Shopping, Demand Composition, and Equilibrium Prices

Lukas Nord*
Minneapolis Fed

December 20, 2023

Abstract. This paper develops an equilibrium theory of expenditure inequality and price dispersion to study how retail prices respond to households' shopping behavior. The theory features incomplete markets, non-homothetic preferences, and equilibrium price dispersion for multiple varieties. Endogenous heterogeneity in the effort to search for prices shifts the price elasticity faced by retailers with the composition of demand. Retailers optimally charge higher margins for goods consumed by low-effort households. Predictions on the shape of price distributions are consistent with evidence from US scanner-data. Heterogeneity in search effort yields differences in prices paid for identical goods and reduces inequality in consumption relative to expenditure. The equilibrium response of posted prices across products doubles this direct effect of search on inequality. The model reconciles conflicting evidence on the cyclicality of retail markups and shows that the response of posted prices to a redistributive tax partially compensates net-contributors for their losses.

Keywords: Household Heterogeneity, Expenditures, Price Search, Markups.
JEL: D10, D31, E21, E30, L11

---

*lksnord.econ@gmail.com. I am indebted to my advisors, Árpád Ábrahám and Russell Cooper, for their continuous support. This work benefited greatly from comments by Annika Bacher, Mark Bils, Jesús Bueren, Giacomo Calzolari, Edouard Challe, Giancarlo Corsetti, Tom Crossley, Alessandro Ferrari, Axelle Ferriere, Philipp Grübener, Kyle Herkenhoff, Wolfram Horn, Greg Kaplan, Ramon Marimon, Guido Menzio, Claudio Michelacci, Virgiliu Midrigan, Kurt Mitman, Alex Monge, Simon Mongey, Giuseppe Moscarini, Alessandra Peter, Krzysztof Pytka, Víctor Ríos-Rull, Todd Schoellman, Kjetil Storesletten, and Joshua Weiss, as well as conference, workshop, and seminar participants at numerous institutions. All remaining errors are mine. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. 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 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

Understanding inequality in households' consumption is essential to infer the welfare consequences of income and wealth inequality. A growing literature emphasizes heterogeneity in prices across households and distinguishes consumption from expenditure inequality (e.g. Aguiar and Hurst 2005, 2007). This distinction is important because posted prices for identical products exhibit significant dispersion and poor households search for bargains to pay less for the same good. Previous work abstracts from any equilibrium effect of this shopping effort on posted prices. However, if buyers search more for cheap offers, retailers face higher competition and optimally reduce the prices they post. This response matters for distinguishing expenditure and consumption inequality because households do not buy the same basket of goods and retailers can discriminate prices across products. It also matters for understanding the impact of aggregate shocks and policies, as retailers' response determines the full effect of changes in shopping effort on prices.

This paper develops an equilibrium theory of expenditure inequality and price dispersion that accounts for the response of posted prices to the shopping behavior of heterogeneous households. First, I show how retailers post prices taking the level of shopping effort as given and test theoretical predictions against large US scanner data on grocery transactions. Second, I quantify the effect of retailers' price posting on the distinction between expenditure and consumption inequality in equilibrium. Finally, I highlight the implications of heterogeneous shopping effort for the cyclicality of retail prices and markups, as well as for the response of prices to redistributive earnings taxes.

The framework developed in this paper is the first to incorporate frictional goods markets in an incomplete markets economy with endogenous price dispersion for multiple goods. Heterogeneous households decide on their spending, savings, and shopping effort. Shopping effort is subject to a utility cost. Households allocate their total expenditure across multiple varieties of a grocery good and an outside good. Consumption baskets vary systematically across households due to non-homothetic preferences over grocery varieties. The markets for grocery varieties are subject to search frictions, extending the seminal framework of Burdett and Judd (1983) to incorporate household heterogeneity and an intensive margin of demand. For every transaction, households have to search for price quotes and draw either one or two offers simultaneously from the equilibrium distribution of posted prices. Higher shopping effort increases the probability that a household observes two prices and can select the cheaper offer before deciding on the quantity purchased. The price distribution for each variety is determined endogenously as the optimal solution to retailers' price posting problem, which trades off higher margins per sale against undercutting simultaneously observed alternative offers.

The first result of this paper is that with heterogeneity in shopping effort, posted prices depend on the composition of demand. The theory suggests that demand-weighted shopping effort is a sufficient statistic for retailers to take into account rich household heterogeneity. I show that the average posted price decreases in demand-weighted effort, driven by a reduction in profit margins. A higher demand-weighted shopping effort means that the average buyer is more likely to observe two prices and substitute towards a cheaper offer. As a result, retailers face a higher average price elasticity. If a larger share of demand comes from households exerting more shopping effort, retailers' best response is to reduce their margins and post lower prices.

In addition, the skewness of posted price distributions strictly increases in demand-weighted search effort. This result provides a testable prediction that is directly linked to the mechanism generating price dispersion. For price dispersion to exist in equilibrium, retailers have to be indifferent between posting low and high prices within a distribution. To keep retailers indifferent when effort increases, the distribution has to become more dense at the bottom and less dense at the top, i.e. its skewness must increase.

Empirical evidence supports the relationship between the skewness and demand-weighted search effort in micro-data on households' grocery transactions from the Nielsen Consumer Panel. I first show that households with higher expenditure exert lower shopping effort, pay higher prices for identical barcodes, and purchase varieties that are more expensive on average. To test the relationship between households' effort and skewness, I exploit variation in demand-weighted shopping effort across products due to differences in households' consumption baskets. In line with theory, the skewness of local, barcode-level price distributions increases in the share of total expenditure for a given barcode stemming from households with higher search effort.

In equilibrium, households generate demand-weighted shopping effort for each good endogenously in response to the distributions of posted prices. I solve for the equilibrium in households' choices and posted price distributions numerically. To quantify the equilibrium effect of shopping on posted prices, I calibrate the model to evidence from the Nielsen Consumer Panel. I match differences in prices paid within and across varieties and heterogeneity in consumption baskets along the expenditure distribution, as well as price dispersion across products. In addition, the model can account for untargeted moments such as the distribution of expenditures in the data.

The calibrated model shows that the equilibrium response of posted prices doubles the effect of shopping effort on inequality. The literature so far has measured the effect of shopping by focusing on differences in prices paid for a given product (see e.g. Aguiar and Hurst 2007; Arslan et al. 2021; Pytka 2018). Under this definition, shopping reduces the cost of consumption for the bottom versus the top expenditure quintile by 2pp. in the model and in the data. However, because high- and low-spending households do not buy the same products, retailers target their prices to the buyers they face and post lower average margins for products in the basket of low-spending (high-search) households. In the model, I show that differences in margins across varieties account for an additional 2.6pp. of the gap in spending between the top and bottom quintiles.

These price differences map into differences in grocery consumption. In a counterfactual with costless search, where all households observe two price draws for each purchase and retailers respond by posting marginal cost, inequality in grocery consumption increases by 8.31%. As retailers target their prices to the shopping effort of the average buyer, households impose a significant externality on each other. Allowing retailers to discriminate and post targeted price distributions for each household would reduce inequality in grocery consumption by 13.41%. Consumption is an appropriate measure for inequality in welfare, as the model-implied cost of shopping effort is flat along the expenditure distribution.

In addition, accounting for heterogeneity in shopping effort has implications for the cyclicality of prices and markups in response to aggregate shocks. I implement an aggregate shock based on the decline in net worth and losses in labor earnings during the Great Recession. The model generates a 0.6% decline in average prices paid upon impact. 0.5

percentage points are accounted for by changes in posted prices as retailers respond with lower markups to an increase in demand-weighted shopping effort. Only 0.1 percentage points can be attributed to a decline in the average price paid relative to the average posted price. This finding shows how focusing on prices paid relative to prices posted understates the effect of shopping on the cost of consumption over the business cycle.

The change in posted prices is almost entirely driven by the decline in wealth. Losses in earnings have little impact on retailers' price posting despite being of similar magnitude. This result arises because wealth losses are relatively more concentrated at the top of the income distribution and earnings losses at the bottom. In response to a loss in her earnings or wealth, any household increases her search effort and reduces consumption. If low-income households consume less, the composition of demand shifts in favor of high-income households with low search effort, incentivizing retailers to raise prices. In response to earnings losses at the bottom of the distribution, this shift in demand composition offsets the increase in individual search effort. In response to a decline in wealth at the top, the increase in individual effort and the composition effect go in the same direction and unambiguously reduce prices and markups. This result reconciles seemingly conflicting empirical evidence, suggesting procyclical price and markup responses to house price shocks (Stroebel and Vavra 2019) and acyclical responses to unemployment fluctuations (Anderson et al. 2020; Coibion et al. 2015). Overall, composition effects reduce the on-impact response of posted prices to the combined shock by one third.

Finally, I show that the response of posted prices to shifts in demand composition partially compensates net contributors to redistributive policies for the decline in their income. A flat tax on labor earnings, rebated lump-sum to all households, redistributes resources towards low-income (high-shopping-effort) households, increases their share in aggregate demand and hence increases demand-weighted shopping effort. In an economy with a higher level of redistribution, retailers optimally choose to reduce their markups and post lower prices. This channel compensates net contributors in the top quintile of expenditures for 4-11% of the decline in their after-tax earnings.

The paper is structured as follows: Section 1.1 discusses related literature. Section 2 presents the theory and discusses the main mechanism. Section 3 provides empirical evidence on shopping effort and price distributions. Section 4 outlines the calibration and quantitative properties of the model. Section 5 studies the implications for inequality. Section 6 presents the results on cyclicality and policies. Section 7 concludes.

1.1 Related Literature

Search Frictions. Seminal contributions on price search in the goods market include Butters (1977), Varian (1980), and Burdett and Judd (1983). I extend the latter to incorporate household heterogeneity and an intensive margin of demand.[1] In models with price search and rich heterogeneity in income and wealth in the tradition of Bewley (1977) and Aiyagari (1994), Arslan et al. (2021) take posted prices as given and Pytka (2018) endogenizes the price distribution for a single good. Both papers focus on the direct effect of shopping on prices paid. The framework presented in this paper is the first with rich heterogeneity, non-homothetic preferences, and endogenous price distributions for multiple varieties. I employ it to study how the equilibrium response of posted prices to shopping affects inequality. Equilibrium effects of shopping effort on posted prices allow Alessandria (2009) to explain movements of relative prices across countries and Kaplan and Menzio (2016) to generate self-fulfilling unemployment fluctuations. Both setups feature price dispersion for a single good and finite types of shoppers. I show how rich heterogeneity can affect the cyclicality of retail prices due to shifts in demand composition when accounting for the incidence of aggregate shocks.[2]

Non-Homotheticities. Recent work integrates non-homothetic preferences in heterogeneous agent models to study e.g. optimal taxation (Ferriere et al. 2023) or monetary policy (Olivi et al. 2023), but focuses on more aggregated categories of goods.[3] Mongey and Waugh (2022) and Waugh (2023) generate non-homotheticities from logit-preferences in an incomplete-markets economy. None of the previous work considers interactions with shopping effort and price dispersion.

[1] The framework of Burdett and Judd (1983) has been widely applied in macroeconomic research, see e.g. Albrecht et al. (2021), Burdett and Menzio (2018), Menzio (2021), and Kaplan and Menzio (2016).

[2] Huo and Ríos-Rull (2015) develop a framework with heterogeneous households and directed search for quantities and show how shifts in demand composition can affect productivity.

[3] The non-homotheticities at the barcode-level considered in this paper are consistent with evidence in e.g. Argente and Lee (2021), Auer et al. (2022), Faber and Fally (2022), Handbury (2021), and Jaravel (2019). Non-homotheticities at this disaggregated level are often interpreted as taste for quality (Bils and Klenow 2001; Bisgaard Larsen and Weissert 2020; Ferraro and Valaitis 2022; Jaimovich et al. 2019).

