

# Shopping, Demand Composition, and Equilibrium Prices

Lukas Nord\*

European University Institute

This Version: June 29, 2022

See latest version [here](#).

**Abstract.** This paper studies the implications of equilibrium interactions between the shopping behavior of heterogeneous households and retailers' price posting. Heterogeneous consumption baskets along the income distribution and higher shopping effort of the poor imply that retailers face different price elasticities depending on their customer base and charge higher markups for goods disproportionately consumed by richer households. First, I formalize this mechanism in a standard model of frictional product markets and derive testable predictions on the shape of price distributions as a function of search effort. Second, I confirm these predictions using supermarket scanner data from the US. Third, I embed the simple framework into a rich incomplete markets model featuring heterogeneous households with non-homothetic preferences as well as endogenous price distributions for multiple varieties. The calibrated model shows that equilibrium interactions between shopping effort, demand composition, and price posting double the contribution of shopping effort to expenditure inequality relative to previous findings. The model also implies that the response of markups to aggregate income shocks crucially depends on the incidence of the shock along the income distribution.

*Keywords:* Household Heterogeneity, Expenditures, Price Search, Markups.

*JEL:* D10, D31, E21, E30, L11

---

\*lukas.nord@eui.eu. I am indebted to my advisors, Árpád Ábrahám and Russell Cooper, for their continuous guidance and support. This work benefited greatly from repeated discussions with Annika Bacher, Edouard Challe, Alessandro Ferrari, Axelle Ferriere, Philipp Grübener, Wolfram Horn, and Alex Monge, as well as comments by my discussant Tsung-Hsien Li and workshop participants at the EUI, IIES, and QMUL. All remaining errors are mine. Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

# 1 Introduction

Macroeconomic research on inequality often uses households' expenditure to proxy for differences in consumption and ultimately welfare, e.g. when assessing the consequences of rising income or wealth inequality. The approximation relies on the assumption that the cost of consumption is identical across households. In reality, this assumption is violated for at least two reasons: First, rich and poor households do not buy the same basket of goods. Second, low-income households are able to exploit price dispersion for identical products, shop for bargains and reduce the price they pay for any given basket. Understanding how this behavior affects the cost of consumption across households is essential to interpret the real inequality implied by any distribution of nominal income, wealth, or expenditures. Previous research has shown both substitution across products and shopping for cheaper prices to be quantitatively important, but has studied each margin in isolation and has taken posted prices as given.

To account for interactions between shopping, substitution, and prices posted it is important to consider the behavior of heterogeneous households in equilibrium with retailers' price setting. From a retailer's perspective, systematic differences in households' consumption baskets lead to heterogeneity in demand composition across goods, i.e. a separation of demand from buyers at different income levels into different varieties. Differences in shopping behavior yield heterogeneous price elasticities of demand across households since the more any buyer searches for cheaper prices the more likely she is to purchase at a cheaper offer. The correlation between households' shopping effort and their consumption baskets implies that retailers face overall higher elasticities of demand for goods in the basket of low-income households. Their best response is to post lower prices by reducing their markups for goods bought by low-income households. Through heterogeneity in consumption baskets, differences in shopping effort can therefore have an equilibrium effect on the prices offered across varieties and not only a direct effect on the prices paid for any given good. In this way shopping effort has additional redistributive effects through the posted market price of consumption.

This paper considers the equilibrium effects of heterogeneity in shopping effort and demand composition on offered prices. I provide evidence on the mechanism outlined above by studying how the shape of price distributions changes across products with the shopping effort of buyers both in theory and in the data. Studying its implications, I show that accounting for effects on offered prices more than doubles the contribution of shopping effort to explaining expenditure inequality, reducing real inequality through the posted price of consumption. This is because demand composition varies across products and

retailers can differentiate price setting across households based on their consumption baskets. In addition, I find that a similar mechanism can have negative consequences in response to aggregate income shocks. If income losses are concentrated among a subset of households they affect the composition of aggregate demand over time. In response to a decline in earnings among low-income households and the ensuing reduction in their consumption, retailers shift the focus of their price setting towards high-income agents and can find it optimal to increase prices, leading to a further decline in real resources.

The analysis proceeds in four steps: First, I formalize the mechanism by focusing on retailers' price setting problem in a standard model of price dispersion under search frictions. In this framework, I derive analytical predictions for how the moments of price distributions co-move with equilibrium shopping effort. Second, I conduct an empirical test of the theoretical predictions providing evidence from US micro data. Third, I develop a quantitative model featuring heterogeneous households choosing their consumption baskets and shopping effort, as well as endogenous price distributions for multiple varieties. I calibrate the model to capture salient features of expenditure inequality and price dispersion. Fourth and finally, I employ the model to highlight the consequences of equilibrium prices responding to households' shopping behavior and demand composition.

To formalize the main mechanism of the paper, I build on a standard model of frictional product markets in the spirit of Burdett and Judd (1983). Buyers have to search for price quotes and for every purchase they make draw either one or two offers from the equilibrium distribution of posted prices. Higher shopping effort increases the probability that a buyer observes two prices for his purchase simultaneously. Fixing a price before meeting any buyers, retailers trade off higher margins per sale against undercutting alternative offers when buyers observe a second price simultaneously. I show analytically that the mean of the ensuing equilibrium price distribution can be decomposed into marginal costs and an average margin, the latter decreasing in equilibrium shopping effort. The model nests two special cases: Bertrand competition if all buyers observe two prices simultaneously and a perfect monopoly if no buyer observes a second price. Shopping effort shifts the equilibrium outcome between these two extremes, increasing the competition among retailers. An additional analytical result shows that as price distributions gradually shift from buyers' maximum willingness to pay (monopoly at zero shopping effort) to marginal cost (Bertrand competition if all buyers observe two prices), their skewness monotonically increases in the shopping effort exerted and is independent of all other parameters of the model. This result presents the skewness of price distributions as an ideal candidate for an empirical test of the mechanism.

For the empirical analysis, I use supermarket scanner data from the Nielsen Consumer Panel made available via the Kilts Center for Marketing at Chicago Booth. The dataset provides granular information on households' grocery purchases and prices paid at the barcode level. Based on the theoretical results, I focus on how the skewness of price distributions co-moves with equilibrium shopping effort across varieties. Defining local price distributions at the level of a barcode in line with the literature, I show that their skewness is decreasing in the share of total expenditure for a given barcode stemming from households at the top of the expenditure or income distribution or having more members employed. This finding is in line with the theoretical predictions, as high-spending, high-income, and employed households are known to exert lower shopping effort and pay higher prices for identical goods. It provides strong supportive evidence for a response of price distributions to shopping effort in line with search theory.

To study equilibrium interactions between households' behavior and retailers' price setting, I introduce a model economy featuring heterogeneous households and endogenous price distributions for multiple varieties. Households differ in their idiosyncratic labor income and beginning of period wealth. They decide on overall spending vs. savings and allocate their total expenditures across varieties. Consumption baskets vary systematically across households due to non-homothetic preferences in the spirit of Handbury (2021). Choosing how much to increase the probability of observing two prices for every purchase they make, households trade off a reduction in average prices paid against the disutility from exerting shopping effort. Equilibrium price distributions are determined endogenously as the solution to a Burdett and Judd (1983) market for each variety. The model matches salient features on the composition of expenditures across households and price dispersion across varieties and is able to reproduce the untargeted distribution of expenditures in the data.

I use the model as a framework to study the equilibrium implications of interactions between shopping behavior and demand composition. It allows me to highlight the consequences of differences in demand composition (i) across varieties and (ii) over time. To capture the first, I employ a decomposition of expenditures to both the model and the Nielsen data and use the model to adjust for differences in posted prices due to heterogeneous shopping effort. For the second, I simulate aggregate income losses for different groups of households and show how offered prices and especially markups respond to changes in households' consumption and shopping choices.

Focusing first on the implications of heterogeneity in demand composition across varieties, I decompose the price a household pays into the contribution of the direct effect of

shopping (difference between price paid and average price of a barcode across households), substitution across varieties (difference between average price of a barcode vs. average price of close substitutes), and expenditures absent both margins. In the data, the bottom quintile of the expenditure distribution pays 4% less than average among a range of close substitutes due to substitution across barcodes, and an additional 1.5% less due to paying less for the same barcode (shopping). The top quintile pays 3.5% above average due to substitution and 0.5% more due to shopping. The model replicates this decomposition of expenditures. It further allows me to adjust for the equilibrium effect of shopping through posted markups. In the baseline decomposition, differences in posted markups are picked up by the average price of a variety and hence allocated to the substitution term. Accounting for the equilibrium effect on offered prices, shopping reduces the cost of consumption for the bottom quintile by 3.5% and increases it for the top quintile by 2%. Equilibrium effects more than double the contribution of shopping to expenditure inequality compared to the partial equilibrium effect of paying more or less for the same good at given offered prices. The insurance provided through posted prices and markups exists because non-homothetic preferences and the implied segregation in demand reduce the externality households exerting low shopping effort impose on others through their effect on retailers' price setting. Despite the findings outlined above, a significant extent of this externalities remains due to the non-zero overlap in households consumption baskets. Allowing households to draw from price distributions targeted to their individual shopping effort, I show that the bottom of the expenditure distribution could reduce their cost of consumption by an additional 15% if good markets were perfectly separated by household type, while the top of the expenditure distribution would pay up to 25% more for the same basket of goods.

To study the consequences of shifts in demand composition over time, I simulate the response of the economy to an aggregate income loss. I consider three scenarios: First, reducing the labor earnings of all households in the economy by 3%. Second, causing the same decline in aggregate income but concentrating the losses only at the top quartile of the labor earnings distribution. Third, concentrating the same aggregate income loss at the bottom quartile of the distribution. The model yields a pro-cyclicality of retail markups that is decreasing in the share of income losses allocated to the bottom of the distribution. In response to an income loss affecting households equally all agents increase their shopping effort, retailers face more competition, and markups decline. Markups decline by more if the losses are concentrated at the top quintile, but increase if losses affect only the lowest earning households. The key to understand this result is a shift in demand composition: To set prices, retailers weigh households' shopping effort by their

share in demand. An income loss concentrated at the bottom of the distribution forces low-income households to reduce consumption and shifts demand composition towards high-income (low-shopping) households, possibly *decreasing* demand weighted shopping despite an *increase* in effort among low-income households. In response to income losses at the top both effects go in the same direction, high-income households searching more for cheap prices and demand being reallocated towards low-income (high-search) agents. The varying cyclicalities of markup responses to income shocks has implications for inequality: High-income households can obtain partial insurance if hit by an aggregate income loss through a decline in the retail markups they face. Low-income households can be hit twice, by a loss in income and a subsequent loss in purchasing power as markups increase in response to an income shock at the bottom of the distribution. This implication of shifts in demand composition over time is in contrast to the findings outlined above for heterogeneity in demand composition across varieties, which dampens inequality among agents.

The remainder of the paper is organized as follows: Section 1.1 discusses the relation to previous literature. Section 2 formalizes and provides empirical evidence on the main mechanism. Section 3 outlines the quantitative model. Section 4 studies implications of heterogeneous shopping effort and differences in demand composition for inequality. Section 5 concludes.

## 1.1 Related Literature

The paper relates to a large and growing empirical literature on expenditure inequality and its drivers, with recent contributions by Aguiar and Bils (2015) or Coibion et al. (2021), sampled e.g. in Attanasio and Pistaferri (2016). Most closely related is the strain of the literature focussing on the consequences of shopping behavior and prices paid, including the seminal contributions by Aguiar and Hurst (2005, 2007) as well as subsequent work by e.g. Broda et al. (2009), Griffith et al. (2009), Aguiar et al. (2013), Nevo and Wong (2019), or Droste et al. (2019). While these papers focus on measuring the direct effect of shopping (paying more or less for a given good) empirically, I combine data and a quantitative model to jointly measure the contributions of shopping and substitution to expenditure inequality and highlight the equilibrium effect of shopping on offered prices.

The theoretical literature on price search in the goods market goes back to the seminal contributions of Butters (1977), Varian (1980), and especially Burdett and Judd (1983). I build on the latter, which has been widely applied in macroeconomic research on topics including price dynamics (Burdett and Menzies 2018) or product differentiation (Albrecht et

al. 2021; Menzio 2021). Building on a similar mechanism of profit margins changing with shopping effort, Kaplan and Menzio (2016) use the framework to show how differences in shopping effort between employed and unemployed households can generate self-fulfilling unemployment fluctuations while Alessandria (2009) applies it to explain fluctuations in relative prices across countries. Two papers study the consequences of search frictions in the goods market for heterogeneous households: Arslan et al. (2021) introduce price search effort into a life-cycle consumption-savings problem but take the distribution of offered prices and the return to search as given. Pytka (2018) additionally endogenizes a single equilibrium price distribution. Both papers focus on the consequences of price search for life-cycle inequality and households' response to idiosyncratic income shocks in a stationary economy, but abstract from heterogeneity in consumption baskets – and therefore from equilibrium effects of shopping across varieties – as well as the response to aggregate income losses. This paper is the first to develop an equilibrium framework with rich household heterogeneity and endogenous price distributions for multiple varieties to study the equilibrium consequences of shopping and heterogeneity in demand composition across goods and over time.

A vast literature exists on heterogeneity in households' consumption baskets, dating back as far as the formulation of Engel's Law in 1857. Most important for the present paper is the strand focussing on non-homotheticities at the barcode level and consequences for e.g. the distribution of gains from product innovations (Jaravel 2019), cost of living differences between poor and rich households over the business cycle (Argente and Lee 2021) or across cities (Handbury 2021), as well as firms' decision of which products to offer (Faber and Fally 2022). Non-homotheticities at this low level are often interpreted as substitution along a quality margin. This margin of adjustment is studied e.g. in Bils and Klenow (2001) and Bisgaard Larsen and Weissert (2020) and its consequences in Jaimovich et al. (2019) or Ferraro and Valaitis (2022). None of these papers considers interactions with households' shopping effort or price dispersion.

The paper also extends the empirical literature on retail prices and markups. Closely related contributions on price dispersion include the seminal work by Kaplan and Menzio (2015) and Kaplan et al. (2019), who provide evidence on the structure of price distributions but do not consider their co-movement with demand composition across products. Stroebl and Vavra (2019) find retail prices to respond pro-cyclically to local variations in house prices, driven mostly through changes in markups, and attribute this pattern to empirically observed changes in shopping behavior. Anderson et al. (2020) find markups paid to covary positively with proxies for local income, driven by local differences in products bought not local variations in markups charged for the same products. Their

findings are well in line with the mechanism proposed in this paper.

