Seven million demand elasticities

Jordan Rosenthal-Kay James Traina Uyen Tran*

November 7, 2021

Abstract

The household's price elasticity of demand is a key input to many economic models' construction of markups and the assessment of consumer surplus. We measure the price elasticity of demand for around 14,000 products by region-year using retail scanner data. In all, we estimate over 7.5 million demand elasticities. We find that the distribution of these elasticities is stationary over time. However, we document substantial spatial heterogeneity in consumers' price sensitivity: consumers in the largest markets are the most price elastic. As demand elasticities are a key input into the measurement of markups, our results suggest that any conclusions that markups are rising in retail markets must be driven by assumptions on conduct.

Keywords: Demand estimation, markups, economic geography

*Rosenthal-Kay: University of Chicago, jrosenthalkay@uchicago.edu; Traina: University of Chicago and Chicago Booth School of Business; Tran: Chicago Booth School of Business. 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.

1 Introduction

This paper reports facts about the demand elasticities for every good in every market and year in the Nielsen Retail Measurement Services (RMS) retail scanner data. We have two key findings: one, that the distribution of demand elasticities across all consumption goods measured in the Nielsen data has not changed over the 2006-2017 period. Moreover, the distribution of time trends in demand elasticities within each good-market is centered at zero and has little spread. Two, despite little variation within or across elasticities over time, there is substantial spatial heterogeneity within goods across markets: consumers in the largest markets are the most price elastic, but there is no relationship between population growth and trends in the demand elasticity time-series. The regions where preferences became more price-inelastic are confined to a handful of major metropolitan areas, including New York, Los Angeles, and San Francisco.

Recent literature documents increasingly concentrated product markets and rising corporate profits. See Berry et al. (2019) for a review. Whether these trends harm consumers depends on the nature of competition and the extent to which firms can raise prices above marginal costs. Consequently, a suite of studies have sought to measure price markups using off-the-shelf quantitative models from industrial organization. These range from from specific studies, e.g., Grieco et al. (2021) for automobiles, or Ganapati and McKibbin (2021) for pharmaceuticals, to economy-wide estimates: e.g., Anderson et al. (2020) use cost data to estimate markups in the retail sector.

Demand elasticities are a key input to the measurement of markups and any assessment of consumer welfare. The more price-inelastic demand is, the more a monopolist can raise price above marginal cost. Inferring markups through demand estimation requires a measure of firms' own price elasticities, cross-price elasticities for many-product firms, firms' market shares, and assumptions on the nature of competition. With these objects, researchers can invert firms' first-order conditions and recover markups.

This work documents that own-price elasticity of demand for products in the RMS data has not changed from 2006 to 2017. Instead, we find the distribution to be remarkably stable, with fairly modest spread. Our results run counter to work finding increasingly inelastic demand elasticities in the retail scanner data driven by niche consumption for a small set of product categories (Brand, 2021).[1] However, Neiman and Vavra (2021) demonstrate that markups need rise even as demand becomes more inelastic due to entry.

Previous work uses the demand-estimation procedure of Berry et al. (1995) (hereon, BLP) to estimate demand elasticities at the module (product category) level. BLP allows for consumer heterogeneity by allowing for random coefficients on consumer demographics in the underlying utility model, meaning that the researcher may infer changes in demand elasticities and markups from secular demographic changes. We mitigate these concerns by estimating demand for the universe of products in the retail scanner data at the region-year level using a log-linear estimation equation and standard set of instrumental variables. We use an empirical Bayes shrinkage technique to correct these estimates for measurement error and impose the prior of downward-sloping demand, so that measurement error and sign changes do not contaminate our estimates of the time-series and spatial properties of the distribution.

Our approach follows the large-scale estimation of preference parameters in the industrial organization, trade, and urban literature. For example, Hitsch et al. (2019) and DellaVigna and Gentzkow (2019) estimate demand elasticities at the brand-store level in the scanner data to understand price dispersion within and across establishments. Broda and Weinstein (2006) estimate elasticities of substitution using trade data to construct ideal price indices. Handbury and Weinstein (2010) and Handbury (Forthcoming) construct regional price indices for goods in the scanner data. We depart from this literature by studying the preference parameter estimates themselves, rather than use them as inputs to construct markups or price indices. Instead, we closely follow Tran (2021), who estimates the effect of broadband exposure on brand-county-year demand elasticities. Our work complements Anderson et al. (2020) who use cost data for retailers to study markups across years and markets in the retail industry: like them, we find substantial spatial variation, but little temporal variation in our estimates of demand elasticities.