Shopping, Retail Prices, and Markups. The paper also extends the empirical literature on shopping, retail prices, and markups. My findings are consistent with previous evidence on heterogeneity in shopping effort and prices paid by Aguiar and Hurst (2005) and subsequent work.[4] Seminal work by Kaplan and Menzio (2015) and Kaplan et al. (2019) provides evidence on price distributions but does not consider their co-movement with demand composition across products. Stroebel and Vavra (2019) find retail prices and markups to respond procyclically to local variations in house prices and attribute this finding to changes in shopping behavior. Anderson et al. (2020) find markups paid to co-vary positively with proxies for local income, driven by differences in products bought. In contemporaneous work, Sangani (2023) provides complementary evidence on higher markups paid by high-income households.[5] Their findings are in line with the theory of this paper. In addition, I provide direct empirical evidence on the main mechanism, testing predictions on the relationship between shopping and the shape of price distributions.

2 A Theory of Inequality and Price Dispersion

A theory of expenditure inequality and price dispersion requires a framework for both the spending choice of households and the price posting decision of retailers. This section introduces households' and retailers' decision problems, solves for the equilibrium distribution of prices, and discusses how shopping effort affects posted retail prices.

2.1 Households with Non-Homothetic Preferences and Shopping

Households are infinitely lived and heterogeneous in their labor earnings zw. w is the common wage rate per unit of labor and z households' idiosyncratic labor productivity, evolving exogenously according to a first order Markov process. Households supply z efficiency units of labor inelastically. In addition, they earn a return r per unit of beginning of period assets a. Households decide jointly on their future asset holdings a', shopping effort s, and their consumption of each variety $j \in J$ of a grocery good $\{c_j\}_{j=1}^{J}$ as well as consumption of an outside (non-grocery) good c_O.

[4]See e.g. Aguiar and Hurst (2007), Aguiar et al. (2013), Broda et al. (2009), Griffith et al. (2009), Nevo and Wong (2019), and Pisano et al. (2022).

[5]Sangani (2023) reconciles these results with an analytical search model with exogenous types and applies this framework to study implications of increasing income inequality for the rise in markups and the distribution of markups across space. My focus is on the feedback from equilibrium prices to inequality and the response of prices to aggregate shocks and redistributive policies.

Households' decision problem can be split into two stages. In a first stage, households divide their resources between savings a' and total expenditure e to solve

$$\begin{aligned} V(z,a) &= \max_{e,a'\geq 0} U(e) + \beta \mathbb{E}_{z'|z} V(z',a') \\ \text{s.t. } & e + a' \leq (1+r)a + zw. \end{aligned} \tag{1}$$

The indirect utility of expenditure $U(e)$ summarizes the second stage, in which households decide on their allocation of consumption across grocery varieties and the outside good as well as their choice for shopping effort, conditional on expenditure.[6] They solve

$$\begin{aligned} U(e) = & \max_{s\in[0,1],\{c_j\}_{j=1}^J, c_O} u(\mathcal{C}) - v(s,\mathcal{C}) \\ \text{s.t.} \quad & \mathcal{C} = (c_G)^{\alpha} (c_O)^{1-\alpha} \\ & c_G = \left[\sum_{j=1}^{J} (\mathcal{C})^{\frac{q_j}{\sigma}} (c_j)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \\ & c_O + \sum_{j=1}^{J} p_j(s) c_j \leq e. \end{aligned} \tag{2}$$

The outside good is taken to be the numeraire and its price is normalized to 1. $u(\cdot)$ are households' preferences over the consumption aggregator $\mathcal{C}$ and $v(s,\mathcal{C})$ is the disutility of exerting shopping effort. I assume that the disutility of effort depends on the level of consumption $\mathcal{C}$ to capture in reduced form that households have to search more often for prices if they have a larger consumption basket.[7] $p_j(s)$ is the per unit price of consuming variety j for a household with shopping effort s, which will be discussed in detail below.

Consumption Allocation. The aggregator $\mathcal{C}$ is a Cobb-Douglas function defined over grocery and non-grocery consumption. Grocery consumption c_G is itself a non-homothetic CES aggregator over varieties $j \in J$ in the spirit of Comin et al. (2021) and Handbury (2021). For given total consumption $\mathcal{C}$ and shopping effort s, it defines a demand system across varieties that can be characterized in terms of expenditure shares ω_j. The optimal

[6] Due to the two stage setup, shopping effort and consumption baskets depend only on spending. The structure allows me to focus on data moments of the expenditure distribution when disciplining the second stage of the household problem.

[7] This mechanism is micro-founded in Pytka (2018) who also provides evidence that conditional on employment high-income households spend more time making purchases and rules out that this is due to shopping as a leisure activity.

allocation satisfies

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

Varieties should be considered close substitutes and can be thought of as different barcodes within a narrowly defined group of products, e.g. different varieties of soft drinks.[8] The parameters $\{q_j\}_{j=1}^J$ govern the expenditure elasticity of demand: With $\mathcal{C}$ increasing in expenditures, the relative expenditure share of variety j vs. variety k $\left(\frac{\omega_j}{\omega_k}\right)$ is increasing in total spending e iff $q_j > q_k$. In line with the literature's interpretation of more expensive varieties among close substitutes as higher quality products, I will refer to varieties with a high q_j as *high-quality.*[9] Under this interpretation, the non-homotheticities considered arise because high-spending households have a stronger taste for quality.

The implied price of the optimal grocery consumption bundle (of one unit c_G) is given as

$$p_G(\mathcal{C}, s) = \left(\sum_{j=1}^{J} \mathcal{C}^{q_j} \left(p_j(s)\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}.$$

The Cobb-Douglas aggregator $\mathcal{C}$ yields optimal shares of expenditures on groceries e_G and the outside good e_O given by

$$\omega_G = \frac{e_G}{e} = \frac{p_G(\mathcal{C}, s) c_G}{e} = \alpha \quad \text{and} \quad \omega_O = \frac{e_O}{e} = \frac{c_O}{e} = 1 - \alpha.$$

For given shopping effort s and expenditure level e, the consumption aggregator is the solution to a non-linear equation

$$\mathcal{C} = \frac{e}{P(\mathcal{C}, s)}, \quad \text{with} \quad P(\mathcal{C}, s) = \left(\frac{p_G(\mathcal{C}, s)}{\alpha}\right)^{\alpha} \left(\frac{1}{1-\alpha}\right)^{1-\alpha}. \tag{3}$$

[8] Appendix B.5 compares consumption baskets along the expenditure distribution in the Nielsen Consumer Panel and shows that significant non-homotheticities arise when defining a product as a barcode as compared to aggregating goods at the level of close substitutes. This is in line with e.g. Jaravel (2019).

[9] See e.g. Bils and Klenow (2001), Bisgaard Larsen and Weissert (2020), or Argente and Lee (2021).

Shopping Effort. Given $p_j(s)$, households' optimal choice of effort equates the marginal benefit and disutility of shopping such that

$$\underbrace{-v_s(s,\mathcal{C})}_{\substack{\text{marginal disutility}\\ \text{of shopping}}} = \underbrace{\frac{\partial P(\mathcal{C},s)}{\partial s}\mathcal{C}}_{\substack{\text{change}\\ \text{in BC}}} \underbrace{\frac{u_{\mathcal{C}}(\mathcal{C}) - v_{\mathcal{C}}(s,\mathcal{C})}{P(\mathcal{C},s) + \frac{\partial P(\mathcal{C},s)}{\partial \mathcal{C}}\mathcal{C}}}_{\substack{\text{marginal benefit}\\ \text{of relaxing BC}}}. \tag{4}$$

The benefit of shopping is the change in the budget constraint times the marginal benefit of additional available resources, here expressed in terms of an increase in the consumption aggregator $\mathcal{C}$. Increasing shopping effort reduces the cost of consumption and relaxes the budget by $\frac{\partial P(\mathcal{C},s)}{\partial s}\mathcal{C}$, where the change in the price index depends on households' consumption basket and the return to shopping for each variety.[10]

Return to Search. The relationship between prices paid for variety j and shopping effort $p_j(s)$ is an equilibrium object and depends on the distribution of posted prices as well as the process with which households sample from this distribution. Households split their total consumption of variety c_j over a unit continuum of transactions τ.[11] For each transactions, a household has to search for prices and samples either one or two draws at random from the equilibrium distribution of posted prices for variety j, denoted $F_j(p)$. Households' shopping effort s determines the share of transactions for which a household observes two prices. Households select the lowest observed price as the effective transaction price and can purchase any quantity at this price. The implied distribution of effective prices across transactions of variety j under effort s is given as

$$G_j(p|s) = (1-s)F_j(p) + sF_j^m(p),$$

where the minimum of two draws from $F_j(p)$ is distributed with $F_j^m(p) = (1-(1-F_j(p))^2)$.

[10] At the optimal solution, the marginal benefit could also be expressed in terms of any of the grocery varieties, the composite grocery good c_G, or the outside good c_O.

[11] Transactions can be interpreted as multiple purchases of the same variety or purchases of vertically differentiated subvarieties of j with identical marginal cost and equilibrium price distributions.

Households have CES preferences over transactions and allocate their spending optimally to minimize the cost of consumption, i.e. households solve

$$\min_{\{c_{j\tau}\}} \int_0^1 p_{j\tau} c_{j\tau} d\tau$$
$$\text{s.t. } c_j = \left(\int_0^1 (c_{j\tau})^{\frac{\sigma_j - 1}{\sigma_j}} d\tau \right)^{\frac{\sigma_j}{\sigma_j - 1}}$$

σ_j is the price elasticity of demand across transactions of variety j. Optimal demand for a transaction with effective price p is given as

$$c_{j\tau} = \left(\frac{p_{j\tau}}{p_j(s)} \right)^{-\sigma_j} c_j, \tag{5}$$

where $p_j(s)$ is the price per unit of c_j given by

$$\begin{aligned} p_j(s) &= \left[\int_0^1 p_{j\tau}^{1-\sigma_j} d\tau \right]^{\frac{1}{1-\sigma_j}} = \left[\int p_{j\tau}^{1-\sigma_j} dG_j(p_{j\tau}|s) \right]^{\frac{1}{1-\sigma_j}} \\ &= \left[(1-s) \int p_{j\tau}^{1-\sigma_j} dF_j(p_{j\tau}) + s \int p_{j\tau}^{1-\sigma_j} dF_j^m(p_{j\tau}) \right]^{\frac{1}{1-\sigma_j}} \end{aligned} \tag{6}$$

As $F_j(p)$ is first order stochastically dominant to $F_j^m(p)$ by definition, prices paid per unit of c_j are decreasing in search effort $\left(\frac{\partial p_j(s)}{\partial s} < 0 \right)$ for any $\sigma_j > 0$.[12]

Assuming a continuum of transactions to eliminate price risk at the household level builds on Pytka (2018). A key difference is that I allow households to adjust quantities based on the effective transaction price, equating the marginal benefit of consumption across transactions with different prices as opposed to imposing equal demand at each purchase. Introducing this quantity choice allows me to extend the canonical price-posting model of Burdett and Judd (1983) to an intensive margin of demand below.

2.2 Retailers' Problem and Posted Price Distributions

Each of the grocery varieties is traded in a separate market in the spirit of Burdett and Judd (1983). Consider the market for variety j, which is sold by a continuum of single-product retailers of mass M_j. Retailers are ex-ante identical, face constant marginal cost

[12] The moments $\mathbb{E}\left[p^{1-\sigma_j}|F_j^m(p)\right]$ and $\mathbb{E}\left[p^{1-\sigma_j}|F_j(p)\right]$ are sufficient statistics to construct the price index as a function of shopping effort $p_j(s)$. This feature simplifies the solution of the model significantly as households do not need to keep track of the entire price distribution.

κ_j, and can freely enter and exit the market subject to a per-period fixed cost of operating K_j. The demand side of the market consists of heterogeneous households as introduced above. Retailers in the market for j take households' choices as well as the distribution of households as given. Households' relevant choices are determined conditional on their expenditure e and it is sufficient to index these choices (households' types) by their spending, with $\lambda(e)$ being the equilibrium distribution over expenditures.