Most closely related is the work of Sangani (2022), providing a complement to this paper. Using a combination of the Nielsen panel and data on wholesale prices, Sangani (2022) shows empirically that markups are on average higher for the goods bought by high-income households and rationalizes this finding with a shopping model. I provide direct empirical evidence on the mechanism, testing theoretical predictions on the relationship between shopping and the shape of price distributions. Combining the single-variety model of Burdett and Judd (1983) with stylized household heterogeneity, Sangani (2022) studies implications of increasing income inequality for the rise in aggregate markups. The focus of my study is the feedback between equilibrium prices and inequality. Hence, I develop a model with rich household heterogeneity in the tradition of Bewley (1977) and Aiyagari (1994), featuring non-homothetic preferences and endogenous price distributions for multiple varieties.

## 2 The Mechanism:

### Equilibrium Shopping and Posted Price Distributions

The main mechanism proposed in this paper is best understood by focussing on the distribution of offered prices for a single variety and how it changes with the shopping effort exerted by the buyers of this variety. I begin by outlining the price posting problem of a single-variety retailer in a market with consumer search and derive the ensuing equilibrium price distribution, building on the work of Burdett and Judd (1983) and Pytka (2018). To provide evidence on the relationship between shopping and posted prices, I characterize analytically how moments of the offered price distribution respond to changes in equilibrium search effort and show that predictions from search theory align well with empirical observations from the Nielsen Consumer Panel micro data.

#### 2.1 Retailers' Price Setting with Consumer Search

Consider the market for a single variety  $j$ , which is produced at homogeneous marginal cost  $\kappa_j$ , and for which consumers have a maximum willingness to pay  $\bar{p}_j$ . The variety is sold by a continuum of homogeneous retailers of measure one. The demand side of the market consists of a continuum of households indexed by their type  $i$ , with  $\lambda_i$  being the distribution over types. A type  $i$  household consumes a quantity  $c_j^i \geq 0$  of variety  $j$ , which he splits into a measure  $c_j^i$  of infinitesimal purchases. The market for the variety is subject to incomplete information as in Burdett and Judd (1983) and Pytka (2018): For



each purchase he makes, the household observes either one or two price postings, drawn at random from the equilibrium distribution of offered prices  $F_j(p)$ . The probability of observing two price draws for any given purchase is determined by the household's shopping effort and given as  $s^i \in [0, 1]$ . Shopping effort in this framework is best interpreted as the intensity with which households search for a second price observation. If a household observes only one price, he makes a purchase as long as this price is below the maximum willingness to pay  $\bar{p}_j$ . If he observes two prices simultaneously, the household purchases at the lowest offer below  $\bar{p}_j$ .

Retailers post prices for variety  $j$  to maximize their profits, taking expectations over which type of household they will meet in the market and how likely these households are to see a second price offer simultaneously. The total profits of a retailer posting price  $p$  are given by

$$\pi_j(p) = \underbrace{C_j}_{\text{demand per retailer}} \underbrace{\left[ \int \frac{\lambda^i c_j^i}{C_j} [(1 - s^i) + s^i 2(1 - F_j(p))] di \right]}_{\text{sales per demand (market share)}} \underbrace{(p - \kappa_j)}_{\text{profit per sale}}$$

where  $C_j = \int \lambda_i c_j^i di$  is total demand for variety  $j$  and  $\frac{\lambda^i c_j^i}{C_j}$  the fraction of demand (fraction of purchases) stemming from households of type  $i$ . In words, profits are given as the margin per sale  $(p - \kappa_j)$  times total demand per retailer ( $C_j$ ) times the market share. To determine the effect of the posted price on the latter, the retailer considers the likelihood with which any buyer he meets in the market observes a second price quote simultaneously: With probability  $\frac{\lambda^i c_j^i}{C_j}$  he meets a type  $i$  household, and with probability  $s^i$  this household has a simultaneous second price observation conditional on being type  $i$ . In the latter case the retailer only makes a sale if his price offer is lower than the second quote, which conditional on posting price  $p$  occurs with probability  $(1 - F_j(p))$ .<sup>1</sup>

As profits are linear in sales and by the law of large numbers, the problem can be simplified to

$$\pi_j(p) = C_j [(1 - \bar{s}_j) + \bar{s}_j 2(1 - F_j(p))] (p - \kappa_j) \quad (1)$$

---

<sup>1</sup>The multiplication of the second term by 2 captures that the retailer can be either the first or second of two price observations.

where

$$\bar{s}_j = \int \underbrace{\frac{\lambda_i c_j^i}{C_j}}_{\text{demand composition}} s^i di \quad (2)$$

Demand weighted, average shopping effort  $\bar{s}_j$  is a sufficient statistic for the retailer to set a price. How households' shopping effort is taken into consideration by the retailer depends on their share in total demand, which from here on out I will refer to as *demand composition*.

Deciding on the price to post in this market, retailers trade off between two forces: On the one hand, a higher price increases the margin earned per sale  $(p - k_j)$ . On the other hand, it increases the probability to be undercut by a competitor  $F_j(p)$  and hence decreases demand at the extensive margin. For the second effect, the equilibrium search effort is key, which determines the likelihood that any given unit of demand observes a second price and therefore the likelihood any retailer has to compete for the purchase. Thereby, together with the endogenous price offer distribution  $F_j(p)$ , the equilibrium shopping effort  $\bar{s}_j$  determines the price elasticity of demand across retailers.

The setup leads to two immediate observations: First, price dispersion through the lens of this model can be interpreted as a distribution of markups over marginal cost  $\kappa_j$ . Second, for any non-degenerate equilibrium price distribution to exist retailers will have to be indifferent between posting a range of prices, i.e. the benefit of earning a higher margin per sale will have to be exactly offset by the loss in demand. For given  $k_j$ ,  $\bar{p}_j$ , and  $0 < \bar{s}_j < 1$  Burdett and Judd (1983) and Pytka (2018) show that a unique and continuous equilibrium distribution of offered prices  $F_j(p)$  exists with compact support  $[\underline{p}_j, \bar{p}_j]$ , where

$$F_j(p) = \begin{cases} 0 & \text{if } p < \underline{p}_j \\ 1 - \frac{1 - \bar{s}_j}{2\bar{s}_j} \frac{\bar{p}_j - p}{p - \kappa_j} & \text{if } p \in [\underline{p}_j, \bar{p}_j] \\ 1 & \text{if } p > \bar{p}_j \end{cases} \quad (3)$$

and

$$\underline{p}_j = \kappa_j + (\bar{p}_j - \kappa_j) \frac{1 - \bar{s}_j}{1 + \bar{s}_j}$$

The distribution of posted prices depends on the marginal cost  $\kappa_j$  and households' maximum willingness to pay  $\bar{p}_j$ , as well as demand weighted shopping effort  $\bar{s}_j$ , but is independent of total demand per retailer  $C_j$  as retailers are infinitesimal and only compete over their share in a total number of sales they take as given. To study the mechanism at the heart of this paper I will focus on how the offered distribution of prices responds

to changes in equilibrium search effort  $\bar{s}_j$ .

## 2.2 The Effect of Shopping on Posted Price Distributions

According to equation (2), changes in  $\bar{s}_j$  can be due to individual households adjusting their shopping effort  $s^i$  or due to differences in demand composition (e.g. across varieties), i.e. how retailers take each household's individual shopping effort into account. In equilibrium, individual shopping, demand composition, and hence demand weighted shopping effort  $\bar{s}_j$  will be determined by households' optimal behavior in response to the offered price distribution. In this section, I take  $\bar{s}_j$  as given and focus on retailers best response to any change in equilibrium search effort. Focussing on how retailers respond to changes in  $\bar{s}_j$  allows me to highlight the effect of differences in both households' shopping effort or demand composition on offered prices.

From equation (3) it is immediate that  $\frac{\partial F_j(p)}{\partial \bar{s}_j} \geq 0$ , i.e. a distribution with lower equilibrium shopping effort  $\bar{s}_j$  has first-order stochastic dominance over any distribution with higher equilibrium shopping effort and hence a greater probability to observe high posted prices. Given the analytical characterization of  $F_j(p)$ , the problem yields closed form expressions for the moments of the distribution. Results for the first three central moments are presented in Proposition 1.

**Proposition 1** *The mean  $\mu_j^F$ , standard deviation  $\sigma_j^F$ , and skewness  $\gamma_j^F$  of the offered price distribution  $F_j(p)$  for given  $\kappa_j$ ,  $\bar{p}_j$ , and  $0 < \bar{s}_j < 1$  can be derived as*

(i)

$$\mu_j^F = \kappa_j + \underbrace{(\bar{p}_j - \kappa_j) \frac{1 - \bar{s}_j}{2\bar{s}_j} \log \left( \frac{1 + \bar{s}_j}{1 - \bar{s}_j} \right)}_{\text{average posted margin}}$$

(ii)

$$\sigma_j^F = \sqrt{(\bar{p}_j - \kappa_j)^2 \left( \frac{1 - \bar{s}_j}{1 + \bar{s}_j} - \left( \frac{1 - \bar{s}_j}{2\bar{s}_j} \right)^2 \log \left( \frac{1 + \bar{s}_j}{1 - \bar{s}_j} \right)^2 \right)}$$

(iii)

$$\gamma_j^F = \frac{\frac{1 - \bar{s}_j}{4\bar{s}_j} \left( 1 - \left( \frac{1 - \bar{s}_j}{1 + \bar{s}_j} \right)^2 \right) - 3 \frac{(1 - \bar{s}_j)^2}{2\bar{s}_j + 2\bar{s}_j^2} \log \left( \frac{1 + \bar{s}_j}{1 - \bar{s}_j} \right) + 2 \left( \frac{1 - \bar{s}_j}{2\bar{s}_j} \right)^3 \log \left( \frac{1 + \bar{s}_j}{1 - \bar{s}_j} \right)^3}{\left( \frac{1 - \bar{s}_j}{1 + \bar{s}_j} - \left( \frac{1 - \bar{s}_j}{2\bar{s}_j} \right)^2 \log \left( \frac{1 + \bar{s}_j}{1 - \bar{s}_j} \right)^2 \right)^{\frac{3}{2}}}$$

**Proof.** *Follows immediately from equation (3) and the standard formulas for the first*

three central moments of any continuous distribution. ■

Focussing first on the mean of offered prices, Proposition 2 shows that the average price posted is increasing in marginal cost  $\kappa_j$  and maximum willingness to pay  $\bar{p}_j$ , but decreasing in equilibrium shopping effort  $\bar{s}_j$ . Figure 1a illustrates this result graphically.

**Proposition 2** *The mean of the offered price distribution  $\mu_j^F$  for  $0 < \bar{s}_j < 1$  is increasing in marginal cost  $\kappa_j$  and maximum willingness to pay  $\bar{p}_j$  but decreasing in equilibrium search effort  $\bar{s}_j$ , i.e.*

$$(i) \frac{\partial \mu_j^F}{\partial \kappa_j} > 0, \quad (ii) \frac{\partial \mu_j^F}{\partial \bar{p}_j} > 0, \quad (iii) \frac{\partial \mu_j^F}{\partial \bar{s}_j} < 0$$

**Proof.** *Follows immediately from taking first derivatives of  $\mu_j^F$ . ■*

The equilibrium effect of shopping on posted prices operates through changes in profit margins over marginal cost  $\kappa_j$ , which are strictly decreasing in equilibrium shopping effort  $\left(\frac{\partial(\mu_j^F - \kappa_j)}{\partial \bar{s}_j} < 0\right)$ . We find the origin of this effect in retailers' profit equation (1): Higher shopping effort  $\bar{s}_j$  makes it more likely that the average buyer a seller meets also observes a second price, and hence tilts sellers' tradeoff between higher margins and retaining demand in favor of the latter. The higher the shopping effort the higher is the implied elasticity of demand: The more buyers in the market search for lower prices the more competition sellers face and the lower their (optimal) margins per sale.

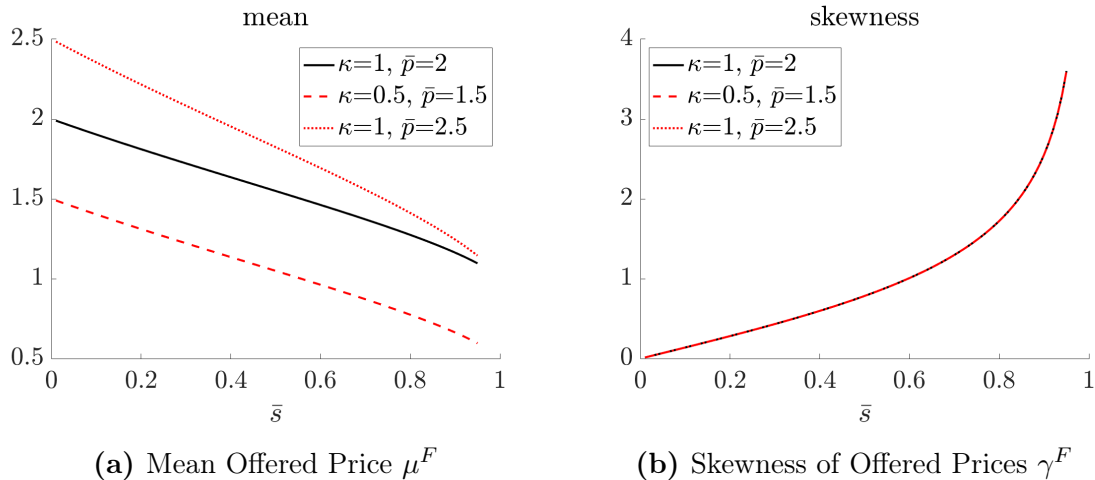
In the limit, the setup approaches two well known special cases: If all buyers observe two prices simultaneously ( $\bar{s}_j = 1$ ), sellers solve a Bertrand competition problem and  $\mu_j^F = \kappa_j$ , i.e. all sellers post marginal cost. If no buyer observes two prices simultaneously ( $\bar{s}_j = 0$ ), all sellers have a monopoly for any buyer they meet and hence all sellers extract buyers maximum willingness to pay such that  $\mu_j^F = \bar{p}_j$ .<sup>2</sup> Households' shopping effort determines a market's position between these two extremes by regulating how much competition sellers face for any given unit of demand they meet.

The result captures the mechanism at the heart of this paper. If goods are bought predominantly by low-searching households, retailers put more weight on their shopping effort, face a low  $\bar{s}_j$  and optimally post higher markups. Taking into account retailers' optimal price posting and differences in demand composition therefore yields equilibrium effects of heterogeneity in shopping effort on posted prices.

While the mean and standard deviation of the posted price distribution are not determined by shopping effort alone, focussing on the skewness provides a sharp, empirically testable prediction. As outlined in Proposition 3 and highlighted graphically in Figure

---

<sup>2</sup>This point has been made by Pytka 2018 for a similar setup with at most 2 price observations.