The rest of the paper is organized as follows: in Section 2 we discuss the data, demand estimation, and procedure to correct for measurement error. In Section 3, we present evidence that demand elasticities have stayed flat over time but vary over space. Section 4 concludes.

[1] Brand (2021) estimates the distribution of own price elasticities for nine product-cateogries: Fruit drinks, soup, cookies, pizza, ice cream, entrees, yogurt, fruit, and light beer. See Döpper et al. (2021) for an expansion of this work to a larger set of product categories.

2 Data and empirical approach

2.1 Data

Our analysis uses the Nielsen Retail Measurement Services data, otherwise knows as the RMS or scanner data. The RMS data is store-level data on quantities, prices, and promotions, covering almost 40,000 stores. The selection of retailers ranges from gas stations to major retailer chains and covers over 50% of spending at all groceries and drugstores.

The data measures sales in terms of quantities and (quantity-weighted) prices at the week level. The data only record prices for positive transactions, and subsequently we omit observations for which the sales volume is zero. Following standard practice the industrial organization and marketing literature, we aggregate universal product code (UPC) sales to the brand level (e.g., Coke), and use this as the definition of product. Each product is uniquely nested in a module: a product category, e.g., sugar. Our data include around 14,000 unique products. In all, we estimate demand elasticities for around 7.5 million unique product-region-years in the data.[2] We use the Nielsen Designated Market Area (DMA) as our definition of region. These regions include major metropolitan areas like New York and Los Angeles, and smaller regions like Juneau, Alaska and Zanesville, Ohio. In all, the Nielsen DMAs partion the country into 209 regions.

2.2 Demand estimation

We estimate a simple log-linear demand system and omit cross-price effects. While BLP estimates allow for flexible substitution patterns, estimation is notoriously numerically unstable, and implementation procedures can often converge to local optima (Knittel and Metaxoglou, 2014). Our estimates, by contrast, are simple and computationally tractable two-stage-least-squares estimates.

We denote quantities q and prices p. For a product j in module m sold at store s in market d observed in week w in quarter t in year y, we estimate log-linear regressions,

$$\log q_{jsw} = \alpha_{jdy} \log p_{jsw} + \theta_{jdty} + u_{jst}, \tag{1}$$

[2]This is about 24% of the possible combination of all product-region-years. It is a small fraction as not all goods are sold in all markets in all years.

where θ_{jdty} are brand-market-quarter-year fixed effects. The variation used to estimate α_{jdy} then is cross-store and week data within a DMA-quarter. The coefficient on price, α_{jdy}, is the price elasticity of demand for product j in region d in year y. Prices and quantities are endogenously determined through market forces, so an OLS regression of quantities on prices cannot yield a consistent estimate of α_{jdy}. To correct for this, we form a shift-share price instrument in the spirit of Hausman (1996). We instrument log prices with the log average price across other stores in the same chain but in different markets. Conditional on our fixed effects, which account for brand-level seasonality, the assumption needed to consistently estimate α_{jdy} is that changes in the pricing decisions that apply to the entire chain are orthogonal to local demand conditions for a given product. This exclusion restriction is consistent with evidence of uniform pricing across stores despite substantial variation in demand elasticities, as documented in DellaVigna and Gentzkow (2019). Under this standard assumption, the interpretation our estimates are that $\hat{\alpha}_{jdy}$ represents the own-price short-run elasticity of demand for a given product-region.

2.3 Elasticity correction

Estimates $\hat{\alpha}_{jdy}$ from the two-stage-least-squares (TSLS) estimation of (1) are extremely noisy. We jointly correct for measurement error and impose the prior that demand slopes downwards with an empirical Bayes procedure. Our procedure uses information from the elasticity distribution within a module-year as a prior to update each individual brand-market-year elasticity estimate. Empirical Bayes models are computationally tractable approximations to hierarchical Bayes models with prior parameters set to their highest likelihood values.