Retailers' Problem. Retailers commit to a price for variety j before meeting any buyers. The expected total profits of a retailer posting price p for variety j are given by

$$\pi_j(p) = (p - \kappa_j) \int \underbrace{\lambda(e) \frac{(1 - s(e)) + s(e)2(1 - F_j(p))}{M_j}}_{\text{extensive margin}} \underbrace{\left[\left(\frac{p}{p_j(s(e))} \right)^{-\sigma_j} c_j(e) \right]}_{\text{intensive margin}} de - K_j,$$

In words, profits are given as the margin per sale $(p - \kappa_j)$ times demand at price p less of the cost of operating. To determine demand at p, retailers need to take expectations about the expenditure type e of the households they meet. Each type e household has $(1 + s(e))$ meetings with a retailer, split equally across M_j retailers, $(1 - s(e))$ of which are linked to transactions with only a single price and $2s(e)$ to transactions with two observed prices. The probability that a second observed price is above p is given by $(1 - F_j(p))$. A retailer posting p hence makes positive sales in $(1 - s(e)) + s(e)2(1 - F_j(p))$ meetings with type e households. Conditional on purchasing from a retailer at price p, households optimally buy $\left(\frac{p}{p_j(s(e))}\right)^{-\sigma_j} c_j(e)$ units to minimize the cost of consumption across transactions. Retailers' profits from posting price p can be simplified to

$$\pi_j(p) = \underbrace{\frac{\bar{C}_j}{M_j}}_{\substack{\text{demand} \\ \text{per retailer}}} \underbrace{\left[(1 - \bar{s}_j) + \bar{s}_j 2(1 - F_j(p)) \right] p^{-\sigma_j}}_{\text{market share}} \underbrace{(p - \kappa_j)}_{\substack{\text{profit} \\ \text{margin}}} - K_j,$$

where $\bar{C}_j = \int \lambda(e) \left[p_j(s(e)) \right]^{\sigma_j} c_j(e) de$ is a notion of total demand for variety j and

$$\bar{s}_j = \int \frac{\lambda(e) \left[p_j(s(e)) \right]^{\sigma_j} c_j(e)}{\bar{C}_j} s(e) de \qquad (7)$$

is the *demand-weighted average search effort* in the market. Deciding on the price to post in this market, retailers trade off between margins per sale and their market share. A higher price increases the margin earned per sale $(p - k_j)$ but decreases demand at both the intensive and extensive margin. The price p that retailers post affects their demand at the *intensive margin*, as households trade off the optimal quantities to purchase across transactions. At the *extensive margin*, the posted price p determines the likelihood $(1 - F_j(p))$ to be the cheaper of two price observations in case a household observes a second price for the current transaction. Demand-weighted search effort $\bar{s}_j$ affects the strength of the second effect as it determines the likelihood that the average buyer observes two prices and therefore the likelihood any retailer has to compete for a transaction. In this sense, $\bar{s}_j$ determines the price elasticity of demand at the extensive margin. While the extensive margin demand elasticity is standard in models building on Burdett and Judd (1983), the intensive margin elasticity is novel to this paper.

Posted Price Distribution. As retailers are homogeneous, a non-degenerate equilibrium price distribution requires them to be indifferent between posting a range of prices, i.e. any price on the support of $F_j(p)$ must yield identical profits. For any price increase on the support of $F_j(p)$, the benefit of earning a higher margin on the current market share has to be exactly offset by the cost of a loss in market share $(\frac{\partial \pi(p)}{\partial p} = 0)$,

$$\underbrace{[(1 - \bar{s}_j) + \bar{s}_j 2(1 - F_j(p))]\, p^{-\sigma_j}}_{\text{current market share}} = \underbrace{\left[\bar{s}_j 2 f_j(p) p^{-\sigma_j} + [(1 - \bar{s}_j) + \bar{s}_j 2(1 - F_j(p))]\, \sigma_j p^{-\sigma_j - 1}\right]}_{\text{loss in market share}} \underbrace{(p - \kappa_j)}_{\substack{\text{current} \\ \text{margin}}}. \tag{8}$$

This differential equation characterizes the equilibrium distribution of posted prices. Demand-weighted average shopping effort $\bar{s}_j$ summarizes all relevant information about the distribution of households and is a sufficient statistic for the retailer to post a price.[13] Solving retailers optimality conditions for the posted price distribution $F_j(p)$ follows similar steps as in Burdett and Judd (1983) and Pytka (2018). First, consider the upper bound of the posted price distribution $\bar{p}_j$.

Lemma 1 *The upper bound of the posted price distribution is given by* $\bar{p}_j = \frac{\sigma_j}{\sigma_j - 1} \kappa_j$.
Proof. *Consider the problem of a retailer that posts the highest price* $\bar{p}_j$. *By definition,* $F_j(\bar{p}_j) = 1$. *Profits at the highest price simplify to*

$$\pi_j(\bar{p}_j) = \frac{\bar{C}_j}{M_j}(1 - \bar{s}_j)\bar{p}_j^{-\sigma_j}(\bar{p}_j - \kappa_j) - K_j$$

[13]Retailers are infinitesimal and take market size $\frac{\bar{C}_j}{M_j}$ as given.

which is maximized at $\bar{p}_j = \frac{\sigma_j}{\sigma_j - 1}\kappa_j$*. If the highest price is below* $\bar{p}_j$ *profits could be increased by moving to* $\bar{p}_j$ *as an increase would not affect demand at the extensive margin. If the highest price was to be above* $\bar{p}_j$*, profits could be increased by reducing the price to* $\bar{p}_j$ *even without considering any potential gain in demand at the extensive margin. Hence* $\bar{p}_j$ *must be the highest price posted in any equilibrium.* ■

The highest posted price is equivalent to the optimal price under monopolistic competition. Unlike in Burdett and Judd (1983) and Pytka (2018), the upper bound of the price distribution is not imposed exogenously but arises endogenously due to retailers optimal tradeoff between profit margins and demand at the intensive margin. Given the upper bound $\bar{p}_j$, a zero profit condition determines entry.

Lemma 2 *The mass of active retailers is given by* $M_j = \frac{\bar{C}_j}{K_j}(1 - \bar{s}_j)\bar{p}_j^{-\sigma_j}(\bar{p}_j - \kappa_j)$*.*
Proof. *Under free entry,* M_j *must satisfy* $\pi_j(\bar{p}_j) = 0$*.* ■

In addition, the properties of $F_j(p)$ carry over from the case without an intensive margin of demand as discussed in Burdett and Judd (1983) and Pytka (2018).

Lemma 3 $F_j(p)$ *is continuous and its support is a connected set.*
Proof. *The proof is equivalent to the argument in Pytka (2018), relying on the presence of profitable deviations in case* $F_j(p)$ *violates either property.* ■

Under Lemmas 1-3, Proposition 1 characterizes the distribution of posted prices.

Proposition 1 *The equilibrium distribution of posted prices is given by*

$$F_j(p) = \begin{cases} 0 & \forall\ p < \underline{p}_j \\ 1 - \frac{1-\bar{s}_j}{2\bar{s}_j}\left[\left(\frac{\bar{p}_j}{p}\right)^{-\sigma_j}\frac{\bar{p}_j - \kappa}{p - \kappa} - 1\right] & \forall\ p \in [\underline{p}_j, \bar{p}_j] \\ 1 & \forall\ p > \bar{p}_j \end{cases} \tag{9}$$

where the lower bound of the support $\underline{p}_j$ *solves*

$$\underline{p}_j^{-\sigma_j}(\underline{p}_j - \kappa) = \frac{1 - \bar{s}_j}{1 + \bar{s}_j}\bar{p}_j^{-\sigma_j}(\bar{p}_j - \kappa) \tag{10}$$

Proof. *Follows from imposing* $\pi_j(p) = \pi_j(\bar{p}_j) = 0$ *for all prices on the support of* $F_j(p)$ *as well as* $F_j(\underline{p}_j) = 0$*. In addition, it must hold that* $\pi_j(p) < 0\ \forall\ p > \bar{p}_j$ *by Lemma 1 and*

$\pi_j(p) < 0 \; \forall \; p < \underline{p}_j$ as a decrease in price at any $p < \bar{p}_j$ reduces $p^{-\sigma_j}(p - \kappa_j)$ and any reduction in prices below $\underline{p}_j$ does not increase demand at the intensive margin. ■

Retailers play a mixed strategy, randomizing prices over the interval $[\underline{p}_j, \bar{p}_j]$ according to the density $f_j(p)$ associated with $F_j(p)$. The distribution of posted prices depends on the marginal cost κ_j and the intensive margin price elasticity σ_j, as well as demand-weighted shopping effort $\bar{s}_j$. It is independent of total demand $\bar{C}_j$ and the cost of operating K_j.[14]

Production and Equilibrium. All grocery varieties are produced and sold to retailers at marginal cost by competitive production firms. The outside good is produced by competitive production firms and sold directly to households in markets without any search frictions. Producers operate linear technologies with labor N as input. I assume production functions $y_O = N_O$ and $y_j = \frac{1}{\kappa_j} N_j$. With the outside good as the numeraire, the equilibrium wage is determined as $w = 1$ and the marginal cost of producing a unit of variety j is given by κ_j. Households' assets are invested in a risk-free bond at exogenous interest rate r. Under these assumptions, the model outcomes can be interpreted as the equilibrium of a small open economy or as the equilibrium in a subregion (state) of a large economy like the US. Appendix C.1 defines an equilibrium formally.

2.3 The Mechanism: Shopping Effort and Posted Prices

The main mechanism of this paper describes how household heterogeneity affects posted price distributions. As demand-weighted effort $\bar{s}_j$ is a sufficient statistic for retailers' pricing decision, the answer to this question is determined in two steps: (i) How does $\bar{s}_j$ change with the distribution of households? and (ii) How does the distribution of posted prices respond to changes in $\bar{s}_j$?

Demand-Weighted Shopping Effort. Focus first on how the distribution of households determines $\bar{s}_j$. Equation (7) implies that a retailer takes into account type e households' shopping effort according to their share in total demand $\frac{\lambda(e)[p_j(s(e))]^{\sigma_j} c_j(e)}{\bar{C}_j}$. From here on out, I will refer to the vector of these shares as *demand composition.* Differences in demand composition shift the weights attached to each household's idiosyncratic search behavior. In this way, heterogeneity in individual effort $s(e)$ creates a role for demand

[14]Beyond ensuring zero profits, the parameter K_j does not influence equilibrium allocations and only the product $M_j K_j$ is uniquely determined by $\pi_j(\bar{p}_j) = 0$ in equilibrium. All results are equivalent to assuming a fixed number of retailers who distribute profits outside of the economy.

composition to affect $\bar{s}_j$ and through it posted prices. $\bar{s}_j$ is higher if a larger share of demand is accounted for by households with a higher shopping effort.