**Figure 1:** Moments of the Offered Price Distribution  $F_j(p)$

1b, the skewness of the offered price distribution  $\gamma^F$  is a function only of the equilibrium shopping effort  $\bar{s}_j$  and independent of  $\kappa_j$  as well as  $\bar{p}_j$ . Furthermore, it is strictly increasing in  $\bar{s}_j$ .

**Proposition 3** *The skewness of the offered price distribution  $\gamma_j^F$  for  $0 < \bar{s}_j < 1$  is increasing in equilibrium search effort  $\bar{s}_j$ , but independent of marginal cost  $\kappa_j$  and maximum willingness to pay  $\bar{p}_j$ , i.e.*

$$(i) \frac{\partial \gamma_j^F}{\partial \kappa_j} = 0, \quad (ii) \frac{\partial \gamma_j^F}{\partial \bar{p}_j} = 0, \quad (iii) \frac{\partial \gamma_j^F}{\partial \bar{s}_j} > 0.$$

**Proof.** *Follows immediately from taking first derivatives of  $\gamma_j^F$ .* ■

The intuition behind this finding is that  $\kappa_j$  and  $\bar{p}_j$  determine the interval of possible prices – no seller will ever post prices below marginal cost or above households maximum willingness to pay as this would yield weakly negative profits – while shopping effort  $\bar{s}_j$  determines the distribution of prices within this interval. The skewness picks up any shift in mass of a distribution over a given interval. If shopping effort is high, prices are concentrated close to marginal cost and the distribution exhibits a lot of mass close to  $\kappa_j$  at the left of the interval and a long right tail (high skewness). If shopping effort decreases, more prices are posted closer to buyers' maximum willingness to pay  $\bar{p}_j$ , the mass of the distribution shifts to the right, and the skewness decreases.

Before turning to an empirical test of Proposition 3, I refer to the appendix for robustness of the main findings with respect to two potential extensions of the model: Retailer entry combined with fixed cost of operating, as well as heterogeneous marginal cost. Appendix A.1 shows that the introduction of fixed cost of operating and entry leaves the solution derived above for the equilibrium distribution of offered prices unchanged.

Appendix A.2 shows that under reasonable calibrations the introduction of heterogeneous marginal cost leaves markups decreasing and the skewness increasing in shopping effort. For the remainder of the paper, I focus on the version with homogeneous marginal cost as the empirically relevant case, as marginal cost in this model should be interpreted as wholesale cost and wholesale price differentiation among retailers within a geographic area is prohibited in the US under the (federal) Robinson-Patman Act as well as various, more commonly applied state level legislations.<sup>3</sup>

## 2.3 Evidence on Demand Composition and Price Distributions

The finding of Proposition 3 makes the relationship between the skewness of price distributions and the shopping effort in a given market a natural candidate for an empirical test of the effect of shopping on offered prices. If posted prices respond to shopping effort, we should expect the skewness of any price distribution to be an increasing function in the demand weighted shopping effort of households buying the good. For an empirical test I rely on differences in shopping effort induced by heterogeneity in demand composition across goods.

To test for an empirical relationship between the shape of price distributions and equilibrium shopping effort, I make use of the Nielsen Consumer Panel micro data waves for 2007-2019. The dataset provides detailed information on the grocery purchases of approximately 60,000 US households per wave, containing both quantities purchased and prices paid for every store visit at the barcode (UPC) level. In addition, the data contains annual information on households' demographic characteristics such as income, household composition, or employment, as well as the place of residence. Further information on the dataset is provided in Appendix B.

A price distribution consists of all transactions observed for a given variety  $j$ , within a region  $r$  and time period  $t$ . In line with the literature (Kaplan and Menzio 2015), I define a variety as a barcode (UPC), a region as a Scantrack Market Area (SMA) and the time period to be a quarter.<sup>4</sup> The price associated with a transaction is defined as the total amount paid less of coupon values, divided by the number of quantities purchased. To

---

<sup>3</sup>See e.g. Nakamura (2008).

<sup>4</sup>The choice for what definition of a region and which time period to consider has to trade off between two forces: A narrow definition ensures that any variation in prices can be allocated to (and exploited by) search frictions, while it also reduces the number of price observations per distribution and hence makes the analysis more noisy. For the ensuing analysis to be valid it is not necessary that households have access to every price within a region, but only that the distribution of prices is identical for any subregion. As Scantrack Markets are defined by industry professionals as target regions for marketing purposes, retailers pricing can be assumed to be sufficiently similar within such regions to ensure identical price distributions throughout.

control for outliers, I drop all transactions for which prices are zero or negative. For the baseline analysis, I consider all price distributions containing at least 25 transactions and compute the skewness of each distribution weighting individual price observations with household weights and quantities purchased.

To proxy for equilibrium shopping effort, I make use of the impact of demand composition on  $\bar{s}$ . Theory suggests that to set their prices, retailers should consider the shopping effort of households weighted by their share in overall demand for the variety they sell. Therefore, for each barcode (UPC) I compute the annual, national expenditure shares stemming from different groups of households, sorted by their shopping effort.<sup>5</sup> I consider separately the five quintiles of the (equivalence scale adjusted) expenditure distribution, four bins of household income, as well as the number of non-employed household heads. Based on results from the literature (see e.g. Aguiar and Hurst 2005; Broda et al. 2009; Kaplan and Menzio 2016) as well as the findings outlined in Section 4.1 below, high-spending, high-income, or households with fewer non-employed members are exerting less shopping effort and pay higher prices for identical goods. To be in line with the predictions from theory, the skewness of price distributions should therefore be decreasing in the expenditure share coming from high-spending or high-income households, but increasing in the share of demand from households with more non-employed members.

To test for the relationship between skewness and shopping derived from theory, I regress the skewness of a price distribution  $(j, r, t)$  on the national expenditure shares of each household group  $g$ , for variety  $j$  in the respective year  $y(t)$ , excluding one household group as a baseline. The specification is given in equation (4). To control for local economic conditions and product characteristics, I include time-region fixed effects  $(\mu_{r,t})$  as well as fixed effects for Nielsen-defined product modules  $(\theta_m)$ .<sup>6</sup> All regressions are weighted by the total amount of expenditures of the price distribution.

$$skew_{j,r,t} = \theta_m + \mu_{r,t} + \sum_{g=2}^G \beta_g share_{j,g,y(t)} + \varepsilon_{j,r,t} \quad (4)$$

Results are reported in Table 1. Standard errors in parentheses are clustered at the barcode-year level. In line with theoretical predictions, the skewness of price distributions

---

<sup>5</sup>I use annual and national shares as Nielsen is representative at this level. Using aggregate rather than local shares is justified by the evidence on uniform price setting of large retail chains across locations (see e.g. DellaVigna and Gentzkow 2019), who should therefore take national rather than local demand composition into consideration for their price setting.

<sup>6</sup>I do not control for UPC fixed effects to exploit variation in expenditure shares across different varieties. Nielsen-defined product modules are the first level of aggregation above UPCs and capture product characteristics at a granular level. Examples of product modules in Nielsen are e.g. “fresh apples” or “fresh oranges” for different categories of fresh fruits.

is monotonically decreasing in the share of expenditure stemming from higher spending households (column (1)). The coefficients should be interpreted as the relative skewness compared to the omitted baseline group. For column (1): If a barcode is bought entirely by households in the fifth quintile of the expenditure distribution, the skewness of its price distribution decreases by 3.4 relative to a barcode bought entirely by the first expenditure quintile. All differences vs. the baseline group are statistically significant at the 99%-level. The finding is robust to measuring expenditure shares conditional on the household head being between age 25-65, to account for spending patterns of student and retiree households (column (2)). Similar findings pertain by income group, again conditioning on working age households (column (3)), while skewness is monotonically increasing in the number of non-employed household heads (column (4)). All findings are well in line with the prediction from Proposition 3, conditional on shopping effort decreasing in expenditures, income, and increasing for non-employed household members.

**Table 1:** Demand Composition and the Skewness of Price Distributions

	by expenditures (1)	by expenditures (working age) (2)		by income (working age) (3)		by non-employment (4)
quintile 2	-1.638*** (0.188)	-1.467*** (0.170)	30k-60k	-0.136 (0.123)	1 non-emp	0.864*** (0.092)
quintile 3	-2.309*** (0.189)	-2.076*** (0.167)	60k-100k	-0.824*** (0.124)	2 non-emp	1.011*** (0.172)
quintile 4	-3.067*** (0.172)	-2.582*** (0.155)	>100k	-0.820*** (0.087)		
quintile 5	-3.412*** (0.153)	-3.007*** (0.142)				
FE module	795	795		795		795
FE Y:Q:SMC	3584	3584		3584		3584
Observations	3,026,551	3,026,404		3,026,404		3,026,551
Adjusted R <sup>2</sup>	0.034	0.033		0.031		0.031

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

In Appendix C.1, I report further robustness with respect to how the skewness of price distributions is measured, exemplarily for the specification of column (1) of Table 1. Table 5 reports results without using weights in the regression, computing the skewness based on unweighted price observations or based on household weights only. All findings are qualitatively and quantitatively robust to using alternative weighting schemes. The pattern of skewness decreasing in expenditure is also robust when using Kelly’s measure of skewness, which is less sensitive to outliers.<sup>7</sup> As Table 6 shows, the patterns remain if considering only price distributions with at least 50 or 100 transactions, and are quantitatively even larger for distributions with more frequent purchases.<sup>8</sup>

<sup>7</sup>Note here that the units of the coefficients are not comparable for Kelly’s measure of skewness and hence no statements can be made about the relative magnitude of the results in column (2) of Table 5.

<sup>8</sup>This is despite controlling for time-region and product module fixed effects, i.e. despite a limited role



All specifications considered suggest the same conclusion: The skewness of price distributions is negatively associated with the share of expenditure stemming from low shopping households. I take this as strong evidence in favor of a response of posted price distributions to demand weighted shopping effort and proceed to studying its implications below.

### 3 Equilibrium of Shopping and Demand Composition

Households' consumption baskets and shopping effort are determined jointly with retailers decision on offered prices in equilibrium. To study implications of the mechanism outlined in the previous section, I therefore develop an equilibrium framework featuring heterogeneous households with non-homothetic preferences and a choice of shopping effort, as well as endogenous price distributions for multiple varieties. The model allows for a decomposition of expenditure inequality that takes into account equilibrium effects of shopping on offered prices and provides a laboratory to study counterfactuals and the consequences of aggregate shocks. In this section, I outline the model and discuss the quantitative implementation and calibration. Results are presented in Section 4.

#### 3.1 Households with Non-Homothetic Preferences and Shopping

Households are heterogeneous in their beginning of period asset holdings  $a$  and idiosyncratic labor productivity  $z$ , which evolves exogenously according to a first order Markov process. They decide jointly on their future asset holdings  $a'$ , quantities consumed of each variety  $\{c_j\}_{j=1}^J$  and shopping effort  $s$ . Households' decision problem can be split into two stages. In a first stage the household divides his resources between savings and total expenditure  $e$ , i.e. the first stage solves

$$V(z, a) = \max_{e, a'} U(e) + \beta \mathbb{E}_{z'|z} V(z', a')$$

$$\text{s.t. } e + a' \leq (1 + r)a + zw$$

---

for composition effects.

The utility of expenditure  $U(e)$  is determined in a second stage by the household's allocation of consumption across varieties and choice for shopping effort, solving

$$U(e) = \max_{s, \{c_j\}_{j=1}^J} u(\mathcal{C}) - v(s, e)$$

$$\mathcal{C} = \left[ \sum_{j=1}^J ((e)^{q_j} c_j)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

$$\sum_{j=1}^J p_j(s) c_j \leq e$$

where  $u(\cdot)$  are households preferences over a non-homothetic consumption aggregator  $\mathcal{C}$  and  $v(s, e)$  is the disutility of exerting shopping effort. The consumption aggregator  $\mathcal{C}$  is a non-homothetic CES function and a simplified version of the preference specification in Handbury (2021). For given total expenditure  $e$  and shopping effort  $s$ , it yields a demand system across varieties that can be defined in terms of expenditure shares  $\omega_j$  where the optimal allocation satisfies

$$\frac{\omega_j}{\omega_k} = e^{(1-\sigma)(q_k - q_j)} \left( \frac{p_j(s)}{p_k(s)} \right)^{1-\sigma}$$

The parameters  $\{q_j\}_{j=1}^J$  govern the expenditure elasticity of demand and for  $\sigma > 1$  (i.e. varieties being substitutes to each other) the relative expenditure share of variety  $j$  vs. variety  $k$   $\left( \frac{\omega_j}{\omega_k} \right)$  is increasing in total expenditures  $e$  iff  $q_j > q_k$ .

The price a household pays for any variety  $j$  is a decreasing function of his shopping effort. Households' optimal choice of shopping effort equalizes the marginal benefits of relaxing the budget constraint by a reduction in prices with the marginal disutility of effort such that

$$\underbrace{u' \left( \frac{e}{\mathbb{P}(s, e)} \right) \frac{e}{\mathbb{P}(s, e)^2} \frac{\partial \mathbb{P}(s, e)}{\partial s}}_{\text{marginal benefit of relaxing BC}} = \underbrace{-v_s(s, e)}_{\text{marginal disutility of shopping}}$$

where

$$\mathbb{P}(s, e) = \left[ \sum_{j=1}^J ((e)^{-q_j} p_j(s))^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

$$\frac{\partial \mathbb{P}(s, e)}{\partial s} = \frac{1}{\mathbb{P}(s, e)} \sum_{j=1}^J [(e)^{-q_j(1-\sigma)} (p_j(s))^{-\sigma} \underbrace{\frac{\partial p_j(s)}{\partial s}}_{\text{return to shopping}}] < 0$$

The relationship between prices and shopping effort, i.e. the return to shopping effort  $\frac{\partial p_j(s)}{\partial s} < 0$ , is an equilibrium object and depends on the distribution of offered prices. I focus on this relationship and equilibrium conditions in the goods market below.

### 3.2 Returns to Shopping and Equilibrium in the Goods Market

The distribution of offered prices for each variety  $F_j(p)$  is determined by retailers' optimal price setting as laid out in Section 2, given equilibrium search effort  $\bar{s}_j$ . Following Pytka (2018), the distribution of effective prices for a single purchase of variety  $j$  by a household exerting shopping effort  $s$  is given as

$$G_j(p|s) = (1-s)F_j(p) + s(1 - (1 - F_j(p))^2)$$

Under the assumptions that households split their total demand for each variety into a continuum of purchases and that every price observation is an i.i.d. random draw from  $F_j(p)$ , the average price paid per purchase of variety  $j$  is given by

$$p_j(s^i) = \mathbb{E}_j^G[p|s^i] = \mu_j^F - s^i (\mu_j^F - \mathbb{E}_j^F[\min\{p', p''\}]) \quad (5)$$

such that

$$\frac{\partial p_j(s^i)}{\partial s^i} = - (\mu_j^F - \mathbb{E}_j^F[\min\{p', p''\}]) = \text{const.} < 0$$

Hence, the two constants  $\mu_j^F$  and  $\frac{\partial p_j(s^i)}{\partial s^i}$  are sufficient statistics to capture the impact of the price distribution of variety  $j$  on households' behavior, i.e. all that households need to know to decide on their demand for each variety and their shopping effort.