Empirical Bayes corrections for noisily estimated measures is common in the literature. For example, hierarchical models have been used to better measure teacher value-added (Kane and Staiger, 2008; Chetty et al., 2014), hospital fixed effects for clinical outcomes (Chandra et al., 2016), and even for demand elasticities using the RMS data (DellaVigna and Gentzkow, 2019; Hitsch et al., 2019; Brand, 2021). Brand (2021) argues that due to growth in sales and the number of products, earlier demand elasticity estimates may be less precisely estimated than those estimated in more recent years. Failure to account for changing estimate precision over time (and space: as regions grow, more varieties may enter the market) may cause the

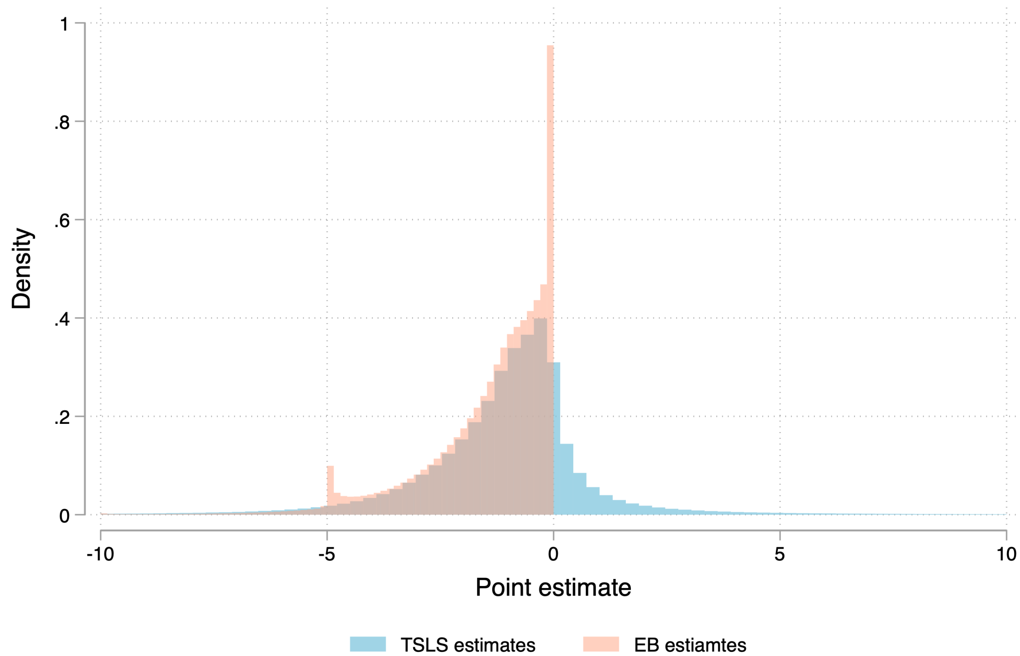


Figure 1: Distribution of OLS and empirical Bayes (EB) estimates of the distribution of product-market-year demand elasticities, $\hat{\alpha}_{jdy}$.

researcher to overstate the time trend for a given elasticity.

Our underlying hierarchical model is,

$$\begin{aligned} \hat{\alpha}_{jdy} \mid \alpha_{jdy} &\sim \; N(\alpha_{jdy}, \sigma_{\alpha}^2) \\ \alpha_{jdy} &\sim \; N_{(-10,0)}(\mu_{my}, \sigma_{\mu}^2) \end{aligned}$$

where in the second level of the hierarchy, we truncate the elasticity distribution over products within a module to make sure demand curves slope downward. The use of a truncated normal allows us to stay within the exponential family of distributions and derive a closed-form expression for the mean of the posterior. In short, as a truncated normal is conjugate with a normal likelihood, the posterior distribution for $\alpha_{jdy} \mid \hat{\alpha}_{jdy}$ is also truncated normal. See Appendix A for details on the estimation procedure.

The empirical Bayes procedure is essentially a signal extraction problem that uses information from estimates from other products within a module-market-year to correct for measurement error and moreover impose the prior from economic theory that demand elasticities are less than or equal to zero.[3] The truncation at -10 has little effect in practice. Less than 0.2% of the IV estimates report an elasticity less than -10. We prefer our truncated-normal empirical Bayes approach to the shrinkage estimators in e.g., DellaVigna and Gentzkow (2019), which combines normal-normal hierarchical model with ad-hoc rules to winsorize the data, dropping estimates whose coefficients or standard errors are too large. In contrast, our procedure allows us to retain all of the data, with minimal bunching in the tails of the distribution of posterior estimates.