Consider next an increase in spending only by households of type e. For convenience, define $\hat{\lambda}(e) = \lambda(e)\,[p_j(s(e))]^{\sigma_j}\,c_j(e)$ such that $\bar{C}_j = \int \hat{\lambda}(e)de$. The derivative of $\bar{s}_j$ w.r.t. to a marginal increase in spending of all households of type e is given by

$$\frac{\partial \bar{s}_j}{\partial e} = \frac{\hat{\lambda}(e)}{\bar{C}_j}\frac{\partial s(e)}{\partial e} + \frac{1}{\bar{C}_j}(s(e) - \bar{s}_j)\frac{\partial \hat{\lambda}(e)}{\partial e}. \tag{11}$$

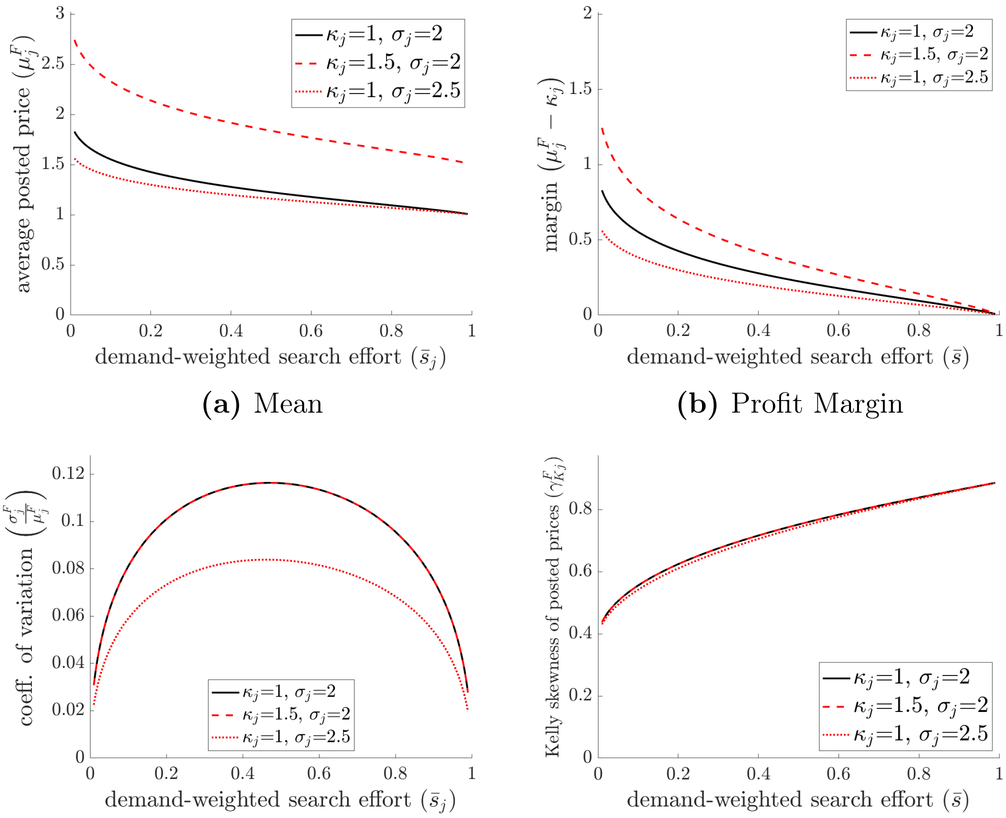
The first term is the direct effect of a change in type e's shopping effort, which is weighted by their share in demand. The second term captures the indirect effect through shifts in demand composition, interacting type e's position in the distribution of shopping effort with their changing weight in demand. Consider $\frac{\partial s(e)}{\partial e} < 0$ and a normal good j, so that $\frac{\partial \hat{\lambda}(e)}{\partial e} > 0$: An increase in spending of type e unambiguously decreases $\bar{s}_j$ if $s(e) < \bar{s}_j$, but is ambiguous otherwise as changes in individual effort and demand composition offset each other. Taking into account shifts in demand composition and households' relative shopping effort is crucial to assess how changes in the distribution of expenditure affect $\bar{s}_j$ and through it posted prices.

Shopping Effort and Posted Prices. Taking households' behavior as given, we can assess how changes in $\bar{s}_j$ affect posted prices. Figure 1 plots selected moments of the posted price distribution $F_j(p)$ for different values of κ_j, σ_j, and $\bar{s}_j$. The average posted price is a strictly decreasing function of demand weighted shopping effort for any (κ_j, σ_j). The effect of shopping on the average price posted operates through changes in profit margins, which are strictly decreasing in $\bar{s}_j$. An increase in $\bar{s}_j$ makes it more likely that the average buyer observes a second price, increases the price elasticity at the extensive margin, and tilts sellers' tradeoff between higher margins and retaining market share in favor of the latter. This is the main mechanism of the paper: If a larger share of demand is accounted for by high-search households, retailers optimally post lower prices.[15]

The setup limits in well known special cases: If all buyers observe two prices simultaneously ($\bar{s}_j = 1$), retailers face Bertrand competition, post marginal cost κ_j, and make zero profits per sale. If no buyer observes two prices simultaneously ($\bar{s}_j = 0$), the economy

[15]Figure 13 in Appendix C.3 highlights the shift in price distributions graphically.

collapses to monopolistic competition and all retailers charge a markup proportionate to the intensive margin elasticity, posting $\bar{p}_j = \frac{\sigma_j}{\sigma_j - 1}\kappa_j$. Shopping effort shifts the economy between these extreme cases endogenously. The limiting cases show that imperfect elasticities at the intensive and extensive margin interact in generating profit opportunities for retailers, with perfect elasticity at either margin ($\bar{s}_j = 0$ or $\sigma_j = \infty$) profits are zero.



(a) Mean

(b) Profit Margin

(c) Coefficient of Variation

(d) Kelly Skewness

Figure 1: Moments of the Posted Price Distribution

Note: Theoretical moments of the posted price distribution $F_j(p)$ as a function of demand-weighted shopping effort $\bar{s}_j$, for different values of marginal cost κ_j and intensive margin elasticity σ_j.

The model has additional implications for the shape of price distributions that can be used to discipline the theory empirically. The coefficient of variation is hump shaped in $\bar{s}_j$, due to the limiting cases without price dispersion discussed above. It decreases in σ_j as a higher intensive margin elasticity reduces the ration between the upper bound of posted prices and marginal cost $\frac{\bar{p}_j}{\kappa_j} = \frac{\sigma_j}{\sigma_j - 1}$ and hence the range of possible prices.

The skewness of the posted price distribution is strictly increasing in $\bar{s}_j$. The intuition for this finding goes back to retailers' indifference condition in equation (8). Given a distribution of posted prices, an increase in demand-weighted search effort increases sales of retailers with low prices and decreases them for retailers with high prices, as households on average find cheaper offers. This increases the benefit of raising prices at the bottom and decreases the benefit at the top. To offset this effect and keep retailers indifferent between posting low and high prices, the loss in market share when raising prices has to increase at the bottom and decrease at the top. This requires the distribution of posted prices to be more dense at the bottom and less dense at the top. A more (less) dense distribution increases (decreases) the number of competitors that undercut a retailer when raising prices marginally. A distribution that is more dense at the bottom and less dense at the top exhibits higher skewness. Other parameters have quantitatively negligible effect on the skewness, making it well suited for an empirical test of the mechanism.[16]

Robustness. Appendix A.1 derives moments of the posted price distribution in closed form for the canonical model of Burdett and Judd (1983) without an intensive margin of demand, confirming the relationship between search effort and the moments of posted price distributions.[17] Appendix A.2 provides an extension of the canonical model to heterogeneity in marginal cost. Under reasonable calibrations heterogeneous cost leave average prices decreasing and the skewness increasing in shopping effort.[18]

3 Data on Inequality, Shopping, and Price Dispersion

This section provides empirical evidence to discipline the theory outlined above. First, I study the role of price differences in accounting for inequality in expenditures. Second, I test the relationship between shopping and the shape of price distributions empirically.

Data. The results presented in this section rely on data from the Nielsen Consumer Panel for 2007-2019, providing information on the grocery purchases of 60,000 US households per year. The data records quantities and prices at the barcode-transactions level, together with annual information on household and product characteristics.[19]

[16]I report Kelly's measure of skewness, which is defined based on the difference between 90th, 50th, and 10th percentile of the distribution and robust to extreme observations. This property makes it appealing for the empirical test conducted in Section 3. Other measures of skewness such as Pearson's moment coefficient of skewness (the standardized third moment) have qualitatively similar properties.

[17]An analytical solution is infeasible with an intensive margin of demand as $\underline{p}_j$ is defined implicitly.

[18]I focus on homogeneous marginal cost as they 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 and more commonly applied state legislation (Nakamura 2008).

[19]Further information on the dataset is provided in Appendix B.1.

3.1 Expenditure Inequality, Prices, and Shopping

A decomposition of households' grocery expenditures shows how price differences within and across varieties contribute to differences in spending across households. I define a variety as a barcode. To make per-unit prices comparable across varieties, I compare only barcodes within a Nielsen product module measured in the same unit and normalize prices and quantities by the size of a product.[20] Total annual grocery expenditure e_i^G of household i is the sum of spending over all barcodes j in all module-unit bins k. Spending per barcode is the quantity-weighted per-unit price p_{ijk} paid by the household for barcode j in bin k times the units consumed c_{ijk}. Define $\hat{p}_{jk}$ as the national, annual, quantity-weighted average price paid for variety j across all households and $\tilde{p}_k$ as the average of $\hat{p}_{jk}$ within bin k.[21] Decompose e_i^G as

$$e_i^G = \textstyle\sum_k \sum_{j \in J_k} p_{ijk} c_{ijk} = \underbrace{\sum_k \sum_{j \in J_k} (p_{ijk} - \hat{p}_{jk}) c_{ijk}}_{\text{within varieties}} + \underbrace{\sum_k \sum_{j \in J_k} (\hat{p}_{jk} - \tilde{p}_k) c_{ijk}}_{\text{across varieties}} + \underbrace{\sum_k \sum_{j \in J_k} \tilde{p}_k c_{ijk}}_{\text{quantities}}. \quad (12)$$

The first term captures price differences within barcodes between what an individual household pays relative to other households. The second term captures differences across barcodes, due to heterogeneity in consumption baskets. The last term summarizes counterfactual spending absent any price differences across purchases in bin k, with any remaining variation within a bin k due to differences in the quantity consumed.

The first panel of Table 1 reports results of the decomposition by quintile of expenditures, expressed as a fraction of grocery spending. Households at the bottom of the distribution have about 5% lower expenditure due to deviations from the average price $\tilde{p}_k$ within module-unit bins. 1.2pp. are due to lower prices within the same barcode. 3.7pp. are due to lower prices across barcodes purchased. At the top of the distribution, price differences increase total spending by 4.3%, 3.5pp. of which due to differences across barcodes.[22]

[20] E.g. I group together all barcodes in the module "fresh apples" that are measured in pieces and divide the price of each barcode by the number of individual apples included to get a price per apple. Most product modules have one dominant unit of measurement and there is no systematic difference in purchases across households in the unit dimension as Appendix B.5 shows.

[21] For the module "fresh apples" measured in pieces, $\tilde{p}_k$ is the average price per apple across all households and all barcodes of apples, $\hat{p}_{jk}$ the average price for one specific barcode of apples across households, and p_{ijk} the average price one specific household pays for one specific barcode.

[22] The reported magnitudes are well in line with the findings of e.g. Aguiar and Hurst (2007) and Bisgaard Larsen and Weissert (2020) under alternative approaches.

Table 1: Prices and Shopping by Expenditure Quintile

	expenditure quintiles				
	Q1	Q2	Q3	Q4	Q5
within variety	-1.256	-0.698	-0.411	-0.054	0.735
	(0.051)	(0.039)	(0.034)	(0.032)	(0.030)
across varieties	-3.712	-2.574	-1.495	0.167	3.517
	(0.094)	(0.070)	(0.069)	(0.062)	(0.059)
both	-4.968	-3.272	-1.906	0.113	4.253
	(0.112)	(0.086)	(0.081)	(0.075)	(0.072)
trips per purchase	0.189	0.145	0.136	0.131	0.130
	(0.001)	(0.0004)	(0.0003)	(0.0003)	(0.0003)

Note: First panel: Decomposition in (12), reported in % of households' total grocery spending and averaged by quintile. Second panel: Trips per purchase (store-days with transactions over store-day-barcodes with transactions). Quintiles sorted within waves on annual spending divided by the square root of household size, applying Nielsen sampling weights. Sample: Waves 2007-2019, households with heads age 25-65. Aggregated with Nielsen sampling weights within waves, each wave weighted equally. Standard errors in brackets bootstrapped with 1000 replications.

Measuring inequality as the ratio between the top and bottom quintile (e.g. Aguiar and Bils 2015), the results of Table 8 imply that price differences can account for about 10% of expenditure inequality. The ratio in spending between the top and bottom quintile would be about 10% lower without variation in prices and identical consumption bundles. This is significant, given the price differences considered are among a narrowly defined set of almost identical products.

Higher prices on a given basket of goods necessarily increase expenditure. Still, the relationship between prices and expenditure levels is not trivially positive: High-spending households could buy a larger basket (higher quantity) but pay less for each individual product or buy cheaper products within each module. The size of the basket is controlled for by dividing each term in (12) by the level of spending.