Solving for an equilibrium in the goods market requires finding a fix-point in households' policies  $\{\{c_j(z, a)\}_{j=1}^J, s(z, a), a'(z, a)\}_{(z,a)}$ , the induced distribution of households across states  $\lambda(z, a)$  and aggregate demand as well as demand weighted shopping effort for each variety  $\{C_j, \bar{s}_j\}_{j=1}^J$  where

$$C_j = \int_{(z,a)} \lambda(z, a) c_j(z, a)$$

and

$$\bar{s}_j = \int_{(z,a)} \frac{\lambda(z, a) c_j(z, a)}{C_j} s(z, a)$$

while the distribution of households across states is the stationary solution to the law of motion

$$\lambda(z', a') = \int_{(z,a)} \lambda(z, a) Pr(z'|z) \mathbb{1}_{a'=a(z,a)}$$

In the full model, the type  $i$  of Section 2 is replaced by households state  $(z, a)$ .

### 3.3 Quantitative Implementation

To study the implications of equilibrium interactions between shopping effort and demand composition, I solve the model numerically. Households' spending-savings problem is solved by a version of the endogenous grid method (EGM) in the spirit of Carroll (2006) where  $U(e)$  is approximated numerically based on the optimal solution to the consumption allocation and shopping problem. At the aggregate, I apply Broyden's method to solve jointly for a fixed point in all  $\{\bar{s}_j\}_{j=1}^J$ . I calibrate the model at annual frequency. The calibration proceeds in three steps: I first calibrate the income process outside of the model, describe functional forms and set some parameters exogenously, and finally calibrate all remaining parameters to match targets on expenditure composition, price dispersion, and model aggregates.

#### Income Process

The process for idiosyncratic labor productivity is calibrated externally. To capture higher moments of income risk as reported e.g. in Guvenen et al. (2021), I follow Ferrier et al. (2022) and Mendicino et al. (2022) and assume an AR(1) process with innovations from a Gaussian mixture, i.e.

$$\log(z') = \rho \log(z) + \varepsilon$$
$$\varepsilon \sim \begin{cases} \mathcal{N}(\mu_{\varepsilon,1}, \sigma_{\varepsilon,1}^2) & \text{with probability } p \\ \mathcal{N}(\mu_{\varepsilon,2}, \sigma_{\varepsilon,2}^2) & \text{with probability } 1 - p \end{cases}$$

I discretize the process with 16 states for  $z$  following the method of Farmer and Toda (2017). The income process requires to calibrate 6 parameters,  $(\rho, p, \mu_{\varepsilon,1}, \sigma_{\varepsilon,1}^2, \mu_{\varepsilon,2}, \sigma_{\varepsilon,2}^2)$ . I impose  $\mu_2 = -\frac{p}{1-p}\mu_1$  to obtain mean zero innovations and calibrate the remaining parameters to match five moments of annual, equivalence scale adjusted, post-tax household labor earnings: The cross sectional variance of earnings, the standard deviation, skewness, kurtosis of annual earnings growth as well as the difference between the 90th and 10th percentile of annual earnings changes. Target values based on PSID data are obtained from De Nardi et al. (2020).<sup>9</sup> All targets are reported in Table 2 along with the model counterparts. The associated parameter values are  $\rho = 0.91$ ,  $\sigma_1 = 0.59$ ,  $\sigma_2 = 0.23$ ,  $p = 0.082$ , and  $\mu_1 = -0.57$ .

---

<sup>9</sup>For more information on how the target values are constructed see their Appendix A.3.

**Table 2:** Calibration Income Process

Targets (Annual)	Model	Data
Cross Sectional Variance (Levels)	0.61	0.57
Standard Deviation of Changes	0.33	0.33
Skewness of Changes	-0.99	-0.98
Kurtosis of Changes	10.6	10.3
P90-P10 of Changes	0.53	0.64

### Functional Forms and Exogenously Set Parameters

I assume standard CRRA preferences for  $u(\cdot)$  and a disutility of shopping effort depending on total expenditures  $e$  such that

$$u(c) = \frac{c^{1-\phi} - 1}{1-\phi} \quad \text{and} \quad v(s, e) = \psi_1 (e)^{\psi_2} \frac{s}{1-s}$$

Disutility of shopping depending on total expenditure captures in reduced form that households have to search more often for prices if they are making more purchases. In this spirit,  $\psi_2$  measures economies of scale in shopping effort.

The calibrated version of the model features three varieties, i.e.  $J = 3$ .<sup>10</sup> Varieties should be considered close substitutes and can be thought of as different barcodes within a Nielsen product module. Think of different varieties as types of apples, differing in their quality, with households substituting towards higher quality (high  $q_j$ ) varieties as they increase spending. I focus on this low level of differentiation across varieties as a significant degree of non-homotheticities can only be observed at this granular definition of a variety.<sup>11</sup> In line with the evidence on low-level elasticities of substitution sampled in Jaravel and Olivi (2021), I set  $\sigma = 2$ . Furthermore, I normalize  $q_2 = 0$  and  $\kappa_1 = 1$ .

The CRRA parameter is set to  $\phi = 2$  in line with the literature and the annual real interest rate to  $r = 0.02$ .

### Endogenously Calibrated Parameters

The remaining parameters to be calibrated are  $(\psi_1, \psi_2, \kappa_2, \kappa_3, \bar{p}_1, \bar{p}_2, \bar{p}_3, \beta, w, q_1, q_3)$ . I impose  $q_1 = -q_3$  and  $\bar{p}_j = a + b(\kappa_j - \kappa_1)$ . This leaves 9 free parameters for which I target 9 moments. Targets can be divided into three groups:

At the aggregate level, I target a wealth to income ratio of 3, an average retail markup of 1.45 following the results of Hall (2018), and normalize average expenditure in line

<sup>10</sup>Considering 4 or 5 varieties does not significantly alter results.

<sup>11</sup>See e.g. Jaravel (2019) and the findings presented in Appendix C.2.

with the Nielsen data.<sup>12</sup> While all parameters can influence all moments, the ones most closely linked to the aggregate targets are  $\psi_1$ , which governs average shopping effort and hence aggregate markups,  $\beta$  for the wealth to income ratio, and the wage rate  $w$  for mean expenditure.

**Table 3:** Targets and Model Fit

Target	Value	Model	Source
baskets overlap Q1 vs. Q5	0.6328	0.6301	Nielsen
$\Delta e$ substitution Q1 vs. Q5	0.072	0.071	Nielsen
$\Delta e$ shopping Q1 vs. Q5	0.02	0.021	Nielsen
$\text{mean}(CoV_j)$	0.1920	0.1927	Nielsen
$CoV_2 - CoV_1$	-0.0120	-0.0074	Nielsen
$CoV_3 - CoV_1$	-0.0201	-0.0240	Nielsen
average markup	1.45	1.4302	Hall (2018)
wealth/income	3	2.99	
mean expenditure	0.964	0.964	Nielsen

A second set of moments targets price dispersion across varieties. Targets for price dispersion are computed based on the same definition of a price distribution as in Section 2.3, i.e. pool transactions for a given barcode within a Scantrack region and a quarter. To account for differences in the average price across barcodes in the data, I focus on the coefficient of variation (CoV), i.e. the standard deviation of prices normalized by the mean. I target the (expenditure weighted) average CoV across all price distributions. In addition, I run the regression in equation (4) for the CoV, including on the right hand side the quintiles of the expenditure distribution and target the implied differences in the CoV across model varieties based on the endogenous demand composition (spending shares across quintiles) in the model. Targets for price dispersion interact most closely with the values for  $\kappa$  and  $\bar{p}$ .

The final set of moments contains targets on expenditure composition across households. To capture the (dis)similarity in households consumption baskets, I target the histogram overlap of expenditure shares  $\omega_j$  between the first and the fifth quintile of the expenditure distribution at the barcode level. Details on the construction of this target are provided in Appendix C.2. The barcode-level overlap between the consumption baskets at the bottom and top of the expenditure distribution is about 63%. In addition, I target the annual savings (as a share of respective total expenditure) of households at the bottom quintile of expenditures relative to the top quintile, due to (i) buying varieties that are cheaper on average (savings from substitution) and (ii) paying less for identical varieties (shopping).

<sup>12</sup>The target for mean expenditure is computed by dividing households annual expenditure in Nielsen by the mean and then truncating the distribution at the top at 4 times mean spending.

**Table 4:** Model Parameters

Parameter	Value
$J$	3
$\sigma$	2.0
$\phi$	2.0
$r$	0.02
$q_j$	$[-0.65 \ 0 \ 0.65]$
$\psi_1$	0.0035
$\psi_2$	0.05
$\kappa_j$	$[1 \ 1.04 \ 1.22]$
$\bar{p}_j$	$2.6 + 1.5(\kappa_j - \kappa_1)$
$\beta$	0.9335
$w$	0.68

The former amounts to about 7.1% and the latter to 2.1% of annual spending.<sup>13</sup> The difference in shopping identifies  $\psi_2$  while the overlap in consumption baskets and savings from substitution interact closely with relative expenditure elasticities  $q$  as well as  $\kappa$  and  $\bar{p}$  across varieties.

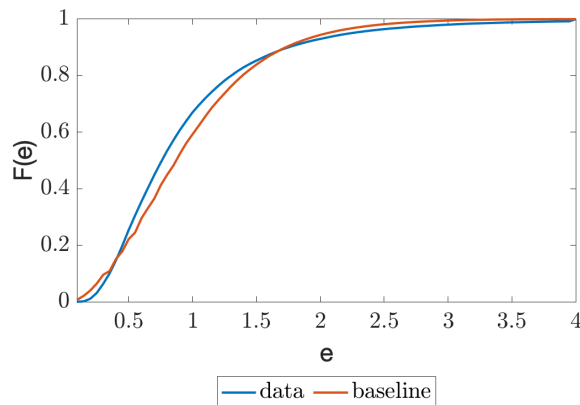
**Figure 2:** Expenditure Distribution - Model vs. Data

Table 3 unites all targets and presents the model fit. The model is able to match all targeted moments. Table 4 summarizes the calibrated parameter values. To provide validation for the calibrated model, Figure 2 plots the model implied distribution of expenditures (baseline) along with its counterpart from the Nielsen dataset. Note here that while I normalize average expenditure and match three moments of expenditure composition across households, no measure of dispersion in household expenditures is targeted in the calibration. Despite this, the model does remarkably well in capturing the empirically observed distribution of expenditures. I take this as confirmation that

<sup>13</sup>More details on these targets in Section 4.1.

the model provides a realistic framework for studying the equilibrium consequences of shopping and non-homotheticities in households' preferences.

## 4 Shopping, Demand Composition, and Inequality

With the calibrated model at hand, this section highlights implications of the equilibrium interaction of shopping and demand composition with retailers' price setting. I focus on implications for inequality among households and the insurance provided by adjustments in equilibrium prices. Section 4.1 studies the consequences of heterogeneity in demand composition across goods in a stationary economy, while Section 4.2 highlights implications of shifts in demand composition over time in response to aggregate income losses.

### 4.1 Equilibrium Prices and Expenditure Inequality

Due to heterogeneity in demand composition across varieties, shopping effort has consequences for the relative prices and especially markups retailers post for different items. It is their optimal response to demand weighted shopping effort varying with the composition of buyers. This section shows the consequences of shopping effort's interaction with heterogeneity in demand composition across varieties for the interpretation of expenditure inequality. I provide a decomposition of expenditures that can be applied to both the data and the model, show that the model fits the data well along this dimension, and finally use the model to adjust the decomposition for the equilibrium effect of heterogeneity in shopping on posted prices.

To decompose expenditure inequality in the data, I sort all barcodes (varieties) into groups of close substitutes, defining a group as all barcodes within a Nielsen product model measured in the same units.<sup>14</sup> Total annual expenditure  $e$  for a given household is the sum of spending over all barcodes in all groups. Spending per barcode is the average price paid by the household per unit of variety  $j$  in barcode group  $k$   $p_{jk}$  times the total quantity consumed  $c_{jk}$ . Both prices and quantities are normalized to the unit of measurement for any given good, e.g. for soft-drinks prices are per milliliters and quantities total milliliters consumed. Further define  $\bar{p}_{jk}$  as the average price paid for variety  $j$  across all households and  $\tilde{p}_k$  as the average of  $\bar{p}_{jk}$  within group  $k$ . As an example for the product module "fresh apples" measured in pieces,  $\tilde{p}_k$  is the average price per apple across all households and

---

<sup>14</sup>I group barcodes with the same unit of measurement to make their per unit prices comparable across products. Most product modules have one dominant unit of measurement and there is no systematic difference across households in the unit dimension as Appendix C.2 shows.



all types of apples,  $\bar{p}_{jk}$  the average price for one specific type of apple across households, and  $p_{jk}$  the average price one specific household pays for one specific type of apple.