Figure 1 displays the histogram of estimated product-region-year demand elasticities before and after the Empirical Bayes' correction. The distribution of TSLS estimates is left-skew and centered below zero, though there is some positive mass, and has very long tails. The EB distribution is by construction shifted left and has considerable mass around zero, due to a nontrivial amount of precisely estimated positive demand elasticities. The truncation at -10 has no visible effect; there is no excess mass piling up to the right of the truncation point. The median of the sales-weighted EB estimates is -2.13, though the distribution is left skew and its mean is -2.60.

3 Results

3.1 Distribution

The demand elasticity time-series Figure 2 displays the time series for the middle 80% of the distribution of demand elasticities, weighted by their sales share. The median elasticity of about -2 stays almost constant over time, becoming slightly more inelastic in 2017. Almost all the other percentiles of the distribution remain constant in time as well.

Almost one quarter of the elasticities in our data are measured *above* -1, indicating *inelastic* demand. This is inconsistent with monopoly pricing, wherein the profit-maximizing monopolist always prices along the elastic portion of the demand curve.

[3] While in theory, demand may slope upwards for Giffen and Veblen goods, evidence for Giffen goods is limited to extremely poor parts of the world (Jensen and Miller, 2008), and the data do not contain products suspected to have the Veblen property, like yachts and luxury cars.

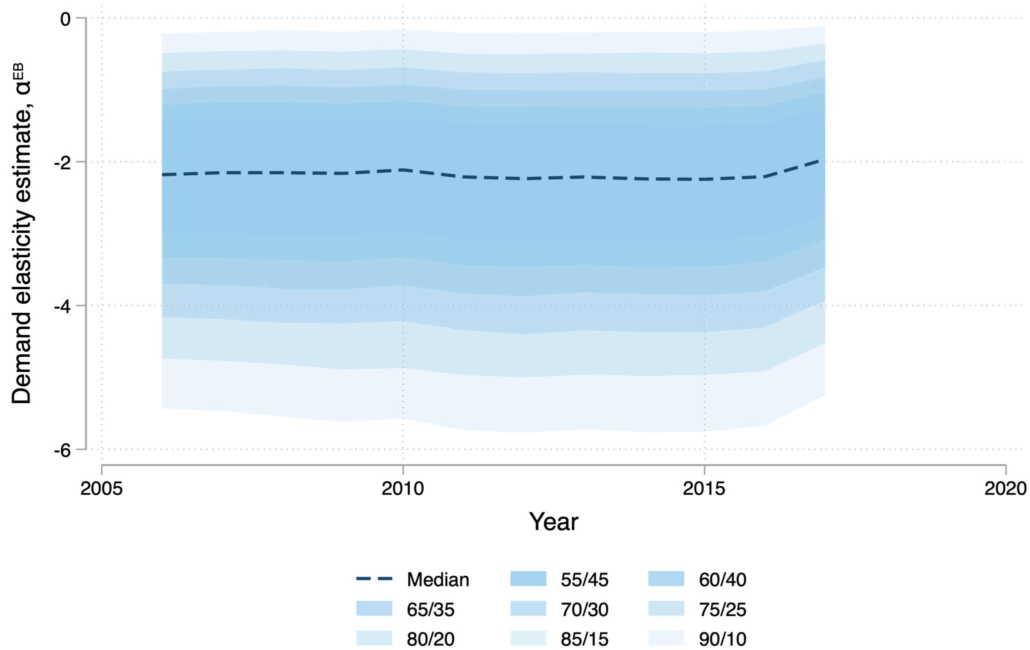


Figure 2: The time series of the distribution of demand elasticities estimates, weighted by revenue. The light-blue colored bands indicate percentiles of the distribution.