Shopping Effort. Nielsen does not provide direct information on the time spent searching for prices. Instead, I rely on two proxies for shopping effort. First, I observe the outcome of the search process – households with higher shopping effort paying less for identical items. The price differences within varieties reported in Table 1 suggest that households with higher spending exert lower effort. An alternative measure of shopping effort is the number of shopping trips a household makes to purchase a given number of items. Kaplan and Menzio (2015) show that an effective way to reduce prices paid is to visit more stores or the same store more often. I define a shopping trip as a visit to a unique store on a unique day. To control for the size of consumption baskets, I divide annual trips by households' purchases, where a purchase is a transaction involving a unique barcode in a given store on a given day. The last panel of Table 1 shows that trips per purchase decrease along the expenditure distribution, the bottom quintile making on average 1.89 trips to purchase 10 items compared to 1.3 for the top quintile.

Robustness. Tables 9 and 10 in Appendix B.3 show that the findings are robust to controlling for other household characteristics commonly associated with shopping effort, such as income or age of the households. Measuring price differences for the same barcode based on a price index as in Aguiar and Hurst (2005) yields quantitatively similar results. Variation in expenditures remain an important driver of price differences even when controlling for further demographics, motivating a theory build around expenditure inequality.[23] Table 8 in Appendix B.3 reports results defining a variety as all barcodes of a given brand within a module. Defining a product as a brand instead of a barcode leaves the overall contribution of price differences to expenditure inequality unchanged, but leads to higher price differences within products and smaller price differences across products by construction. As I interpret price differences within products as a measure of search effort, I focus on the barcode based definition as the conservative approach for the remainder of the paper. For the baseline results, I compute average prices for a barcode at the national level. Appendix B.2 shows that decompositions based on local average prices yield similar results compared to those under national average prices once controlling for a potential small-sample bias at the local level. Considering the tradeoff of introducing a bias into estimates at the local level and the similarity of unbiased estimates to the national definition, I apply national average prices for all baseline results.

[23]The strong relation between households' shopping effort and expenditure levels is in line with the mechanism introduced in Pytka (2018). As households with higher expenditure make more purchases, they have to search more often to achieve the same average reduction in prices. Hence, reducing the average price paid becomes more costly as the size of a household's basket increases. I incorporate this effect in reduced form in the model by allowing the disutility of effort to depend on the size of the consumption bundle. In the calibrated model, I show that this effect is quantitatively dominated by differences in the marginal utility and price of consumption along the distribution, which should be picked up by expenditure levels empirically.

3.2 Demand Composition and Price Distributions

Testing for a reduction in retailers' margins in response to higher shopping effort would require data on markups at the seller-good level. For a test of the relationship between the skewness of price distributions and effort, it is sufficient to observe price distributions and demand-weighted shopping effort. Both are available in the Nielsen dataset.

Demand-Weighted Shopping Effort. According to theory, retailers should consider the shopping effort of households weighted by their share in overall demand for the variety they sell. I exploit variation in demand composition across products and compute for each barcode the national, annual expenditure shares stemming from different groups of households, sorted by their shopping effort.[24] Building on the results in Section 3.1 and the previous literature, I consider separately the five quintiles of the expenditure distribution, four bins of household income, as well as the number of non-employed household heads.[25] In addition, I sort households into quintiles by shopping effort, measured either based on differences in prices paid for the same barcode or by trips per purchase. To be in line with the predictions from Section 2.3, the skewness of price distributions should be decreasing in the expenditure share coming from high-spending, high-income or high-price households, but increasing in the share of demand from households with more non-employed heads or more trips per purchase.

Price Distributions. A price distribution consists of all transactions observed for a barcode j, within a region r and time period t. In line with Kaplan and Menzio (2015), I define a region as a Scantrack Market Area (SMA) and the time period to be a quarter.[26] The price associated with a transaction is defined as the total amount paid less of coupon values, divided by the quantity purchased. To control for outliers, I drop all transactions

[24] I use national-annual shares for two reasons: First, many retailers engage in uniform price setting across locations (see e.g. DellaVigna and Gentzkow 2019), making national rather than local demand composition the relevant statistic for their price setting. Second, national-annual spending shares do not exclusively rely on the same transactions used to construct the local-quarterly price distributions, serving as an instrument for local demand composition.

[25] See e.g. Aguiar and Hurst (2005, 2007), Kaplan and Menzio (2016) and Pytka (2018).

[26] The choice for what definition of a region and which time period to consider trades off between two forces: A narrow definition ensures that any variation in prices can be confidently 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.

for which the reported amount paid less of coupons is zero or negative. For the baseline analysis, I consider all price distributions containing at least 25 transactions and compute Kelly's measure of skewness for each distribution, weighting individual price observations with household weights and quantities purchased.

Estimation. To test for the predicted relationship between skewness and shopping, 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 $y(t)$. I run separate regressions defining groups based on expenditure or shopping quintiles, income bins, and the number of non-employed household heads, excluding the lowest quintile, the lowest income bin and households with no non-employed head respectively as a baseline, and estimate

$$skew_{j,r,t} = \theta_m + \mu_{r,t} + \sum_{g=2}^{G} \beta_g share_{j,g,y(t)} + \varepsilon_{j,r,t} \tag{13}$$

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 (θ_m).[27] The included fixed effects demean the skewness by product category and by region at a given point in time. Therefore, the coefficients of interest β_g are identified by the covariation of demand composition and differences in the skewness of distributions among closely substitutable barcodes within a given region and period.

Table 2 reports the results. The skewness of price distributions is monotonically decreasing in the share of expenditure stemming from high-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 0.33 relative to a barcode bought entirely by the first quintile. All differences w.r.t. the baseline group are statistically significant at the 1%-level. The finding is robust to measuring expenditure shares conditional on the (male) household head being between age 25-65 to account for spending patterns of student and retiree households (2). Similar findings pertain when sorting households by price differences within barcodes (3), or

[27]Nielsen-defined product modules are the first level of aggregation above barcodes 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.

Table 2: Demand Composition and the Skewness of Price Distributions

	by expenditures		by shopping			by income		by employment
	all ages	working age	prices	trips		(working age)		(all ages)
	(1)	(2)	(3)	(4)		(5)		(6)
quintile 2	−0.125***	−0.130***	-0.143***	0.066***	income	−0.012	1 non-employed	0.033***
	(0.035)	(0.031)	(0.022)	(0.023)	30k-60k	(0.018)	household head	(0.015)
quintile 3	−0.195***	−0.175***	−0.306***	0.096***	income	−0.012	2 non-employed	0.133***
	(0.032)	(0.029)	(0.020)	(0.023)	60k-100k	(0.018)	household heads	(0.027)
quintile 4	−0.244***	−0.223***	−0.342***	0.172***	income	−0.038**		
	(0.029)	(0.027)	(0.019)	(0.021)	>100k	(0.016)		
quintile 5	−0.330***	−0.299***	−0.261***	0.168***				
	(0.027)	(0.025)	(0.017)	(0.019)				
FE product module	X	X	X	X		X		X
FE quarter-SMA	X	X	X	X		X		X
Observations	3,026,551	3,026,404	3,026,551	3,026,551		3,026,404		3,026,551

Note: Regression of Kelly's measure of skewness on demand shares by household groups. Price distributions defined as all transactions of a barcode within a Scantrack Market Region and quarter. Demand shares defined as the share of national annual spending on a barcode by each group of households. Data from Nielsen Consumer Panel waves 2007-2019. Observations weighted by total sales in given price distribution. Standard errors clustered at the barcode-year level. *p<0.1; **p<0.05; ***p<0.01.

income group (conditioning on working age households, (5)). In addition, the skewness is increasing in the share of demand from households with more trips per purchase (4) or more non-employed household heads (6). All specifications suggest one conclusion: The skewness of price distributions decreases in the share of expenditure stemming from low-effort households. This confirms the predictions of Section 2.3 and provides strong evidence in favor of the theoretical relationship between search effort and posted prices.

Robustness. In Appendix B.4, I report further robustness with respect to how the skewness of price distributions is measured. Table 11 reports results for Pearson's measure of skewness.[28] Table 12 reports results controlling for barcode instead of product-module fixed effects. Table 13 replicates column (1) of Table 2 under different weights and for price distributions from the 2019 wave of the Nielsen Retail Scanner Data.[29] Findings are qualitatively robust to these alternative approaches. Table 14 shows that results become quantitatively stronger if considering only price distributions with at least 50

[28]The units of the coefficients are not comparable to Kelly's measure of skewness, so no statements can be made about the relative magnitude of the results in Table 11.

[29]The Retail Scanner datasets collects information on prices and quantities sold at the barcode level from a representative sample of US grocery chains. Details are discussed in Appendix B.1. Due to the large size of the Retail Scanner data I am restricted to using one wave of observations. Despite the large sample size of the Retail Panel, this limits the available variation as I am restricted to using information on demand composition from the COnsumer Panel wave 2019. I therefore rely on price distributions from the Consumer Panel waves 2007-2019 for the baseline results.

or 100 transactions. The robustness tests alleviate concerns that price distributions are constructed based on transaction data sampled from households. As the findings are robust to not weighting by quantities purchased, focusing on distributions with many transactions (where each posted price has a high chance of entering the sample), and are qualitatively similar when constructing price distributions based on data collected from stores instead of households, it is unlikely that households' purchase behavior is driving the results.

4 Calibration and Model Properties

4.1 Calibration

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 inequality, price dispersion, and macro aggregates.

Income Process. I assume an AR(1) with innovations from a Gaussian mixture, to capture higher moments of income risk. I target the cross-sectional variance of income, as well as moments of the distribution of income changes. Data targets are obtained from De Nardi et al. (2020). Details on the calibration are delegated to Appendix C.2.[30]

Functional Forms and External Parameters. I assume CRRA preferences for $u(\cdot)$ and a disutility of shopping effort as a function of $\mathcal{C}$ such that

$$u(\mathcal{C}) = \frac{\mathcal{C}^{1-\phi} - 1}{1-\phi} \quad \text{and} \quad v(s, \mathcal{C}) = \psi_1 \mathcal{C}^{\psi_2} \frac{s^2}{1-s}.$$

The term $\frac{s^2}{1-s}$ ensures that households will prefer an interior solution for s.[31] For $\psi_1 > 0$ and $\psi_2 > 0$ the disutility of effort increases in $\mathcal{C}$, capturing the need to shop more often for larger baskets. ψ_2 determines the economies of scale in shopping effort.

[30] I thank Gonzalo Paz-Pardo for making the specific moments available.

[31] It yields $v(0, \mathcal{C}) = 0$, $v(1, \mathcal{C}) = \infty$, $v_s(0, \mathcal{C}) = 0$, $v_s(1, \mathcal{C}) = \infty$. Under these assumptions households optimally choose $0 < s < 1$ iff $\frac{\partial p_j(s^i)}{\partial s^i} < 0$, facilitating the computational solution of the model.

The calibrated version of the model features three varieties (levels of quality), i.e. $J = 3$.[32] In line with the evidence on low-level elasticities of substitution sampled in Jaravel and Olivi (2021), I set $\sigma = 2$.[33] I normalize the marginal cost of the lowest quality to $\kappa_1 = 1$ and set $q_3 = 0$ to ensure that the grocery price index $p_G(\mathcal{C}, s)$ is increasing in $\mathcal{C}$. The CRRA parameter is set to $\phi = 2$ and the annual real interest rate to $r = 0.02$. Based on Broda and Parker (2014), I set $\alpha = 0.35$ to the share of non-durable and services consumption covered by the Nielsen dataset for which I can observe price differences. This is a conservative choice, as search frictions and price dispersion can be expected to matter beyond the products covered by Nielsen.

Internal Parameters. The remaining parameters to be calibrated are $(\psi_1, \psi_2, \kappa_2, \kappa_3, \sigma_1, \sigma_2, \sigma_3, \beta, q_1, q_2)$. As they do not influence allocations, I do not need to account for the fixed cost K_j in the calibration. I impose $q_1 = 2q_2 < 0$ and $2(\kappa_2 - \kappa_1) = \kappa_3 - \kappa_1$. This leaves eight parameters for which I target eight moments, divided into three groups.