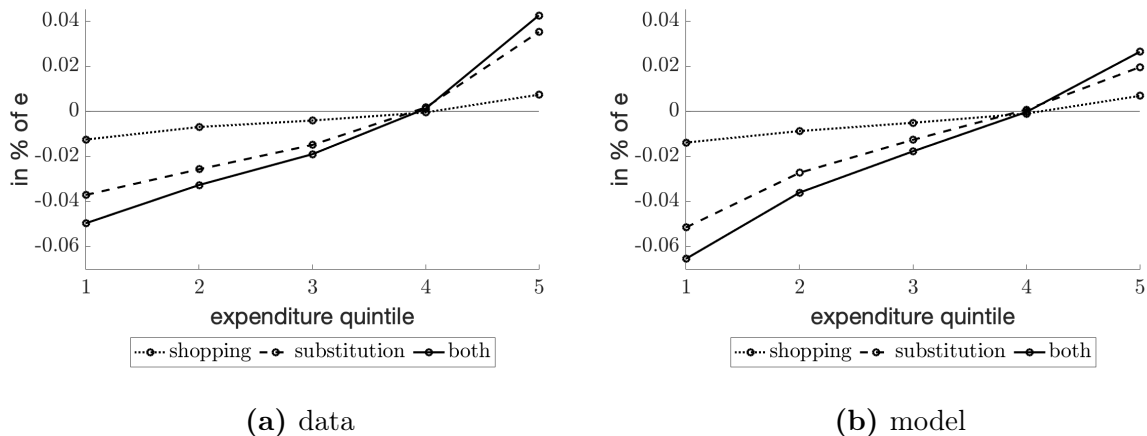
We can rewrite total expenditures, decomposing the price a household pays per unit of variety  $j$  such that

$$\begin{aligned} e &= \sum_k \sum_{j \in J_k} e_{jk} = \sum_k \sum_{j \in J_k} p_{jk} c_{jk} \\ &= \sum_k \sum_{j \in J_k} \underbrace{(p_{jk} - \bar{p}_{jk}) c_{jk}}_{\text{shopping (direct)}} + \underbrace{(\bar{p}_{jk} - \tilde{p}_k) c_{jk}}_{\text{substitution}} + \underbrace{\tilde{p}_k c_{jk}}_{\text{quantity}} \end{aligned}$$

The first term is the difference between what the individual household pays for the same variety relative to other households, which can be interpreted as the direct effect of differences in shopping behavior, i.e. for a given price distribution paying more or less for the same variety. The second term can be seen as the effect of substitution among similar goods (within a module/unit bin). The larger this term is, the more expensive products a household buys from a range of similar varieties. The literature (e.g. Argente and Lee 2021; Bisgaard Larsen and Weissert 2020; Jaimovich et al. 2019) has interpreted a higher price among close substitutes as households buying higher quality products. The last term summarizes households' counterfactual expenditure if there was no difference in shopping effort and substitution across close substitutes, i.e. if everyone pays the same price for all apples. The same decomposition can be applied to the model, where I consider only a single (representative) group  $k$  and all three varieties  $j$  are considered close substitutes within that group.

Figure 3a shows the results of the empirical decomposition by quintile of the expenditure distribution, expressed as a fraction of total expenditure. Households at the bottom of the distribution have about about 5.5% lower expenditure due to price differences within module/unit bins, of which 1.5pp. are due to the direct effect of shopping and 4pp. due to substitution across varieties. At the top of the distribution, price differences increase total spending by 4%, 3.5pp. of which due to substitution. In between, the contribution of both shopping and substitution is monotonically increasing in the expenditure of households. The magnitudes of the contribution of the (direct) effect of shopping and substitution are well in line with the findings e.g. Aguiar and Hurst (2007) or Bisgaard Larsen and Weissert (2020) obtain with alternative methodologies.

We can also interpret these findings in terms of their contribution to expenditure inequality. Inequality is often measured as a ratio between the top and the bottom of the distribution (e.g. Aguiar and Bils 2015). Define  $\tilde{e} = \sum_k \sum_{j \in J_k} \tilde{p}_k c_{jk}$  as a household's



**Figure 3:** Decomposition of Expenditures

counterfactual spending without any differences in shopping effort and consumption baskets. The results of Figure 3a imply that expenditure inequality measured as the ratio between spending of the top to bottom quintile of the expenditure distribution would be about 10% lower when measured in counterfactual expenditure  $\tilde{e}$  compared to true expenditure  $e$ .<sup>15</sup>

In Table 7 in Appendix C.3 I regress the contribution of shopping and substitution on households' expenditure to show that both are robust to controlling for income and household size as well as household and year-state fixed effects. To relate the lower price paid for identical varieties to household shopping effort, I show in a similar regression that the number of trips per purchase made, a common measure for shopping effort, are a decreasing function of household expenditure.<sup>16</sup>

Figure 3b shows that the model is able to reproduce the empirical patterns. It is important to note that only the differences for shopping and substitution between the lowest and highest quintile are included in the set of targeted moments. The model does well at reproducing the levels and shape of the empirical counterparts.

In the previous decomposition, the substitution term picks up differences in the posted price distributions across varieties. The model allows to further account for the equilibrium effect of shopping on offered prices. As Proposition 1 has shown, the effect of shopping on posted prices operates through the markups retailers set in response to differences in  $\bar{s}$ . While I do not observe markups in the data, the model allows me to

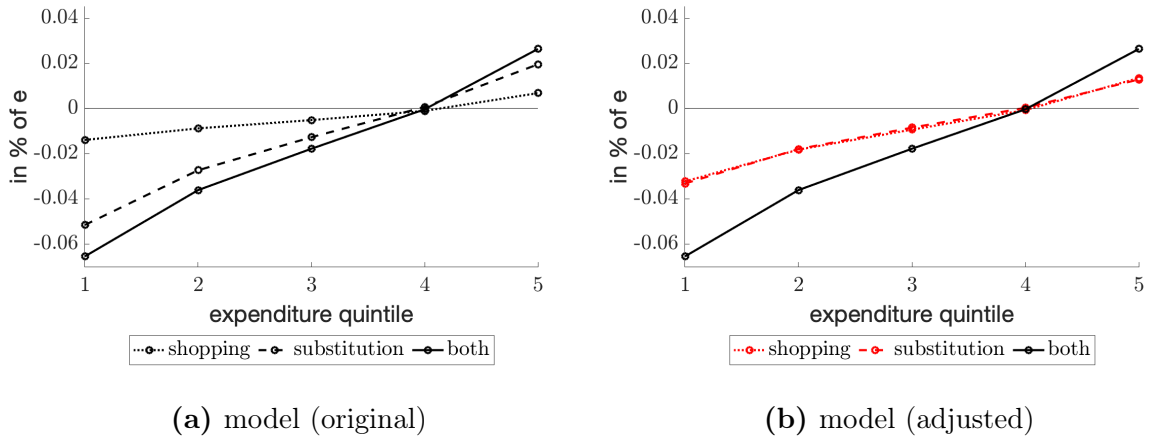
<sup>15</sup>Figure 12 in Appendix C.3 shows that this number increases to 12% when considering the bottom vs. top 10% of the expenditure distribution.

<sup>16</sup>See e.g. Aguiar and Hurst (2007) or Droste et al. (2019). Trips are normalized by the number of purchases to control for shopping needs. Kaplan and Menzio (2015) show that a reduction in prices is most effectively obtain by visiting more stores, or the same store more frequently, i.e. increasing the number of trips per purchase.

attribute differences in average markups across varieties to the equilibrium effect of shopping. The substitution effect is reduced to differences in marginal costs. The adjusted decomposition is given as

$$\begin{aligned}
e &= \sum_k \sum_{j \in J_k} \underbrace{(p_{jk} - \bar{p}_{jk})c_{jk}}_{\text{shopping (direct)}} + \underbrace{(\bar{p}_{jk} - \tilde{p}_k)c_{jk}}_{\text{substitution}} + \underbrace{\tilde{p}c_{jk}}_{\text{quantity}} \\
&= \sum_k \sum_{j \in J_k} \underbrace{(p_{jk} - \bar{p}_{jk})c_{jk} + ((\bar{p}_{jk} - \kappa_{jk}) - (\tilde{p}_k - \tilde{\kappa}_k))c_{jk}}_{\text{shopping (direct+equilibrium)}} + \underbrace{(\kappa_{jk} - \tilde{\kappa}_k)c_{jk}}_{\text{substitution}} + \underbrace{\tilde{p}_k c_{jk}}_{\text{quantity}}
\end{aligned}$$

Figure 4 presents the results for the adjusted decomposition alongside the original results from above. It shows that accounting for the equilibrium effect of differences in  $\bar{s}$  on offered prices the contribution of shopping to expenditure inequality more than doubles, making it equally important as differences in marginal costs across varieties. According to this decomposition, shopping alone through its direct and equilibrium effects can account for 5% of inequality in household expenditures.



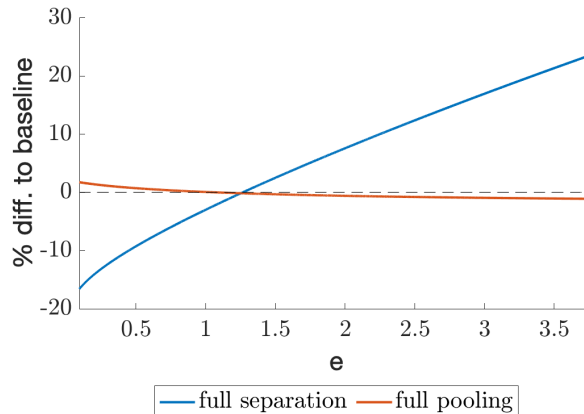
**Figure 4:** Decomposition of Expenditures - Equilibrium Effects

In combination with non-homotheticities in preferences and the ensuing heterogeneity in demand composition across varieties, shopping can provide significant insurance to households through its effect on equilibrium prices. The previous literature such as Aguiar and Hurst (2007), Pytka (2018), or Arslan et al. (2021) has focussed on how households can reduce the price they pay for a given variety after e.g. a decline in income. The findings outlined above show that additional insurance comes from the effect of low income households' collective shopping effort on the prices for the goods they purchase.

In this sense, non-homotheticities and the partial separation of demand into different varieties reduces the externalities households exerting low shopping effort impose on their high-shopping counterparts. The model economy allows for a quantification of this reduction in externalities as well as the remaining externality due to the non-zero overlap in

households' consumption baskets. To do so, I construct counterfactual price distributions based on alternative values of  $\bar{s}$  across varieties.

For a *full pooling* counterfactual I equalize  $\bar{s}$  across all varieties, computing an average shopping effort by weighting individuals shopping policies with their share in total demand across all varieties such that  $\bar{s} = \int_{(z,a)} \frac{\lambda(z,a) \sum_{j=1}^J c_j(z,a)}{\sum_{j=1}^J C_j} s(z,a)$ . I fix households' consumption baskets and shopping policies at their equilibrium values and compute the change in prices and hence total expenditures if they were drawing from the distributions retailers would post if they faced the counterfactual  $\bar{s}$  for all varieties.<sup>17</sup> The *full pooling* line in Figure 5 plots the percentage change in expenditures across households at different expenditure levels over the range of the model implied expenditure distribution. If all households were pooled in a single price distribution, households at the bottom of the expenditure distribution would spend 2% more on their consumption bundle while households at the top could save 1% while buying the same basket. This is in line with the results from the expenditure decomposition presented above.



**Figure 5:** Shopping Externality

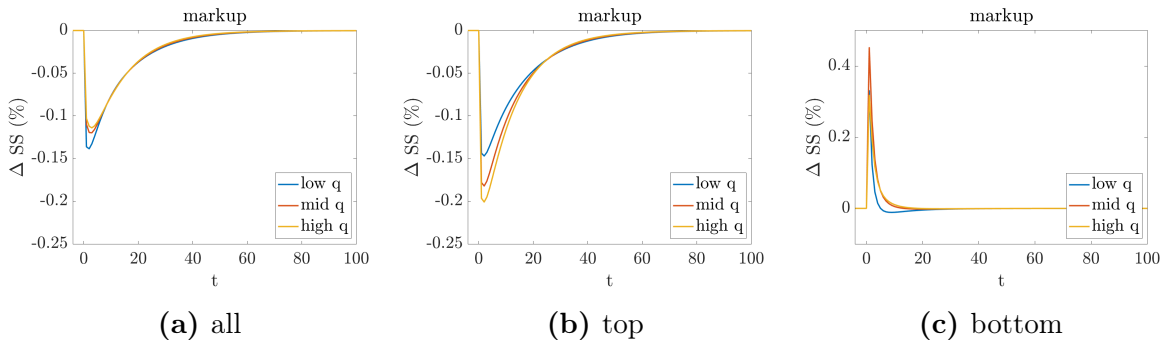
To measure how much externality remains, I allow a household in state  $(z, a)$  to draw for each variety from a price distribution that retailers' would choose to post if all other households were exerting shopping effort  $s(z, a)$ , i.e. the same shopping effort as a type  $(z, a)$  household. In this counterfactual, I obtain *full separation*. Figure 5 shows that the remaining externality is sizeable. Holding shopping effort and consumption baskets fixed, households at the bottom of the expenditure distribution could save an additional 15% if they were drawing prices from a targeted price distribution, i.e. if retailers could perfectly discriminate between household types. The highest spending households on the other hand would pay up to 25% more in a world with perfect discrimination. The large size of the remaining externality can be accounted for by the generally higher spending levels

<sup>17</sup>The exercise can be interpreted as the change in a Laspeyres price index.

at the top of the expenditure distribution. While high-spending households consume relatively less of the goods that are important in low-spending households' baskets, they still account for a sizeable share of expenditures across all varieties due to their higher level of expenditures.<sup>18</sup>

## 4.2 Demand Shocks and the Cyclicalities of Markups

While the previous section has focussed on the implications of heterogeneity in demand composition across goods, this section studies the consequences of an interplay between shopping and demand composition over time. In the model economy, I simulate a decline in aggregate income of 3% holding all parameters fixed at steady state values. I consider three different scenarios: In a first scenario *all* households are affected, i.e. I reduce the labor earnings of all households by 3%. In a second scenario, I reduce aggregate income again by 3% but concentrate the losses only on the *top* 25% households in the labor earnings distribution – each of them affected proportionately to their labor earnings. In a third scenario, I concentrate the same aggregate loss among the *bottom* 25% of the labor earnings distribution.

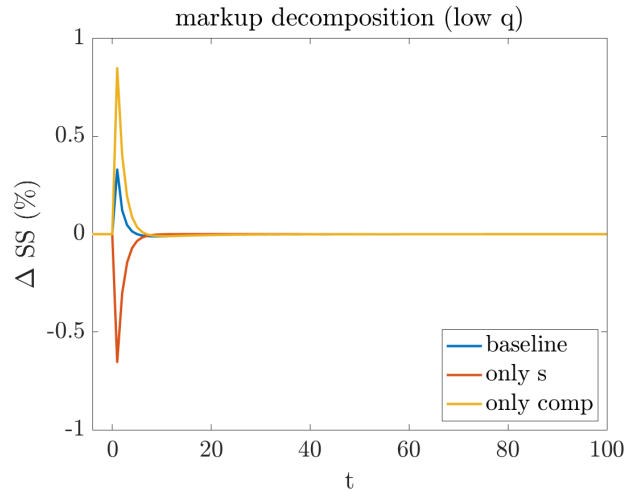


**Figure 6:** Markups and Aggregate Income Losses

Figure 6 plots the response of average posted markups by variety. Focussing first on the case with all households affected, markups decline in response to a loss in aggregate income. If losses are focussed at the top of the distribution, markups become even more pro-cyclical, especially for varieties with higher  $q$  which are disproportionately consumed by high-income households. In response to the same loss in aggregate income concentrated among the bottom of the earnings distribution, markups in the model economy become counter-cyclical. Hence the model implies that the procyclicality of retail markups in response to aggregate income losses is decreasing in the share of the loss accounted for by the bottom of the earnings distribution.

<sup>18</sup>Appendix C.2 shows that this is in line with the data.