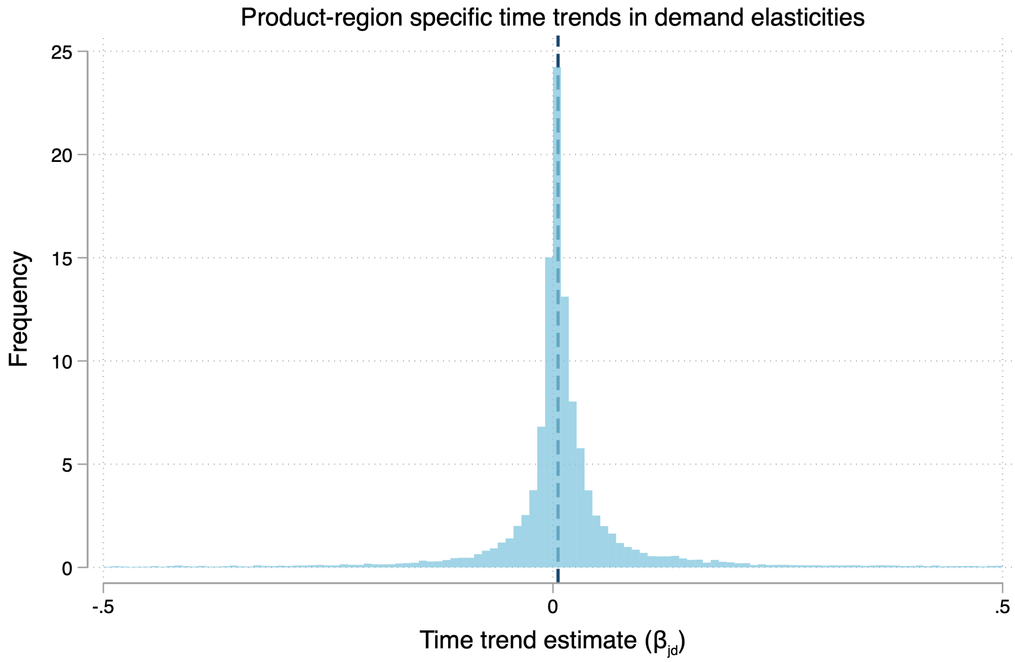
Product-specific trends We estimate product-region time trends through a linear regression,

$$\log(-\hat{\alpha}_{jdy}) = \gamma_{jd} + \beta_{jd} \times y + \upsilon_{jdy}. \tag{2}$$

The transformation of the regressand means that the unique product-region coefficient β_{jd} has the interpretation of the average annual percent change in the demand elasticity. Moreover, we flip the sign on the elasticities, so that $\beta_{jd} < 0$ indicates demand becoming *more inelastic*.

We estimate product-region unique time trends by first differencing (2) and recovering the fixed effects, $\hat{\beta}_{jd}$. This ameliorates the need to estimate an extrodinary number of fixed effects to account for product-region specific intercept terms. Figure 3 displays the middle 98% of the distribution weighted by sales volume.

The median of the distribution (dashed line) is to the right of zero: more than half the estimated elasticities are trending more elastic in time. The distribution is fairly tightly estimated around the mean; the middle 50% of the distribution estimates an annual percent change between -0.7% and 3%.



.

Figure 3: The distribution of individual time trends. The dashed line represents the median of the distribution.

3.2 Spatial variation in elasticities

Figure 4 shows the relationship between the sales-weighted average demand elasticity and market population. Though in the cross-section, demand is most elastic in the largest markets, there is no relationship between population trend and the trend in the sales-weighted average demand elasticity.

Figure 5 maps the median sales-weighted average demand elasticity time trend by market. The figure shows that downwards trends – i.e., demand becoming more inelastic – is confined to a handful of regions in the country, most of them major metropolitan areas.

4 Conclusion

In this paper, we reported evidence that the household's price elasticity of demand for different has not changed over time, using information from nearly 14,000 products across 11 years in over 200 markets in the United States. However, while there has been no change in preference parameters, we find there is substantial spatial heterogeneity. Larger markets have more inelastic demand. While this relationship holds in the cross-section, we find there is no rela-

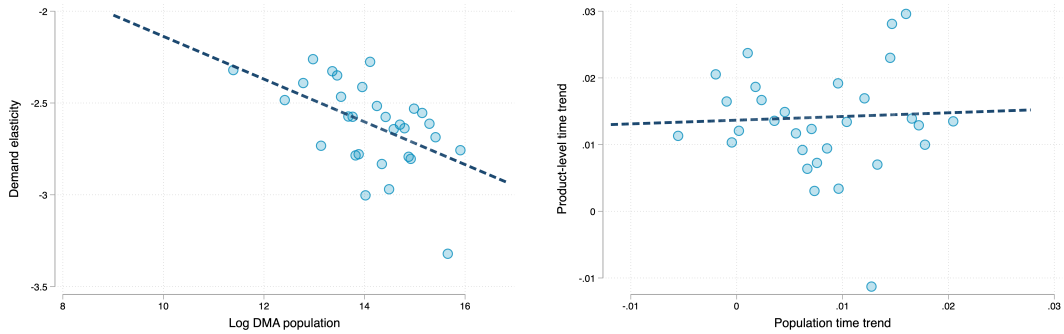


Figure 4: Correlations between region population and demand elasticities. Left: cross section, Right: time trends