At the aggregate level, I target a wealth-to-income ratio of 3. While all parameters can influence all moments, the one most closely linked to the wealth to income ratio is β. Furthermore, I compute total sales of retailers over the cost of goods sold as a measure of the average retail markup. I target a value of 1.39 computed based on the US Census' Annual Retail Trade Survey.[34] The aggregate markup is closely related to ψ_1, which governs average shopping effort and hence the average price elasticity of demand at the extensive margin.

A second set of moments targets price dispersion across varieties. These moments are closely related to the return to search. Capturing the right returns to search across varieties (and therefore across consumption baskets) is important for the correct identification of the elasticity of shopping to expenditures. Targets for price dispersion are computed from the Nielsen Consumer Panel based on the same definition of a price dis-

[32] $J = 3$ is the minimum number required to generate realistic separation in consumption baskets.

[33] I abstract from non-homotheticicities in price elasticities across varieties as Handbury (2021) and Faber and Fally (2022) report these to be of negligible importance for the Nielsen sample. This choice does not affect retailers' price posting decisions as they take total demand for their variety as given and only take the price elasticity across transactions into account.

[34] I use data for 2007-2019 and take sales divided by purchases net of the change in inventories for *food and beverage stores*, *health and personal care stores*, and *general merchandise stores* as the categories most closely reflecting the retailers covered in Nielsen. I weight markups across categories by total sales. The targeted value lies within the set of results reported for retail markups in the literature, ranging from 1.31 (Sangani 2023) to 1.45 (Hall 2018).

Table 3: Calibration Targets and Model Fit

Target	Data	Model	Source
basket overlap (Q1 vs. Q5)	63.28%	63.34%	Nielsen (2007-2019)
Δp across varieties (Q1 vs. Q5)	7.23%	7.23%	Nielsen (2007-2019)
Δp within varieties (Q1 vs. Q5)	1.99%	1.99%	Nielsen (2007-2019)
mean(CoV_j)	0.1920	0.1935	Nielsen (2007-2019)
$CoV_2 - CoV_1$	-0.0119	-0.0115	Nielsen (2007-2019)
$CoV_3 - CoV_1$	-0.0203	-0.0201	Nielsen (2007-2019)
average markup	1.39	1.39	ARTS (2007-2019)
wealth/income	3.00	3.00	

Note: Results of the internal calibration of $(\psi_1, \psi_2, \kappa_2, \kappa_3, \sigma_1, \sigma_2, \sigma_3, \beta, q_1, q_2)$.

tribution as in Section 3.2 (SMA-quarter). To account for differences in the average price across barcodes in the data, I focus on the coefficient of variation (CoV). I target the expenditure-weighted average CoV across all price distributions. In addition, I estimate equation (13) for the CoV, including on the right-hand side the quintiles of the expenditure distribution. I target the implied differences in the CoV across varieties based on the endogenous demand composition (spending shares across quintiles) in the model.[35] Targets for price dispersion interact most closely with the values for σ_j.[36]

The final set of moments contains targets on expenditure composition across households, again measured from the Nielsen data. This set of targets identifies the elasticities of consumption baskets and shopping effort to households' expenditure and with them the main mechanism of this paper. To discipline how consumption baskets change across households, I target the (dis)similarity in expenditure shares ω_j at the barcode level between the first and the fifth quintile of the expenditure distribution. For this purpose, I interpret the vector of expenditure shares for a group of households as a discrete distribution over the universe of available varieties (barcodes) and measure the similarity between two such distributions as the histogram overlap.[37] The barcode-level overlap between consumption baskets at the bottom and top of the expenditure distribution is 63%. In addition, I target the annual savings as a share of respective grocery expenditure of households at the bottom quintile of expenditures relative to the top quintile,

[35]Targeting the coefficient of variation is equivalent to normalizing all prices by the mean and computing the standard deviation of normalized prices as e.g. in Kaplan and Menzio (2015).

[36]Figure 1 shows that conditional on demand weighted search effort $\bar{s}_j$ the CoV directly identifies σ_j.

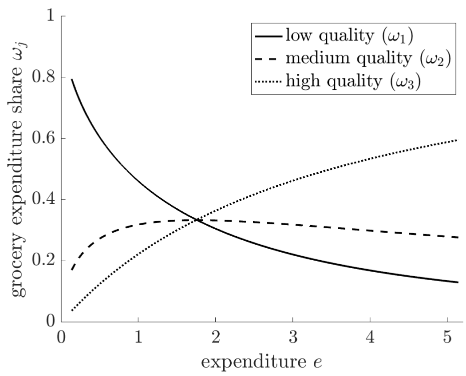
[37]Full details on the construction of this target are provided in Appendix B.5.

due to (i) buying similar varieties that are cheaper on average and (ii) paying less for identical varieties. Target values are given by the differences between the first and fifth quintile in Table 1. Given the return to search (matching the CoV), difference in prices paid for identical varieties identifies ψ_2, which governs the shape of households' shopping policy along the expenditure distribution. The overlap in consumption baskets and price differences across varieties interact closely with q_j and κ_j across varieties respectively.

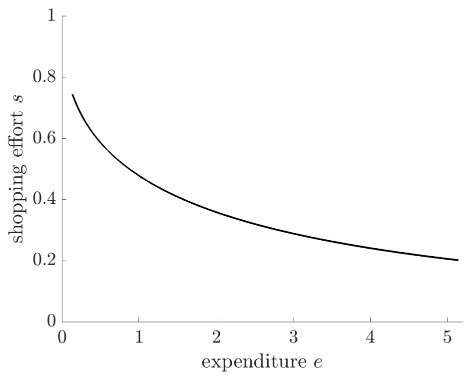
Table 3 summarizes all targets and shows that the model is able to match the moments considered. The calibration yields $q_j = \{-1.76,\ -0.88,\ 0\}$, $\psi_1 = 0.012$, $\psi_2 = 0.33$, $\kappa_j = \{1,\ 1.098,\ 1.196\}$, $\sigma_j = \{1.46,\ 1.495,\ 1.5213\}$, and $\beta = 0.933$. Two results are noteworthy: First, $\psi_2 < 1$ implies increasing returns to scale in shopping effort as the utility cost of exerting a given effort s increases less than one-for-one with the consumption aggregator $\mathcal{C}$. Second, the intensive margin price elasticities across transactions σ_j are very similar across varieties. This resonates with the evidence in Handbury (2021) and Faber and Fally (2022) that intensive margin price elasticities are similar across income groups.

4.2 Model Properties and Validation

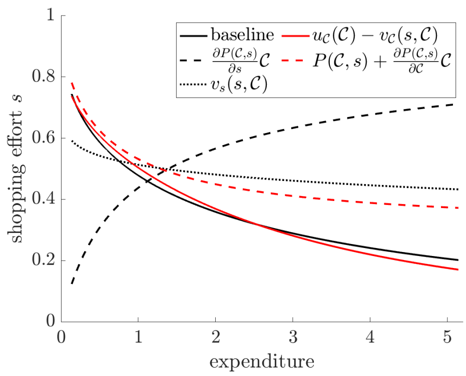
Policy Functions. Figure 2a reports that households with higher expenditures allocate a larger share of their consumption basket to varieties with higher quality (higher elasticity q_j) due to their non-homothetic preferences. Figure 2b shows that shopping effort is decreasing in households' total expenditure e. Figure 2c decomposes the differences in shopping effort across households into the main components of equation (4). The figure shows that the decline in shopping effort over expenditures is predominantly driven by the declining marginal utility of consumption ($u_{\mathcal{C}}(\mathcal{C}) - v_{\mathcal{C}}(s, \mathcal{C})$) and the increasing price of consumption ($P(\mathcal{C}, s) + \frac{\partial P(\mathcal{C},s)}{\partial \mathcal{C}}\mathcal{C}$) due to substitution towards (more costly) higher quality varieties, both reducing the marginal benefits of additional resources. The return to shopping alone ($\frac{\partial P(\mathcal{C},s)}{\partial s}\mathcal{C}$) would imply a shopping effort increasing in expenditures. The dependence of the disutility on the size of consumption bundles ($v_s(s, \mathcal{C})$) plays a quantitatively smaller role in explaining differences in shopping effort. This result suggests that alternative specifications of $v(s, \cdot)$, including e.g. employment status, that match the same targets on price differences across households and the distribution of expenditures would require a similarly small elasticity of the disutility of shopping and are unlikely to change results in a significant manner.



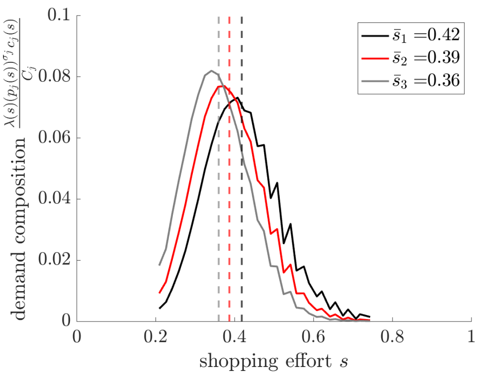
(a) Expenditure Shares



(b) Shopping Policy



(c) Shopping Policy (Drivers)



(d) Demand Composition

Figure 2: Consumption Baskets and Shopping Effort

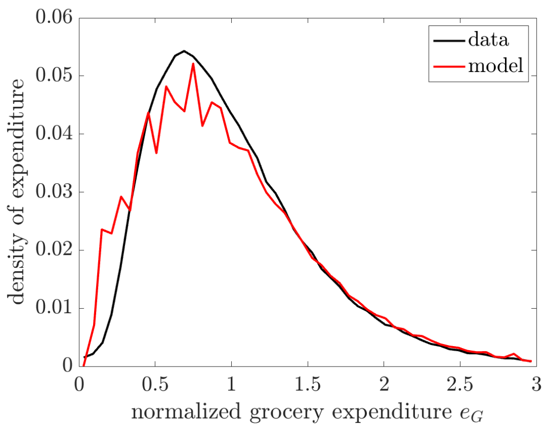
Note: Panels (a) and (b): Model implied expenditure shares ω_j across grocery varieties j and shopping effort s as a function of total household expenditure e. Panel (c): Decomposition of shopping effort based on equation (4). Counterfactual effort computed varying one term at a time while holding all other terms constant at their respective population average and solving $v_s(s, \mathcal{C})$ for the implied counterfactual s. Panel (d): Model implied distribution of demand weights in constructing $\bar{s}_j$ by grocery varieties j as a function of household shopping effort s.

Demand Composition. Figure 2d plots the composition of demand over households' shopping effort for each of the grocery varieties, interacting the policies for shopping and consumption baskets. The distribution of demand for higher quality varieties is shifted towards households with lower shopping effort, resulting in a lower demand weighted effort $\bar{s}_j$ for varieties with higher quality q_j.[38]

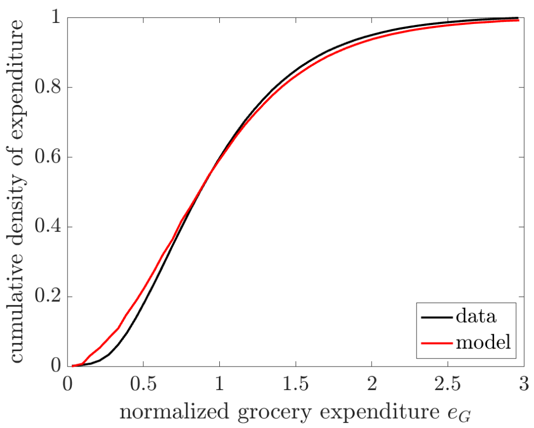
[38]Table 16 in Appendix C.3 shows that households in the bottom quintile of the expenditure distribution account for 11.32% of the spending on the low quality variety, they account for only 2.46% of spending on the high quality variety. Households in the top quintile of the expenditure distribution account for 51.26% of spending on the highest quality variety but only 26.34% on the lowest quality variety. Intuitively, each household is more important for the demand of varieties in her own consumption basket. Table 17 in Appendix C.3 shows this both for the model and the data. For a more formal discussion of how to measure this in the data refer to Appendix B.5.