To understand the mechanism behind this result, I focus on the response of the low  $q$  variety to a loss concentrated at the bottom of the earnings distribution. In the model, with parameters fixed at steady state levels markups respond only to changes in demand weighted shopping effort  $\bar{s}$ . As equation (2) shows, changes in equilibrium shopping effort can be driven either by changes in households' shopping effort or in demand composition, i.e. the expenditure share stemming from low- vs. high-shopping households. I can consider both drivers separately. Figure 7 plots the equilibrium response as a baseline together with two counterfactuals: First, I fix demand composition at the steady state level and change *only shopping policies* to households' optimal choices along the path. In a second counterfactual I fix households' shopping policies and adjust *only demand composition*.



**Figure 7:** Markup Decomposition

When adjusting only households' shopping effort, markups decline in response to the income loss. This is because households at the bottom have become poorer and try to compensate by increasing their search for cheaper prices and reducing the cost of consumption. What is responsible for the overall increase in markups is the change in demand composition over time. As the bottom of the distribution loses income and has to reduce expenditures, high-income households account for a larger share of demand. These households were exerting lower shopping effort to begin with and therefore any shift of demand composition towards them reduces  $\bar{s}$  and with it the price elasticity of demand. Retailers face lower competition and find it optimal to increase markups. The second effect dominates the increase in shopping effort at the bottom of the distribution and overall markups increase. For an income loss concentrated at the top of the earnings distribution both effects work in the same direction: High-income households increase their effort and demand composition shifts towards low-income (high-shopping) households.

The findings are in sharp contrast to the results of the previous section: While heterogeneity in demand composition across goods provides additional insurance by reducing the markups charged on goods bought by low-income households, shifts in demand composition over time amplify inequalities: High-earning households are partially compensated through a decline in posted markups when they are hit by an aggregate income loss. Low-earning households might see markups rise if they lose income as retailers adjust their price setting to their declining share in aggregate demand.

While an empirical investigation of the cyclicity of retail markups is beyond the scope of this paper, empirical findings in the literature provide support to the results presented above. Focussing on retail markups and using similar datasets to Nielsen but enriched with data on wholesale cost, Anderson et al. (2020) find acyclical markups in response to local unemployment shocks while Stroebel and Vavra (2019) find strongly procyclical retail markups in response to changes in local house prices. This evidence can be reconciled with the findings outlined above under the reasonable assumption that unemployment risk is concentrated at the bottom of the income distribution while house prices affect only homeowners who are concentrated at the middle and top of the distribution.

## 5 Conclusion

This paper revisits the role of households' shopping behavior for inequality in consumption expenditures, studying equilibrium interactions between heterogeneity in shopping effort, demand composition, and optimal price setting. Heterogeneous consumption baskets along the income distribution and higher shopping effort of the poor imply that retailers face different price elasticities depending on their customer base and charge higher markups for goods disproportionately consumed by richer households. I highlight the basic intuitions in a standard framework of price posting under search frictions and test theoretical predictions on the shape of offered price distributions in the Nielsen Consumer Panel, finding strong empirical support for the proposed mechanism. To study implications for inequality, I embed the simple framework in a quantitative model featuring rich household heterogeneity, non-homothetic preferences, and endogenous shopping effort, as well as endogenously determined price distributions for multiple varieties. The calibrated model replicates salient features of expenditure inequality and price dispersion. It shows that equilibrium interactions between shopping effort, demand composition across varieties, and price posting double the contribution of shopping effort to expenditure inequality relative to previous findings. The model also implies that the response of markups to aggregate income shocks crucially depends on the incidence of the shock

along the income distribution, due to shifts in demand composition over time. In ongoing work, I study implications for the effectiveness of redistributive policies and the design of consumption taxes.



## References

- Aguiar, Mark and Mark Bilts (2015). “Has Consumption Inequality mirrored Income Inequality?” *American Economic Review* 105.9, pp. 2725–56.
- Aguiar, Mark and Erik Hurst (2005). “Consumption versus Expenditure.” *Journal of Political Economy* 113.5, pp. 919–948.
- (2007). “Life-cycle Prices and Production.” *American Economic Review* 97.5, pp. 1533–1559.
- Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis (2013). “Time Use during the Great Recession.” *American Economic Review* 103.5, pp. 1664–96.
- Aiyagari, S. Rao (1994). “Uninsured Idiosyncratic Risk and Aggregate Saving.” *The Quarterly Journal of Economics* 109.3, pp. 659–684.
- Albrecht, James, Guido Menzio, and Susan Vroman (2021). “Vertical Differentiation in Frictional Product Markets.”
- Alessandria, George (2009). “Consumer Search, Price Dispersion, and International Relative Price Fluctuations.” *International Economic Review* 50.3, pp. 803–829.
- Anderson, Eric, Sérgio Rebelo, Arlene Wong, et al. (2020). “Markups Across Space and Time.”
- Argente, David and Munseob Lee (2021). “Cost of Living Inequality during the Great Recession.” *Journal of the European Economic Association* 19.2, pp. 913–952.
- Arslan, Yavuz, Bulent Guler, and Temel Taskin (2021). “Price Search, Consumption Inequality and Expenditure Inequality over the Life-Cycle.” *International Economic Review* 62.1, pp. 295–320.
- Attanasio, Orazio P. and Luigi Pistaferri (2016). “Consumption Inequality.” *Journal of Economic Perspectives* 30.2, pp. 3–28.
- Bewley, Truman (1977). “The Permanent Income Hypothesis: A Theoretical Formulation.” *Journal of Economic Theory* 16.2, pp. 252–292.
- Bilts, Mark and Peter J Klenow (2001). “Quantifying Quality Growth.” *American Economic Review* 91.4, pp. 1006–1030.
- Bisgaard Larsen, Rasmus and Christoffer J. Weissert (2020). “Quality and Consumption Basket Heterogeneity - Transitory Shocks and Implications for Consumption-Savings Behaviour.”

- Broda, Christian and Jonathan A Parker (2014). “The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption.” *Journal of Monetary Economics* 68, S20–S36.
- Broda, Christian, Ephraim Leibtag, and David E Weinstein (2009). “The Role of Prices in Measuring the Poor’s Living Standards.” *Journal of Economic Perspectives* 23.2, pp. 77–97.
- Burdett, Kenneth and Kenneth L Judd (1983). “Equilibrium Price Dispersion.” *Econometrica*, pp. 955–969.
- Burdett, Kenneth and Guido Menzio (2018). “The (Q, S, s) Pricing Rule.” *The Review of Economic Studies* 85.2, pp. 892–928.
- Burdett, Kenneth and Dale T Mortensen (1998). “Wage Differentials, Employer Size, and Unemployment.” *International Economic Review*, pp. 257–273.
- Butters, Gerard R. (1977). “Equilibrium Distributions of Sales and Advertising Prices.” *The Review of Economic Studies* 44.3, pp. 465–491.
- Carroll, Christopher D. (2006). “The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems.” *Economics Letters* 91.3, pp. 312–320.
- Coibion, Olivier, Yuriy Gorodnichenko, and Dmitri Koustas (2021). “Consumption Inequality and the Frequency of Purchases.” *American Economic Journal: Macroeconomics* 13.4, pp. 449–82.
- De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz-Pardo (2020). “Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare.” *Journal of the European Economic Association* 18.2, pp. 890–926.
- DellaVigna, Stefano and Matthew Gentzkow (2019). “Uniform Pricing in US Retail Chains.” *The Quarterly Journal of Economics* 134.4, pp. 2011–2084.
- Droste, Michael, Luigi Pisano, Luca Riva, and Andrea Stella (2019). “Price Heterogeneity and Consumption Inequality.”
- Faber, Benjamin and Thibault Fally (2022). “Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data.” *The Review of Economic Studies* 89.3, pp. 1420–1459.
- Farmer, Leland E and Alexis Akira Toda (2017). “Discretizing Nonlinear, Non-Gaussian Markov Processes with Exact Conditional Moments.” *Quantitative Economics* 8.2, pp. 651–683.
- Ferraro, Domenico and Vytautas Valaitis (2022). “Wealth and Hours.”

- Ferrier, Axelle, Philipp Grübener, Gaston Navarro, and Oliko Vardishvili (2022). “On the Optimal Design of Transfers and Income-Tax Progressivity.”
- Griffith, Rachel, Ephraim Leibtag, Andrew Leicester, and Aviv Nevo (2009). “Consumer Shopping Behavior: How much do Consumers Save?” *Journal of Economic Perspectives* 23.2, pp. 99–120.
- Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2021). “What do Data on Millions of US Workers reveal about Lifecycle Earnings Dynamics?” *Econometrica* 89.5, pp. 2303–2339.
- Hall, Robert E. (2018). “New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy.”
- Handbury, Jessie (2021). “Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living across US Cities.” *Econometrica* 89.6, pp. 2679–2715.
- Jaimovich, Nir, Sergio Rebelo, and Arlene Wong (2019). “Trading down and the Business Cycle.” *Journal of Monetary Economics* 102, pp. 96–121.
- Jaravel, Xavier (2019). “The Unequal Gains from Product Innovations: Evidence from the US Retail Sector.” *The Quarterly Journal of Economics* 134.2, pp. 715–783.
- Jaravel, Xavier and Alan Olivi (2021). “Prices, Non-homotheticities, and Optimal Taxation.” *Available at SSRN 3383693*.
- Kaplan, Greg and Guido Menzio (2015). “The Morphology of Price Dispersion.” *International Economic Review* 56.4, pp. 1165–1206.
- (2016). “Shopping Externalities and Self-Fulfilling Unemployment Fluctuations.” *Journal of Political Economy* 124.3, pp. 771–825.
- Kaplan, Greg, Guido Menzio, Leena Rudanko, and Nicholas Trachter (2019). “Relative Price Dispersion: Evidence and Theory.” *American Economic Journal: Microeconomics* 11.3, pp. 68–124.
- Mendicino, Caterina, Lukas Nord, and Marcel Peruffo (2022). “Distributive Effects of Banking Sector Losses.”
- Menzio, Guido (2021). “Optimal Product Design: Implications for Competition and Growth under Declining Search Frictions.”
- Michelacci, Claudio, Luigi Paciello, and Andrea Pozzi (2020). “The Extensive Margin of Aggregate Consumption Demand.”
- Mortensen, Dale (2003). *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT press.

- Nakamura, Emi (2008). “Pass-Through in Retail and Wholesale.” *American Economic Review* 98.2, pp. 430–37.
- Nevo, Aviv and Arlene Wong (2019). “The Elasticity of Substitution between Time and Market Goods: Evidence from the Great Recession.” *International Economic Review* 60.1, pp. 25–51.
- Pytka, Krzysztof (2018). “Shopping Effort in Self-Insurance Economies.”
- Sangani, Kunal (2022). “Markups Across the Income Distribution: Measurement and Implications.”
- Stroebel, Johannes and Joseph Vavra (2019). “House Prices, Local Demand, and Retail Prices.” *Journal of Political Economy* 127.3, pp. 1391–1436.
- Varian, Hal R (1980). “A Model of Sales.” *The American Economic Review* 70.4, pp. 651–659.

## A Extensions to the Retailer Problem

### A.1 Entry and Fixed Cost of Operating

Consider a version of the retailer problem outlined in Section 2.1 where retailers selling variety  $j$  are subject to a per-period fixed cost of operating  $K_j$  and the total mass of retailers  $M_j$  is determined endogenously by free entry. With all other notation as before, the total profits of a retailer posting price  $p$  in the adjusted model are given by

$$\tilde{\pi}_j(p) = \frac{C_j}{M_j} [(1 - \bar{s}_j) + \bar{s}_j 2(1 - F_j(p))] (p - \kappa_j) - K_j = \frac{\pi_j(p)}{M_j} - K_j$$

where  $\pi_j(p)$  is retailers' profits of posting  $p$  in the version of the model without fixed cost of operating and a fixed mass one of retailers. To sustain an equilibrium distribution of posted prices retailers have to be again indifferent between all prices offered. Take two offered prices  $p_1$  and  $p_2$ , indifference requires

$$\begin{aligned} \tilde{\pi}_j(p_1) &= \tilde{\pi}_j(p_2) \\ \Rightarrow \frac{\pi_j(p_1)}{M_j} - K_j &= \frac{\pi_j(p_2)}{M_j} - K_j \\ \Rightarrow \pi_j(p_1) &= \pi_j(p_2) \end{aligned}$$

The indifference condition between prices is independent of  $M_j$  and  $K_j$ , i.e. independent of entry and fixed cost of operating, and identical to the condition in the original model. This implies the distribution of offered prices  $F_j(p)$  is identical to the model without entry and fixed cost. To solve the model with fixed cost and entry, one can therefore first recover the offered price distribution as well as the constant profits at any price on the support of  $F_j(p)$ , denoted  $\bar{\pi}_j$ , in the original model and solve for  $M_j$  given this solution. Free entry requires zero total profits of operating, i.e.  $\tilde{\pi}_j(p) = 0$ . The equilibrium number of retailers therefore given by  $M_j = \frac{\bar{\pi}_j}{K_j}$ .

### A.2 Heterogeneous Marginal Cost

Take the setup from Section 2.1 but consider a continuous distribution of retailers over marginal cost, with CDF  $\Gamma_j(\kappa)$  and support  $[\underline{\kappa}_j, \bar{\kappa}_j]$  and assume  $\bar{\kappa}_j = \bar{p}_j$ . I.e. consider a distribution of active firms for which the support has to end at the maximum willingness to pay, as no retailer with marginal cost above  $\bar{p}_j$  could ever make a sale with positive profits. Profits of a retailer with marginal cost  $\kappa$  of posting price  $p$  for variety  $j$  are given

by

$$\pi_j(p, \kappa) = (p - \kappa) ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p))) C_j$$

Define  $p(\kappa)$  as the set of prices maximizing  $\pi_j(p, \kappa)$  for given  $F_j(p)$ , i.e. the indifference set of posted prices for a retailer with marginal cost  $\kappa$ .

### A.2.1 Solving for the Distribution of Offered Prices

To solve for the equilibrium distribution of offered prices I follow closely the steps of Burdett and Mortensen (1998) or Mortensen (2003) for a similar model of wage posting.

#### 1. Properties of the Distribution

By similar argument as in Burdett and Judd (1983), the offered distribution  $F_j(p)$  has no mass points, has a connected support, and the upper bound of the support of  $F_j(p)$  is  $\bar{p}_j$ . Intuitively all three can be shown by providing a profitable deviation in price posting if an offered price distribution is violating one of the three conditions.