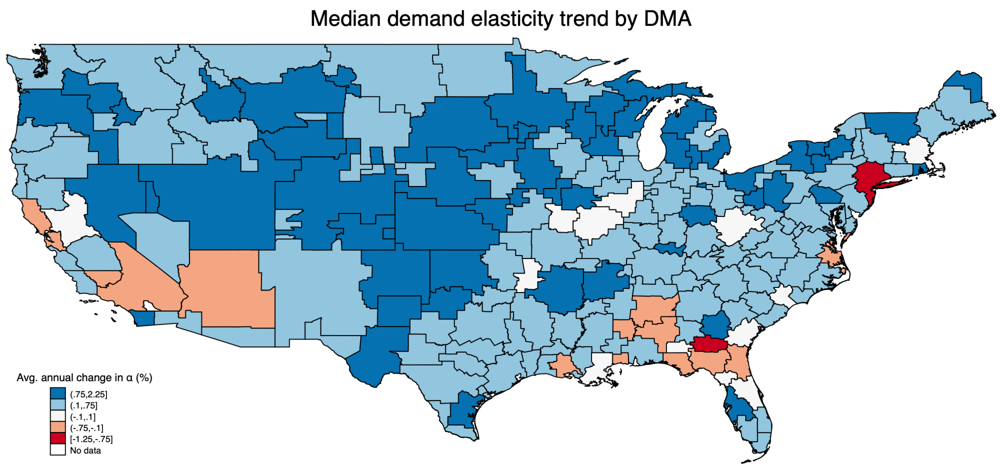


Figure 5: Correlations between region population and demand elasticities. Left: cross section, Right: time trends

tionship between population growth and changes in demand parameters. Areas where demand became more inelastic over the study period are confined to a handful of major metropolitan areas.

Our results suggest that trends in unobservables, recovered from economic models, in which preference parameters are a key input, are driven by other features of the data. In particular, trends in markups or consumer surplus must come from assumptions on the nature of competition and changes in market shares and cross-price elasticities. Nearly a quarter of our estimated elasticities fall above -1, which is inconsistent with monopoly pricing, suggesting the standard assumption of differentiated Bertrand competition may fail in practice.

References

Eric Anderson, Sergio Rebelo, and Arlene Wong. Markups across space and time. *Working paper*, 2020.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4), 1995.

Steven Berry, Martin Gaynor, and Fiona Scott Morton. Do increasing markups matter? lessons from empirical industrial organization. *Journal of Economic Perspectives*, 33(3), 2019.

James Brand. Differences in differentiation: Rising variety and markups in retail food stores. *Working paper*, 2021.

Christian Broda and David E. Weinstein. Globalization and the gains from variety. *The Quarterly Journal of Economics*, 121(2):541–585, 2006.

Amitabh Chandra, Amy Finkelstein, Adam Sacarny, and Chad Syverson. Health care exceptionalism? performance and allocation in the us health care sector. *American Econoimc Review*, 106(8), 2016.

Raj Chetty, John N. Friedman, and Jonah E. Rockoff. Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Econoimc Review*, 104(9), 2014.

Stefano DellaVigna and Matthew Gentzkow. Uniform pricing in u.s. retail chains. *The Quarterly Journal of Economics*, 134:2011–2084, 2019.

Hendrik Döpper, Alexander MacKay, Nathan H. Miller, and Joel Stiebale. Rising markups and the role of consumer preferences. *Working paper*, 2021.

Sharat Ganapati and Rebecca McKibbin. Markups and fixed costs in generic and off-patent pharmaceutical markets. *Working paper*, 2021.

Paul L. E. Grieco, Charles Murry, and Ali Yurukoglu. The evolution of market power in the us auto industry. *Working paper*, 2021.

Jessie Handbury. Are poor cities cheap for everyone? non-homotheticity and the cost of living across u.s. cities. *Econometrica*, Forthcoming.

Jessie Handbury and David E. Weinstein. Goods prices and availability in cities. *Review of Economic Studies*, 82(1):258–296, 2010.

Jerry A. Hausman. Valuation of new goods under perfect and imperfect competition. *The Economics of New Goods*, 1996.