Skewness. The skewness of price distributions decreases in the demand share of high spending households, qualitatively in line with the empirical results in Section 3.2. Based on the model implied demand composition across varieties, the coefficients of the first column of Table 2 would predict a -0.06 difference in Kelly's measure of skewness between varieties 3 and 1 vs. -0.03 in the model.[39] Hence, the differences in skewness generated by the model can account for 50% of the untargeted empirical results.[40]

Inequality. Shopping effort and the allocation of consumption across varieties are determined conditional on expenditure e. Matching the distribution of expenditures therefore ensures a realistic distribution of shopping and consumption policies in the model. Figure 3 plots the model implied distribution of grocery expenditures along with its equivalent from the Nielsen dataset. While the calibration targets include moments of the labor earnings process and match the empirical overlap in consumption baskets, as well as price differences within and across varieties, the dispersion in households' expenditure is not included in the calibration. The model does remarkably well in capturing the empirically observed distribution of grocery expenditures, establishing it as a suitable framework for studying the relationship between expenditure inequality and posted prices.[41]



(a) Density



(b) Cumulative Density

Figure 3: Distribution of Grocery Expenditures – Model vs. Data

Note: Fit of the model implied distribution of grocery expenditures $e_G = \alpha e$. Model and data distributions normalized to mean 1 and sorted into 50 equally spaced bins between 0 and the maximum expenditure level. Data from Nielsen Consumer Panel 2007-2019. Household-level grocery expenditures are equivalence scale adjusted and deflated with the Urban CPI to 2019. I winsorize the top 1% of data.

[39]The first column of Table 11 predicts a -0.55 difference between Pearson's measure of skewness of varieties 3 and 1 vs. -0.22 in the model.

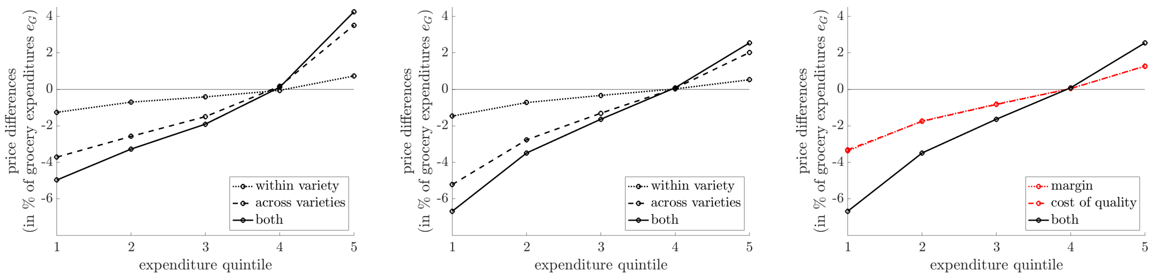
[40]Table 18 in Appendix C.3 provides additional information on the model implied price distributions.

[41]Table 19 in Appendix C.3 reports an equally good fit for the distribution of total disposable income.

5 Implications for Inequality

The calibrated model measures the contribution of heterogeneous shopping effort to inequality in two steps: First, I revisit the decomposition of price differences across households through the lens of the model and compute expenditure inequality under counterfactual shopping effort. Second, I solve for optimal consumption under counterfactual prices and study the welfare implications of heterogeneity in shopping effort.

Price Differences in the Model. I begin by revisiting the decomposition of households' grocery expenditures in equation (12) in the model. To compare the model to the data, I treat each variety j as a barcode and all three varieties as belonging to a single product bin k.[42] Figure 4 plots the decomposition of expenditures for both the data (4a) and the model (4b). Only the differences within and across varieties between the lowest and highest quintile are included in the set of targeted moments. The model does well at reproducing the levels and slope of both margins along the entire expenditure distribution.



(a) Data **(b)** Model **(c)** Model (Adjusted)

Figure 4: Price Differences along the Expenditure Distribution

Note: Price differences by expenditure quintile in the model and data, as a share of households' grocery spending. *Within variety*: difference between the price paid by a given household and the average price for a given product. *Across varieties*: difference in average prices across different products. In the *adjusted* decomposition, *margin* refers to margin differences within and across products (direct and equilibrium effect of shopping) and *cost of quality* to differences in marginal costs across products. Data from the Nielsen Consumer Panel 2007-2019, as reported in Table 1.

Price differences across households can be due to differences in marginal costs for the products bought or differences in margins paid. While a higher marginal cost cannot be avoided to purchase a preferred good, a higher margin implies that it would be feasible to

[42] To be consistent with the data, I compute average prices paid across transactions for each household instead of defining the barcode-level average price as $p_j(s)$.

achieve the same consumption allocation with lower expenditures. Higher margins reflect the cost imperfect competition in the product market – due to the presence of search frictions – imposes on households. The model can separate margins and marginal cost.

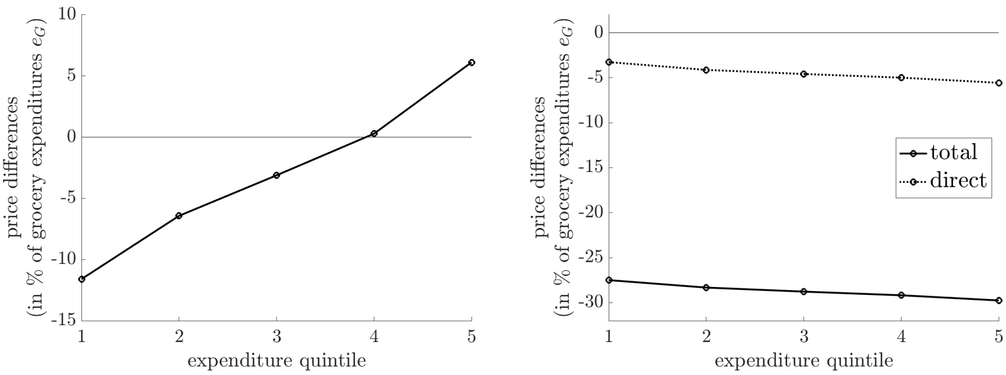
I adjust the decomposition in equation (12) to separate the effect of differences in margins and marginal cost. Differences in prices within varieties are entirely due to margin differences. They capture the direct effect of shopping effort, paying less for an identical product given the posted price distribution. Differences in average prices across varieties can be due to either marginal costs or margins. As shown in Section 2.3, average posted margins capture the equilibrium effect of search frictions on posted prices. This effect operates through retailers' response to demand-weighted shopping effort $\bar{s}_j$, which changes across varieties due to differences in demand composition. If all households where to observe two price draws for each transaction ($\bar{s}_j = 1 \ \forall j$), profit margins would be zero and any difference in price paid would be due to differences in marginal cost. I therefore attribute differences in average margins ($\hat{p}_{jk} - \kappa_{jk}$) across varieties to the equilibrium effect of shopping. This leaves differences in marginal cost of variety κ_j relative to the average marginal cost in bin k ($\tilde{\kappa}_k$) to explain the remaining price gap across households. In line with the interpretation of higher q_j as quality, I will refer to this term as *cost of quality*. I adjust the decomposition accordingly.[43]

$$e_i^G = \underbrace{\sum_k \sum_{j \in J_k} (p_{ijk} - \hat{p}_{jk}) c_{ijk} + ((\hat{p}_{jk} - \kappa_{jk}) - (\tilde{p}_k - \tilde{\kappa}_k)) c_{ijk}}_{\text{profit margins (direct \& equilibrium effect of shopping)}} + \underbrace{\sum_k \sum_{j \in J_k} (\kappa_{jk} - \tilde{\kappa}_k) c_{ijk}}_{\text{cost of quality}} + \sum_k \sum_{j \in J_k} \tilde{p}_k c_{ijk}$$

Figure 4c plots the adjusted decomposition in the model. Differences in margins account for a 4.6% difference in prices paid between the top and bottom quintile, about half of the effect of price differences for similar product on inequality in grocery spending. 2.6pp. of this effect is driven by the differences in profit margins across varieties, while the remaining effect is due to differences in profit margins within varieties. In this sense,

[43] Figure 15 in Appendix C.3 provides an alternative decomposition, assuming homogeneous shopping effort across households. Equalizing shopping effort across households under constant consumption bundles reduces inequality in spending by 4pp. 1.5pp. (38%) of which accounted for by the equilibrium effect. Difference to the decomposition in Figure 4 are accounted for by the non-linearities in $p_j(s)$ and remaining differences in profit margins due to the interaction of $\bar{s}_j$ with κ_j and σ_j.

accounting for the equilibrium effect of shopping on prices across varieties more than doubles the contribution of shopping effort to expenditure inequality.



(a) Price Discrimination **(b)** Costless Shopping

Figure 5: Counterfactual Shopping Effort and Prices

Note: Change in households' grocery spending under constant consumption bundles and alternative grocery prices. Baseline is the calibrated steady-state. Panel (a): Counterfactual assumes individually targeted price distributions for each variety, determined as the best response to individual search effort $\bar{s}_j = s(e)$. Panel (c): Costless search, i.e. $s(e) = \bar{s}_j = 1$. Counterfactuals adjust only individual effort given price distributions (direct) and collapse price distributions to κ_j (total) respectively.

Externalities. As retailers' price posting targets the average buyer in the market, each agents' effort affects the prices faced by all other households. Non-homotheticities and the partial separation of demand into different varieties reduce this externality. Nevertheless, a significant externality remains: Figure 5a shows that the bottom quintile would pay 12% less and the top quintile 6% more if households where to draw their prices from targeted price distributions (i.e. if $\bar{s}_j = s(e)$), while holding effort and consumption bundles constant.[44] The large size of the remaining externality can be accounted for by the higher spending levels of households at the top of the expenditure distribution, making them important for the demand of all varieties.[45]

Eliminating the Cost of Search. To measure the overall effect of shopping frictions, I eliminate the cost of search entirely and assume all households observe two prices per transaction ($s(e) = \bar{s}_j = 1$), again holding consumption bundles constant. I first adjust only individual effort (direct) before allowing posted prices to respond (total), moving the economy to Bertrand competition and collapsing all price distributions to κ_j. Bertrand

[44] Sangani (2023) finds similar magnitudes for the remaining externality despite a different approach.

[45] Appendix B.5 shows that this is in line with the data. Retailers do not significantly lose profits due to a lack of price discrimination. Holding households' policy functions constant, sales would increase by only 0.69% under perfect price discrimination. Lower profits on low-search households are almost perfectly offset by higher profits on high-search households.

competition reduces prices for all households, but more so for the top quintile, where households pay higher profit margins. Moving to Bertrand competition would decrease the cost of grocery consumption by 27.49% for the bottom quintile and 29.75% for the top quintile, increasing inequality in grocery spending by 2.26%.[46]

Effect on Grocery Consumption. To study the effect of price differences on grocery consumption, I fix the distribution of expenditures to the baseline economy and allow households to reoptimize their consumption bundles under alternative prices. Differences to the price changes reported above arise for three reasons: First, relative prices across varieties change and households can substitute towards a cheaper bundle. Second, as overall consumption $\mathcal{C}$ changes, non-homotheticicities affect the optimal bundle and the cost of additional consumption. Third, any counterfactual that eliminates price dispersion at the household level leads to consumption gains due to imperfect substitutability of purchases across transactions.[47] The first panel of Table 4 reports the implied change in grocery consumption by expenditure quintile for all counterfactuals introduced above. The difference between the response of the first and fifth quintile can be interpreted as the change in consumption inequality, measured as the ratio of top to bottom consumption.[48]

To measure the effect of shopping on inequality in grocery consumption, I rely on the decomposition of expenditures above. First, I eliminate price dispersion and let households pay the marginal cost for a given variety plus the average profit margin $(\kappa_j + (\tilde{p}_k - \tilde{\kappa}_k))$. Second, I let households pay the average price for a variety $(\hat{p}_{jk})$, isolating the direct effect of individual effort on prices paid. With no difference in profit margins, consumption inequality would increase by 3.74%. 1.84pp. are accounted for by margin differences within varieties (direct effect of shopping), while the rest is accounted for by margin differences across products (equilibrium effect of shopping). Again, the equilibrium effect doubles the contribution of shopping effort to inequality in consumption.

Allowing perfect price discrimination and letting all households draw prices from a targeted distribution would reduce inequality in grocery consumption by 13.42%. Removing the cost of search increases consumption along the entire distribution, but more so at the

[46] The difference to the decomposition above is accounted for by the fact that the no-cost counterfactual changes aggregate spending, while the baseline decomposition holds aggregate spending constant.