#### 2. Prices Posted Are Weakly Increasing in Marginal Cost

We can show that for any  $\kappa'' > \kappa'$ ,  $p' \in p(\kappa')$ , and  $p'' \in p(\kappa'')$  it has to hold that  $\pi_j(p', \kappa') > \pi_j(p'', \kappa'')$  and  $p'' \geq p'$ , i.e. profits are strictly decreasing and prices are weakly increasing in marginal costs. To do so, note the following

$$\pi'_j(p', \kappa') = (p' - \kappa') ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p'))) C_j \quad (\star 1)$$

$$\geq (p'' - \kappa') ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p''))) C_j \quad (\star 2)$$

$$> (p'' - \kappa'') ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p''))) C_j = \pi_j(p'', \kappa'') \quad (\star 3)$$

$$\geq (p' - \kappa'') ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p'))) C_j \quad (\star 4)$$

Where the steps from  $(\star 1)$  to  $(\star 2)$  and  $(\star 3)$  to  $(\star 4)$  follow from the optimality of  $p' \in p(\kappa')$  and  $p'' \in p(\kappa'')$  respectively and the step from  $(\star 2)$  to  $(\star 3)$  from  $\kappa'' > \kappa'$ . From above it is immediately clear that  $\pi_j(p', \kappa') > \pi_j(p'', \kappa'')$ , i.e. profits are strictly decreasing in  $\kappa$ . To see that  $p'' \geq p'$  note that  $(\star 1) - (\star 4) \geq (\star 2) - (\star 3) > 0$  and hence

$$(\kappa'' - \kappa') ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p'))) \geq (\kappa'' - \kappa') ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p'')))$$

which yields  $F_j(p'') \geq F_j(p')$  and therefore, as any cumulative distribution cannot be decreasing,  $p'' \geq p'$ . So any price optimal at  $\kappa'$  cannot be higher than any price optimal at  $\kappa''$ . Hence  $p(\kappa')$  and  $p(\kappa'')$  can intersect in at most one boundary point.

With a continuous distribution of marginal cost, the latter also implies that  $p(\kappa)$  has to be single valued.

### 3. The Price Distribution is a Shifted Distribution of Marginal Cost

By the single value property of  $p(\kappa)$

$$F_j(p) = F_j(p(\kappa)) = \Gamma_j(\kappa)$$

and hence

$$F'_j(p(\kappa)) = f_j(p(\kappa)) = \frac{\Gamma'_j(\kappa)}{p'(\kappa)}$$

### 4. The Price Function $p(\kappa)$ Solves Retailers Profit Maximization

Analogue to before, the profits of a retailer with marginal cost  $\kappa$  posting price  $p$  are given by

$$\pi_j(p, \kappa) = (p - \kappa) ((1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p)))$$

and the profit maximizing price satisfies

$$\frac{\partial \pi}{\partial p} = (1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p)) - (p - \kappa)2\bar{s}_j F'_j(p) = 0$$

which yields

$$1 = \frac{(p - \kappa)2\bar{s}_j F'_j(p)}{(1 - \bar{s}_j) + 2\bar{s}_j(1 - F_j(p))}$$

and by the result of 3.)

$$p'(\kappa) = \frac{(p(\kappa) - \kappa)2\bar{s}_j \Gamma'_j(\kappa)}{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))}$$

This differential equation together with the boundary condition  $p(\bar{\kappa}) = \bar{\kappa} = \bar{p}$  pins down the unique solution to  $p(\kappa)$  and hence to  $F'_j(p)$ . The boundary condition holds because the upper bound of any price distribution has to be at  $\bar{p}$  (else there are profitable deviations) and a firm with marginal cost  $\kappa = \bar{\kappa} = \bar{p}$  will only be willing to post this price.

### 5. Obtaining a Solution

Define

$$T(k) = \log((1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(k)))$$

such that

$$T'(k) = \frac{-2\bar{s}_j\Gamma_j(\kappa)}{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))}$$

We can the rewrite the first difference equation pinning down the pricing function as

$$p'(\kappa) = -(p(\kappa) - \kappa)T'(\kappa) \Rightarrow p'(\kappa) + T'(\kappa)p(\kappa) = \kappa T'(\kappa)$$

Therefore any solution has to satisfy (multiply both sides by  $e^{T(\kappa)}$  before integrating)

$$p(\kappa)e^{T(\kappa)} = \int_{\underline{\kappa}}^{\kappa} xT'(x)e^{T(x)}dx + A = \kappa e^{T(\kappa)} - \underline{\kappa}e^{T(\underline{\kappa})} - \int_{\underline{\kappa}}^{\kappa} e^{T(x)}dx + A$$

where the second equality follows from integration by parts. Hence

$$p(\kappa) = \kappa + e^{-T(\kappa)} \left[ A - \underline{\kappa}e^{T(\underline{\kappa})} - \int_{\underline{\kappa}}^{\kappa} e^{T(x)}dx \right]$$

Using the boundary condition  $p(\bar{\kappa}) = \bar{\kappa}$  it follows that

$$A = \underline{\kappa}e^{T(\underline{\kappa})} + \int_{\underline{\kappa}}^{\bar{\kappa}} e^{T(x)}dx$$

The solution to the pricing function and the distribution of offered prices is hence given as

$$p(\kappa) = \kappa + e^{-T(\kappa)} \int_{\underline{\kappa}}^{\bar{\kappa}} e^{T(x)}dx = \kappa + \int_{\underline{\kappa}}^{\bar{\kappa}} \frac{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(x))}{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))} dx$$

with derivative

$$p'(\kappa) = 1 - 1 - (-2\bar{s}_j\Gamma_j'(\kappa)) \int_{\underline{\kappa}}^{\bar{\kappa}} \frac{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(x))}{[(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))]^2} dx > 0$$

and therefore

$$F'(p) = \frac{\Gamma_j'(\kappa)}{p'(\kappa)} = \frac{1}{2\bar{s}_j \int_{\underline{\kappa}}^{\bar{\kappa}} \frac{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(x))}{[(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))]^2} dx} = \frac{[(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))]^2}{2\bar{s}_j(\bar{\kappa} - \kappa)(1 + \bar{s}_j) - 4\bar{s}_j^2 \int_{\underline{\kappa}}^{\bar{\kappa}} \Gamma_j(x) dx}$$

We cannot conclude anything on how the profit margins (markups) per sale are changing with marginal cost  $\kappa$ . To see this note that



$$p(\kappa) - \kappa = \int_{\kappa}^{\bar{\kappa}} \frac{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(x))}{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))} dx$$

and hence

$$\frac{\partial(p(\kappa) - \kappa)}{\partial \kappa} = p'(\kappa) - 1 = -1 + 2\bar{s}_j \Gamma_j'(\kappa) \int_{\kappa}^{\bar{\kappa}} \frac{(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(x))}{[(1 - \bar{s}_j) + 2\bar{s}_j(1 - \Gamma_j(\kappa))]^2} dx$$

So whether markups are increasing or decreasing in marginal costs depends on the shape of the distribution  $\Gamma_j(\kappa)$ . Intuitively, retailers' optimization trades off higher margins (markups) against a decrease in demand, where the latter depends on the distribution of prices which in turn depends on the distribution of marginal costs.

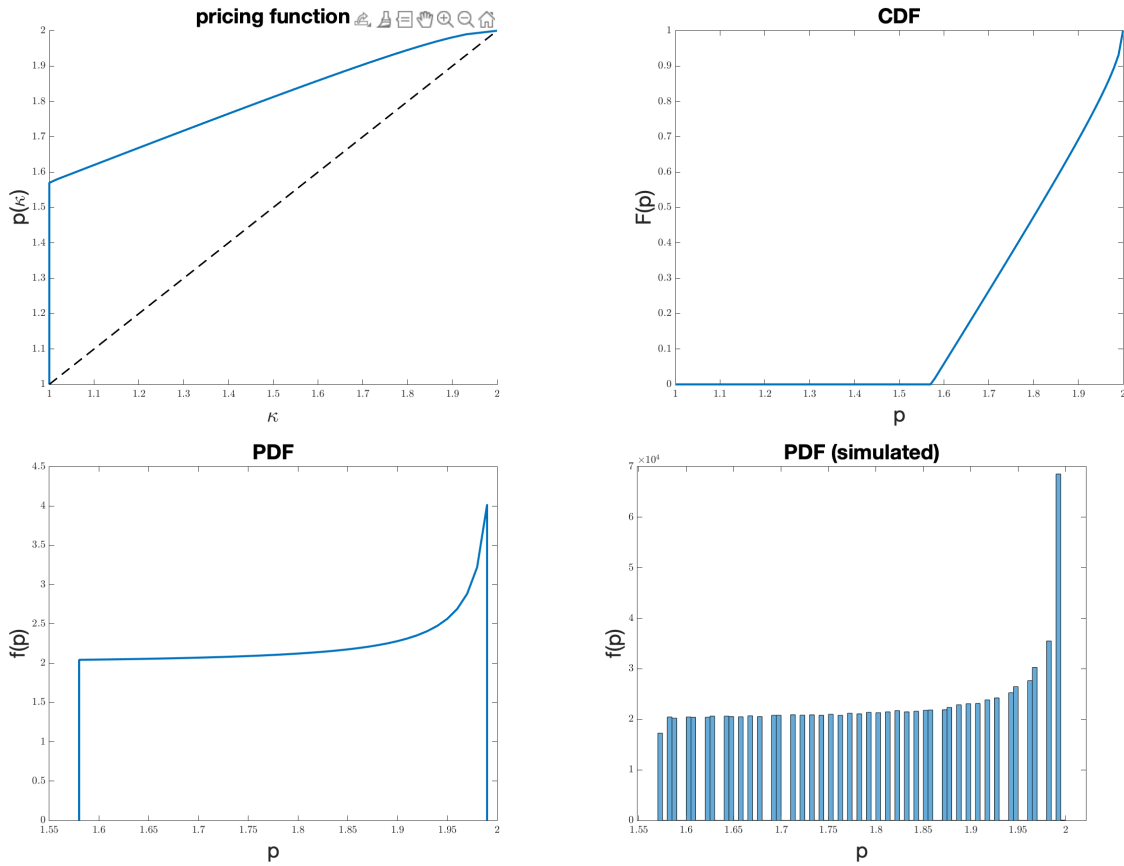
### A.2.2 Quantitative Results under Uniform Marginal Costs

While an analytical characterization of how the moments of the price distribution respond to changes in equilibrium shopping effort  $\bar{s}_j$  under heterogeneous marginal cost is beyond the scope of this paper, I show robustness of the analytical results for the baseline model by reporting numerical simulations. I assume a uniform distribution of marginal cost over  $[\underline{\kappa}_j, \bar{\kappa}_j]$  and consider parameterizations with  $\bar{p}_j \in \{1, 2, 3, 4, 5\}$ ,  $\kappa_{min} \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$  such that  $\underline{\kappa}_j = \kappa_{min}\bar{p}_j$  and  $\bar{\kappa}_j = \bar{p}_j$ . I take  $\bar{p}_j = 3$  and  $\kappa_{min} = 0$  as the baseline and vary one parameter at a time, simulating 1,000,000 price draws for each combination of parameters and computing the mean and skewness of the offered price distribution.

To highlight the properties of a solution to the model with heterogeneous  $\kappa$ , Figure 8 plots the pricing function  $p(k)$  and CDF  $F_j(p)$  as well as the analytical and simulated PDF of a single calibrated version with  $\bar{\kappa} = \bar{p} = 2$ ,  $\underline{\kappa} = 1$ ,  $\bar{s} = 0.75$ .

Figure 9 recovers the result of skewness being a strictly increasing function of average search effort  $\bar{s}$ . Other parameters do not have considerable influence on the skewness of the price distribution. For the mean of posted prices the main mechanism pertains: For any combination of parameters the average posted price is decreasing in shopping effort. This is because the pricing function gets more and more concentrated at the maximum willingness to pay when  $\bar{s}$  goes to zero.

Results for other types of distributions (exponential, logistic) as well as a version with a discrete set of marginal-cost types yield similar conclusions: While under some calibra-



**Figure 8:** Uniform Distribution - Example

Calibration with  $\bar{p} = 2$ ,  $\underline{\kappa} = 1$ ,  $\bar{s} = 0.75$ .

tions small regions of skewness decreasing in shopping effort are possible, these usually exist only for  $\bar{s}_j \approx 1$  and are associated with counterfactually low levels of price dispersion. Further results are available upon request. Exploiting the skewness of price distributions for an empirical test of the mechanism is therefore a reasonable approximation even in a world with potentially heterogeneous marginal cost.

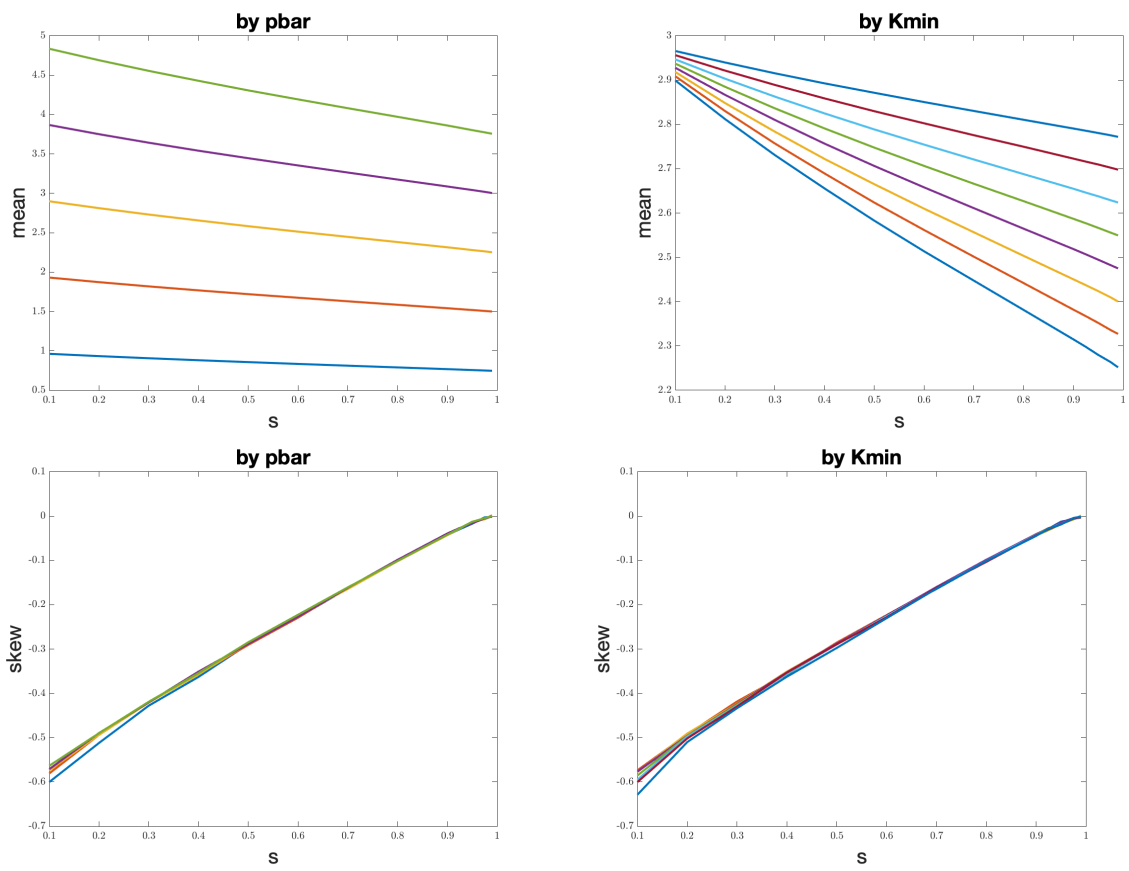


Figure 9: Uniform Distribution - Simulations

## B The Nielsen Dataset

All empirical results are based on the Nielsen Consumer Panel, provided via the Kilts Center for Marketing at Chicago Booth. The dataset is a nationally representative, annual panel of around 60,000 US households who report on their consumption expenditures at daily as well as demographic information at annual frequency. Demographic variables include e.g. information on household composition, age, education, occupation, employment status, income, and location of residence. Nielsen applies several quality checks such as minimum reporting restrictions before making the data available and keeps households in the sample for multiple waves, with a median of three waves per household. Households in the sample are provided with a device to record the prices and quantities of all purchases made in stores by scanning the barcodes of the items they bought (or record prices manually if the store is not participating in Nielsen’s sample). The focus of the dataset is on grocery and drug stores, supermarkets and superstores, covering approximately 15% of total consumer spending, 40% of spending excluding durable goods, and 70% of spending excluding durables and services, i.e. all results can be seen as representative for non-durable consumption.<sup>19</sup>