Günter J. Hitsch, Ali Hortaçsu, and Xiliang Lin. Prices and promotions in u.s. retail markets: Evidence from big data. *NBER Working Paper*, 2019.

Robert T. Jensen and Nolan H. Miller. Giffen behavior and subsistence consumption. *American Econoimc Review*, 98(4), 2008.

Thomas J. Kane and Douglas O. Staiger. Estimating teacher impacts on student achievement: An experimental evaluation. *NBER working paper 14607*, 2008.

Christopher R. Knittel and Konstantinos Metaxoglou. Estimation of random-coefficient demand models: Two empiricists' perspective. *Review of Economics and Statistics*, 96(1), 2014.

Carl N. Morris. Parametric empirical bayes inference: Theory and applications. *Journal of the American Statistical Association*, 78(381):47–55, 1983.

Brent Neiman and Joseph Vavra. The rise of niche consumption. *Working paper*, 2021.

Esteban Rossi-Hansberg and Chang-Tai Hsieh. The industrial revolution in services. *NBER working paper*, (25968), 2021.

Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter. Diverging trends in national and local concentration. *NBER Macroeconomics Annual 2020*, 35:115–150, 2020.

Uyen Tran. High-speed internet and price elasticities. *Working paper*, 2021.

A Shrinkage estimation for demand elasticities

Recall the hierarchical model is,

$$\hat{\alpha}^{OLS} \mid \alpha \sim N(\alpha, \sigma_\alpha^2)$$
$$\alpha \sim N_{(-\infty,0)}(\mu, \sigma_\mu^2)$$

Our empirical Bayes estimate of α, $\hat{\alpha}^{EB}$ is defined as the expectation of $\mathbb{E}[\alpha \mid \hat{\alpha}^{OLS}]$. The associated posterior density has the form,

$$f(\alpha \mid \hat{\alpha}^{OLS}) \propto f(\hat{\alpha}^{OLS} \mid \alpha)\pi(\alpha)$$

where $f(\hat{\alpha} \mid \alpha)$ is the normal likelihood of observing the OLS estimate given the underlying parameter, and $\pi(\alpha)$ is our prior on the parameter. As the product of two normal densities is also normal, it follows that the product of a normal distribution with a truncated normal is also truncated normal. We write the parameters of the posterior distribution θ and τ. The posterior truncated normal has the same location (mean) and scale (variance) parameters as had the prior been normal, since the use of the truncated distribution simply removes part of the support and rescales the posterior. Thus we can apply the standard normal likelihood and normal prior empirical Bayes shrinkage procedure to recover these parameters,

$$\theta = \frac{\hat{\sigma}_\alpha^{-2}}{\hat{\sigma}_\alpha^{-2} + \hat{\sigma}_\mu^{-2}}\hat{\alpha}^{OLS} + \frac{\hat{\sigma}_\mu^{-2}}{\hat{\sigma}_\alpha^{-2} + \hat{\sigma}_\mu^{-2}}\hat{\mu}$$
$$\tau^2 = \hat{\sigma}_\alpha^2 \frac{\hat{\sigma}_\mu^{-2}}{\hat{\sigma}_\alpha^{-2} + \hat{\sigma}_\mu^{-2}}$$

Finally, we recover the mean of the posterior distribution $f(\alpha \mid \hat{\alpha}^{OLS})$ by using the Inverse Mills ratio, common to selection-correction procedures, which provides an analytic expression for the moments of the truncated normal distribution,

$$\hat{\alpha}^{EB} = \mathbb{E}[\alpha \mid \hat{\alpha}^{OLS}] = \theta - \tau \frac{\phi((-10-\theta)/\tau) - \phi(-\theta/\tau)}{\Phi((-10-\theta)/\tau) - \Phi(-\theta/\tau)}$$

where ϕ and Φ are the PDF and CDF for a standard normal distribution.

Following Morris (1983) We form $\hat{\mu}$ by taking the weighted mean of $\hat{\alpha}$ within a module-region-year, where the weights are $1/(se(\hat{\alpha}) + se(\hat{\mu}))$. As $\hat{\mu}$ appears on both the left and right hand side, we repeatedly estimate $\hat{\mu}$ until convergence starting with a guess of $se(\hat{\mu}) = 0$. In practice, we run this procedure for 4 iterations per estimator. In most cases, it converges in fewer, though in a handful of cases, the iteration error upon completion is larger than our tolerance of 1e-5, though seldom exceeding 1e-4.