of $ln(\mu_t)$:

$$\sigma^2_{ln(\mu_t)} = \sigma^2_{e_t} \frac{1}{e_t^2(e_t - 1)^2}. \tag{7}$$

Similarly, we have the covariance between $ln(\mu_t)$ and $ln(\mu_{t_0})$:

$$\sigma_{ln(\mu_t),ln(\mu_{t_0})} = \sigma_{e_t,e_{t_0}} \frac{1}{e_t(e_t - 1)} \frac{1}{e_{t_0}(e_{t_0} - 1)}, \tag{8}$$

where $\sigma_{e_t,e_{t_0}}$ denotes the covariance between e_t and e_{t_0}.

Hence, the variance of the cumulative percentage change in markup $ln(\mu_t) - ln(\mu_{t_0})$ will be:

$$\sigma^2_{ln(\mu_t)-ln(\mu_{t_0})} = \sigma^2_{e_t} \frac{1}{e_t^2(e_t - 1)^2} - 2\sigma_{e_t,e_{t_0}} \frac{1}{e_t(e_t - 1)} \frac{1}{e_{t_0}(e_{t_0} - 1)} + \sigma^2_{e_{t_0}} \frac{1}{e_{t_0}^2(e_{t_0} - 1)^2}, \tag{9}$$

where the variances of demand elasticities e_t and e_{t_0} and their covariances are jointly estimated within each sample, either IRI or Nielsen, as described in the note for Figure 6 of Section 3.4. When calculating the cumulative percentage changes in markups and their variances, we simply plug in associated estimates into (5) and (9), respectively. Taking the square root of these variances, we obtain their standard errors.

C.2 Markup trend and its standard errors

Suppose that there is a linear trend in elasticities:

$$e_t^{trend} = a + bt, \tag{10}$$

where a and b are jointly estimated by $\hat{a}$ and $\hat{b}$. Suppose that the estimators have a joint asymptotic normal distribution with an available estimate $\widehat{\Sigma_{ab}}$ for the corresponding asymptotic variance-covariance matrix.

The linear trend in elasticities corresponds to a non-linear trend in ln markups:

$$MarkupTrend_t = ln(1 + \frac{1}{a + bt - 1}). \tag{11}$$

We are interested in the average markup growth per year over 2001-2020 along this trend (in the absence of the cyclical component). In the regression for the trend and cyclical variation in elasticity, we set 2006 as the reference year, i.e., t is the year relative to 2006. Then, the average markup growth is estimated by

$$\widehat{AMG} = \frac{1}{T_2 - T_1}\left(ln(1 + \frac{1}{\hat{a} + T_2\hat{b} - 1}) - ln(1 + \frac{1}{\hat{a} + T_1\hat{b} - 1})\right), \tag{12}$$

where $T_1 = 2001 - 2006$, $T_2 = 2020 - 2006$.

One can compute standard errors for this quantity using the Delta method. Let's first compute the gradient of AMG w.r.t. a and b, evaluated at their estimates $\hat{a}$ and $\hat{b}$:

$$\frac{\widehat{\partial AMG}}{\partial a} = \frac{1}{T_2 - T_1}\left(-\frac{1}{(\hat{a} + T_2\hat{b})(\hat{a} + T_2\hat{b} - 1)} + \frac{1}{(\hat{a} + T_1\hat{b})(\hat{a} + T_1\hat{b} - 1)}\right), \tag{13}$$

$$\frac{\widehat{\partial AMG}}{\partial b} = \frac{1}{T_2 - T_1}\left(-\frac{T_2}{(\hat{a} + T_2\hat{b})(\hat{a} + T_2\hat{b} - 1)} + \frac{T_1}{(\hat{a} + T_1\hat{b})(\hat{a} + T_1\hat{b} - 1)}\right). \tag{14}$$

Then, we can estimate the variance of $\widehat{AMG}$ using the following formula:

$$\widehat{\sigma^2_{AMG}} = \left(\frac{\widehat{\partial AMG}}{\partial a}, \frac{\widehat{\partial AMG}}{\partial b}\right)\widehat{\Sigma_{ab}}\left(\frac{\widehat{\partial AMG}}{\partial a}, \frac{\widehat{\partial AMG}}{\partial b}\right)'. \tag{15}$$

Taking the square root, we obtain the estimated standard error. However, our regression model for the trend and cyclical variation in elasticity includes market-good-specific fixed effects, which do not have well-defined variances. To proceed, we regard $\hat{a}$ (embedded in these fixed effects) as fixed and only consider the randomness of $\hat{b}$ in practice. Similarly, we only consider the randomness of the cyclical variation in elasticity relative to its trend when computing the standard error of the cyclical variation in markup relative to its trend.