Expenditure is reported at the barcode level. Nielsen organizes all barcodes into 10 departments (e.g. dry groceries or fresh foods), which are then divided into 125 product groups (e.g. snacks vs. pasta within dry groceries), and further split into about 1,100 product modules (e.g. potato chips vs. tortilla chips within snacks). Within product modules each variety is uniquely identified by its Universal Product Code (UPC), examples of a UPC are e.g. a box of Pringles Sour Cream and Onion or a bag of Lay’s BBQ within the module potato chips. For each purchase of a barcode at a store at a given day, Nielsen records the quantity bought, the total price of the transaction, the value of all coupons used as well as the unique store identifier of the location where the purchase was made. Households’ purchases can further be grouped into shopping trips, where a trip consists of all purchases of any barcode made by a household in a given store on a given day.

Data is provided in annual waves and I use the waves of 2007-2019. Data is also available for the period 2004-2006, but I focus on the later period due to a sample break between 2006 and 2007. All empirical results remain qualitatively unchanged if earlier waves are included. Across all households the dataset contains a total of about 7.5 million shopping trips and around 50 million purchases from a universe of 500,000 UPCs per wave.

---

<sup>19</sup>For further details on the dataset see e.g. Argente and Lee (2021), Kaplan and Menzio (2015), Droste et al. (2019), Broda and Parker (2014) or Michelacci et al. (2020).

No data on wealth is available in the Nielsen panel and income data is only available as categorical variable and reported for the previous calendar year, while expenditure on the consumption categories covered in Nielsen is well measured. This is why for all baseline results on heterogeneity across households, I sort by their position in the expenditure distribution. Whenever I refer to expenditure, I adjust households' total annual expenditure measured in the Nielsen dataset by the square root of household size and (where applicable) sort them into quintiles/deciles based on their position in the expenditure distribution in the year of observation. Wherever dollar values are reported, these are inflation adjusted to 2019 USD.

## C Additional Empirical Results

### C.1 Evidence on Demand Composition and Price Distributions

Table 5: Robustness

	baseline	Kelly	unweighted regression	unweighted skewness	only HH weights skewness
	(1)	(2)	(3)	(4)	(5)
quintile 2	-1.638*** (0.188)	-0.135*** (0.036)	-1.802*** (0.099)	-1.450*** (0.153)	-1.627*** (0.181)
quintile 3	-2.309*** (0.189)	-0.226*** (0.035)	-2.583*** (0.095)	-2.100*** (0.152)	-2.354*** (0.180)
quintile 4	-3.067*** (0.172)	-0.282*** (0.034)	-3.374*** (0.089)	-2.793*** (0.148)	-3.062*** (0.169)
quintile 5	-3.412*** (0.153)	-0.382*** (0.030)	-4.066*** (0.082)	-3.151*** (0.133)	-3.425*** (0.151)
FE module	795	777	795	795	795
FE Y:Q:SMC	3584	3584	3584	3584	3584
Observations	3,026,551	2,832,442	3,026,551	3,026,551	3,026,551
R <sup>2</sup>	0.035	0.027	0.034	0.054	0.036
Adjusted R <sup>2</sup>	0.034	0.026	0.033	0.053	0.035

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

**Table 6:** Minimum Number of Observations

	Nmin = 25	Nmin = 50	Nmin = 100
	(1)	(2)	(3)
quintile 2	-1.638*** (0.188)	-2.289*** (0.398)	-3.162*** (0.873)
quintile 3	-2.309*** (0.189)	-2.929*** (0.412)	-4.268*** (0.840)
quintile 4	-3.067*** (0.172)	-3.797*** (0.374)	-5.186*** (0.811)
quintile 5	-3.412*** (0.153)	-4.654*** (0.325)	-6.436*** (0.709)
FE module	795	539	268
FE Y:Q:SMC	3584	3584	3534
Observations	3,026,551	803,604	202,067
R <sup>2</sup>	0.035	0.059	0.111
Adjusted R <sup>2</sup>	0.034	0.054	0.094

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

## C.2 Consumption Baskets and Separation in the Goods Market

To quantify non-homotheticities in the data, one needs to find a measure for the similarity of consumption baskets. I define the consumption basket of any group  $g$  of households  $i$  via the share of their annual total expenditure allocated to each good  $\omega_j^g$ . This *basket share* of good  $j$  for group  $g$  in a given year are given as

$$\omega_j^g = \frac{\sum_{i \in g} e_j^i}{\sum_{j \in J} \sum_{i \in g} e_j^i}$$

The vector of basket shares for any given group can be seen as a distribution over a discrete set of alternatives – the universe of available products. We can hence measure the similarity of two such vectors, i.e. the consumption baskets of two groups of households  $g$  and  $h$ , by computing the histogram overlap  $\Omega^{g,h}$ , given as

$$\Omega^{g,h} = \sum_{j \in J} \min \{ \omega_j^g, \omega_j^h \}$$

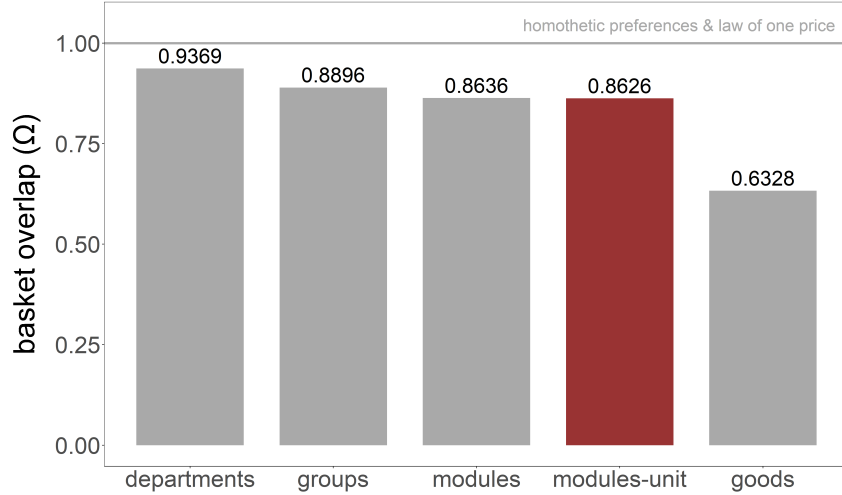
Note that under homothetic preferences and the law of one price we would have that  $\omega_j^g = \omega_j^h \forall j, g, h$  and hence  $\Omega^{g,h} = 1$ , so any deviation of the overlap from one can be interpreted as a deviation from these assumptions. We conduct this analysis by groups of households to smooth out variation in taste within groups and compute statistics at the annual frequency to average out seasonal fluctuations.

Figure 10 reports the histogram overlap between the first and fifth quintile of the distribution of annual expenditures, defining a “product” at different levels of aggregation. If products are broadly defined, e.g. at the Nielsen department level, the overlap in consumption baskets is as high as 94% and even when considering product modules it is still as high as 86%. Only at the lowest level of aggregation where products are unique UPCs (the goods level) the overlap decreases drastically to 63%. I.e. consumption baskets of high and low expenditure households exhibit a substantial mismatch driven by variation in purchases of closely substitutable goods within Nielsen-defined product modules. For the empirical decompositions in Section 4.1 it is also important to note that conditioning on units of measurement within product modules does not alter the overlap substantially compared to considering the entire module, i.e. there are no notable non-homotheticities by unit of measurement.<sup>20</sup>

---

<sup>20</sup>The overlap between any other two quintiles of the expenditure distribution exhibits similar patterns. Overlap at any level of aggregation decreases monotonically in the distance (difference in total expenditures) between two groups.





**Figure 10:** Consumption Basket Overlap - Top vs. Bottom Expenditure Quintile

Complementary evidence to the missing overlap in consumption baskets is a measure of how important the demand of other households is for the goods that any group of households buys. First, to determine how important demand from any group of households  $g$  is for a given good  $j$ , we define the *good expenditure share* ( $GS$ ) of group  $h$  for good  $j$  as

$$GS_j^h = \frac{\sum_{i \in h} e_j^i}{\sum_{g \in G} \sum_{i \in g} e_j^i}$$

We can then weight the good shares of group  $h$  with the consumption basket of group  $g$  to compute the *cross market share* ( $CMS$ ) of group  $h$  for the basket of group  $g$ , defined as

$$CMS^{gh} = \sum_{j \in J} \omega_{jt}^g GS_j^h$$

This statistic can be interpreted as the average expenditure share of group  $h$  in the basket of group  $g$  and hence measures how important group  $h$  is for the demand of goods that group  $g$  buys.

Figure 11 plots the cross market shares by quintile of the expenditure distribution at the good (UPC) level. It shows that each group of households is substantially overrepresented in their own consumption baskets. E.g. the lowest expenditure quintile is twice as important for their own consumption basket as for the basket of the highest expenditure quintile.

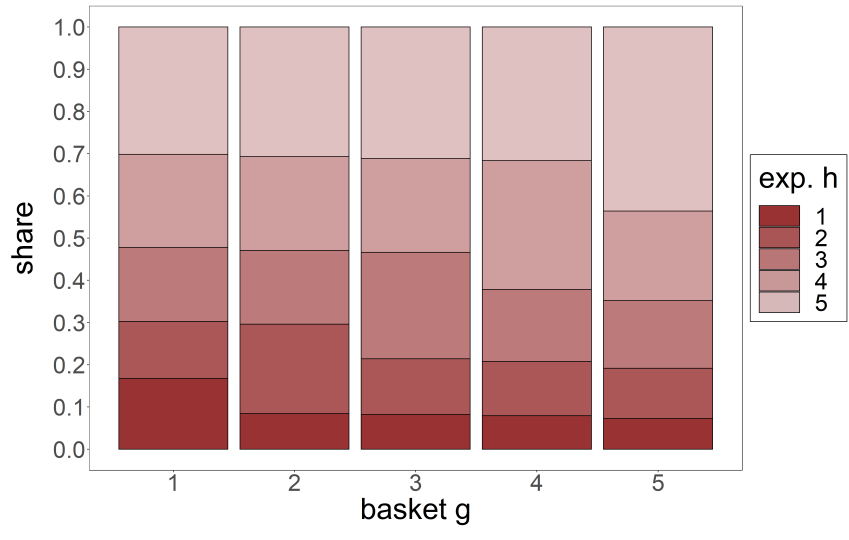


Figure 11: Cross Market Shares

### C.3 Expenditure Decomposition

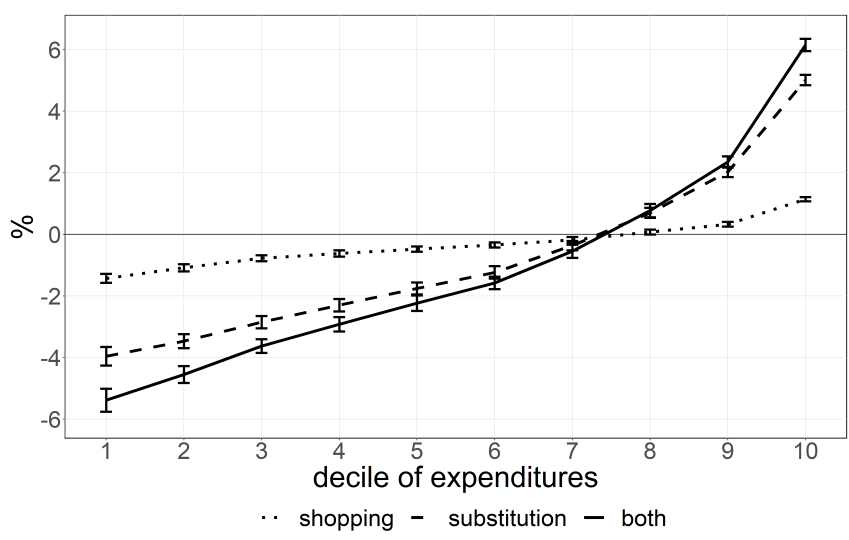


Figure 12: Expenditure Decomposition – Deciles

**Table 7:** Expenditure Decomposition - Regressions

	shopping	substitution	trips per purchase
	(1)	(2)	(3)
log(expenditure)	0.952*** (0.069)	3.399*** (0.147)	-0.042*** (0.001)
prime age	0.074 (0.053)	0.302*** (0.111)	-0.001** (0.001)
HH size	0.506*** (0.073)	0.520*** (0.146)	-0.032*** (0.001)
income 30k-60k	0.086* (0.050)	0.613*** (0.101)	-0.001** (0.001)
income 60k-100k	0.235*** (0.063)	1.178*** (0.128)	-0.002*** (0.001)
income >100k	0.431*** (0.080)	1.631*** (0.161)	-0.002*** (0.001)
mean			0.15
year:state FE	X	X	X
HH FE	X	X	X
Observations	801,398	801,398	801,398
R <sup>2</sup>	0.818	0.803	0.888
Adjusted R <sup>2</sup>	0.765	0.745	0.855

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01.