[47] This is the standard cost of price (markup) dispersion across varieties in a model with CES-demand.

[48] This holds as $log\left(\frac{x}{y}\right) = log(x) - log(y)$.

top where households pay higher profits margins. Allowing costless search would increase inequality in consumption by 8.31%. Only 2.39pp. (30%) of the increase in inequality is due to the direct effect of allowing households to observe two prices for all transactions, highlighting further the importance of the equilibrium response of posted prices when it comes to the effect of shopping on inequality.[49]

Table 4: Counterfactuals: Consumption and Welfare

	counterfactual		**quintile of expenditures**					**difference**
	p_j	$v(\cdot)$	Q1	Q2	Q3	Q4	Q5	**Q5-Q1**
grocery consumption c_G								
no margin difference	$\kappa_j + (\tilde{p}_k - \tilde{\kappa}_k)$	-	-1.40	-0.24	0.49	1.21	2.35	3.74
average price paid	$\hat{p}_{jk}$	-	-0.18	0.43	0.78	1.13	1.65	1.84
price discrimination	$\bar{s}j = s(e)$	-	8.81	4.91	2.40	-0.19	-4.61	-13.41
no cost of search	κ_j	-	24.66	27.14	28.76	30.38	32.97	8.31
no cost – direct effect	$\left(\mathbb{E}\left[p^{1-\sigma_j} \vert F_j^m(p)\right]\right)^{\frac{1}{1-\sigma_j}}$	-	2.30	3.07	3.53	3.98	4.69	2.39
total consumption $\mathcal{C}$								
no margin difference	$\kappa_j + (\tilde{p}_k - \tilde{\kappa}_k)$	-	-0.49	-0.08	0.17	0.42	0.81	1.31
average price paid	$\hat{p}_{jk}$	-	-0.06	0.15	0.27	0.39	0.58	0.64
price discrimination	$\bar{s}_j = s(e)$	-	2.99	1.69	0.83	-0.07	-1.64	-4.64
no cost of search	κ_j	-	8.02	8.77	9.25	9.73	10.49	2.47
no cost – direct effect	$\left(\mathbb{E}\left[p^{1-\sigma_j} \vert F_j^m(p)\right]\right)^{\frac{1}{1-\sigma_j}}$	-	0.80	1.06	1.22	1.38	1.62	0.82
welfare $\Delta(e)$ **(one period)**								
no margin difference	$\kappa_j + (\tilde{p}_k - \tilde{\kappa}_k)$	0	-0.15	0.27	0.52	0.76	1.12	1.27
average price paid	$\hat{p}_{jk}$	0	0.28	0.50	0.62	0.73	0.88	0.60
price discrimination	$\bar{s}_j = s(e)$	> 0	2.99	1.69	0.83	-0.07	-1.64	-4.64
no cost of search	κ_j	0	8.43	9.20	9.67	10.14	10.86	2.43
no cost – direct effect	$\left(\mathbb{E}\left[p^{1-\sigma_j} \vert F_j^m(p)\right]\right)^{\frac{1}{1-\sigma_j}}$	0	1.15	1.43	1.58	1.72	1.93	0.78
no disutility	$p_j(s(e))$	0	0.34	0.35	0.35	0.33	0.30	-0.05

Note: %-change in consumption and welfare under counterfactual prices and disutility of effort, holding the distribution of expenditures constant. p_j: Assumed price for variety j. $v(\cdot)$: Assumed disutility.

Welfare. To go from households' grocery consumption to welfare requires two steps. First, the second panel of Table 4 shows that the patterns reported for grocery consumption hold also for total consumption. The response of overall consumption is attenuated by about 1/3 as groceries account for 35% of households' overall expenditures. The assumption of frictionless markets for the outside good makes all results with respect to the effect of shopping on total consumption and welfare a lower bound. Second, studying the welfare implications of shopping effort requires to account for the disutility of shopping.

[49] The effect of costless search on inequality in consumption is larger than the gap in prices alone would suggest. This is due to a change in the optimal consumption bundle due to non-homotheticities and given the large decline in prices. The resulting increase in the cost of consumption is stronger for the bottom quintile as the highest-quality variety already accounts for a large share of the basket at the top. In addition, the top quintile benefits more from eliminating price dispersion across transactions.

I therefore compute the compensating percentage change $\Delta(e)$ in households' total consumption $\mathcal{C}(e)$ in the baseline steady state (SS) to make a household with expenditure e indifferent between one period under the baseline or with counterfactual prices (CF). $\Delta(e)$ is defined implicitly by

$$u\left((1+\Delta(e))\mathcal{C}^{SS}(e)\right) - v\left(s^{SS}(e), (1+\Delta(e))\mathcal{C}^{SS}(e)\right) = u\left(\mathcal{C}^{CF}(e)\right) - v\left(s^{CF}(e), \mathcal{C}^{CF}(e)\right) \quad (14)$$

The third panel of Table 4 reports the implied compensating change in consumption by expenditure quintile. It shows that the change in welfare is quantitatively almost identical to the change in total consumption. This result arises because the disutility of shopping effort is approximately flat along the expenditure distribution. I isolate this disutility cost by assuming a counterfactual with prices and consumption bundles as in the baseline economy but zero disutility, compensating households only for the cost of search. The disutility cost of shopping is about 0.34% for all quintiles. Along the distribution, the change in the disutility due to the decrease of shopping effort s in expenditure is offset by an increase in the size of the consumption bundle $\mathcal{C}$.[50]

Overall, the findings show the importance of the equilibrium effect of shopping on posted prices for its contribution to inequality in expenditures, consumption and welfare. The previous literature (Aguiar and Hurst 2007; Arslan et al. 2021; Pytka 2018, e.g.) has focused on how households can reduce the price they pay for a given variety. The findings above show that the effect of shopping goes further, due to the response of prices posted.

6 Implications for Average Prices and Markups

The relation between demand-weighted effort and posted prices makes the average price and markup in the economy a function of the distribution of expenditure. In this section, I exploit the dynamic structure of the model to study how aggregate shocks and policy changes shape this distribution and through it demand weighted effort and posted prices.

[50] As a baseline, I have computed welfare changes in a single period under constant expenditures. Table 20 in Appendix C.3 exploits the dynamic structure of the model and computes welfare over the infinite horizon, allowing households to change both their consumption bundles and spending-savings choices. The change in welfare is defined similar to equation (14) as the percentage change in lifetime consumption to make the household indifferent between living in the baseline or in the counterfactual economy. The results are qualitatively similar to the findings for single-period welfare but differences across households are quantitatively attenuated due to mean-reversion in states (z, a) over time.

6.1 Shopping and Prices over the Business Cycle

Empirical work on the cyclical properties of retail prices and markups in response to aggregate demand shocks remains conflicted, reporting acyclical price and markup responses to local unemployment shocks (Anderson et al. 2020; Coibion et al. 2015) but procyclical responses to changes in house prices (Stroebel and Vavra 2019). I revisit these findings in the model economy by focusing on the Great Recession period around 2008. The results in this section extend the work of Kaplan and Menzio (2016) on the role of shopping for aggregate fluctuations to rich household heterogeneity.[51] I show, that shifts in demand composition can account for the conflicting empirical evidence, making the cyclicality of prices and markups a function of the incidence of aggregate shocks.

The Great Recession saw both substantial earnings losses due to an increase in unemployment and losses in wealth in response to the decline in house prices. I simulate this shock as an unexpected one-time loss in households' net worth and persistent earnings losses differentiated by households' labor productivity state z. I choose an equal loss in wealth of 15% for all households, to match the decline in households' net worth between the last quarter of 2007 and first quarter of 2009 as reported in the US Financial Accounts (Table Z.1).[52] For losses in labor earnings along the income distribution, I take the estimates of Heathcote et al. (2020) for earnings changes by income in 2008-2010 for the first three periods after the shock and let earnings return to their steady state level by $t = 6$. The implied earnings losses are heavily concentrated among low-income households.[53] The decline in wealth amounts to roughly 46% of aggregate annual income in the model and the present discounted value of cumulative earnings losses to 22%.

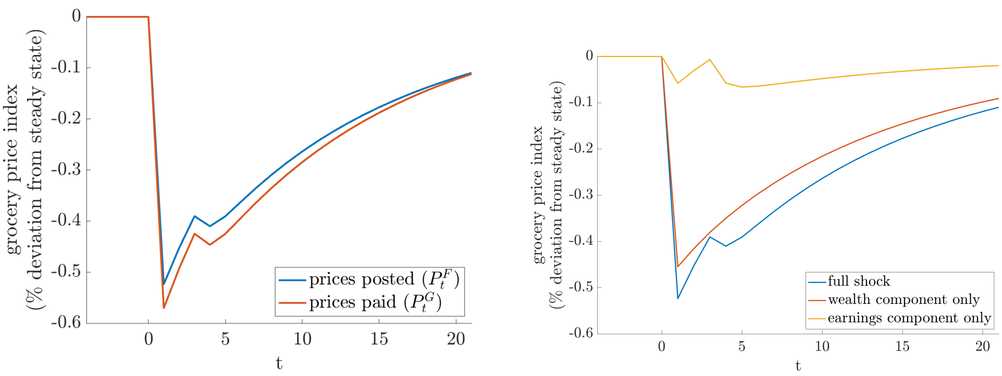
A Laspeyres index of average prices across grocery varieties, given by

$$P_t^l = \frac{\sum_{j=1}^{J} C_j^{SS} \mu_{jt}^l}{\sum_{j=1}^{J} C_j^{SS}},$$

[51] Kaplan and Menzio (2016) consider demand composition across two types of agents: Employed and unemployed. In order to sustain multiple equilibria and self fulfilling unemployment fluctuations, their framework requires a decline in aggregate income to be associated with an increasing role for (high-search-effort) unemployed households and a resulting decline in prices and profit margins for firms.

[52] This choice is in line with the decline in wealth growth for the top two quintiles of the wealth distribution (those households holding significant wealth) reported in Krueger et al. (2016).

[53] Table 21 in Appendix C.3 maps Heathcote et al. (2020) into labor productivity in the model.



(a) Prices Posted vs. Paid

(b) Response to Earnings vs. Wealth

Figure 6: Price Dynamics during the Great Recession

Note: Model implied response of an aggregate Laspeyres index $P_t^l = \frac{\sum_{j=1}^J C_j^{SS} \mu_{jt}^l}{\sum_{j=1}^J C_j^{SS}}$ to the Great Recession shock (15% loss in wealth and earnings losses from Heathcote et al. (2020)). Panel (a): Prices posted (P_t^F) and prices paid (P_t^G). Panel (b): Full response of posted prices decomposed into response to loss in earnings and loss in wealth.

measures the cyclicality of aggregate price changes. $C_j^{SS} = \int\int \lambda(e)c_j(p,e)dG_j(p|s(e))de$ is the total quantity of variety j consumed in steady state. The Laspeyres index abstracts from changes in baskets when aggregating across varieties and isolates prices changes. All parameters are kept at their steady-state values throughout, any response of prices is therefore driven by changes in households' (demand-weighted) shopping effort.

Figure 6a plots the response of the aggregate Laspeyres index of grocery prices, separately for prices posted P_t^F and prices paid P_t^G. Changes in households' shopping behavior reduced prices paid by about 0.6% during the Great Recession. This effect is predominantly driven by a reduction in posted prices by about 0.5pp. as retailers' respond to households' choices, making the equilibrium response of posted prices the dominant force behind the effect of shopping on fluctuations in paid retail prices during the Great Recession.[54]

Next, I show that the change in the index of posted prices is almost entirely driven by the decline in wealth. To isolate their respective effects, I simulate the loss in wealth and the loss in labor earnings separately. Figure 6b plots the response of the posted price index P_t^F for each of the two components and for the combined response. The combined

[54] For comparison of the magnitudes consider that the CPI for Food at Home, the subindex closest to the coverage of the Nielsen data, grew by 0.5pp. in 2009 relative to an average growth of 2.1pp. over the 2000-2021 period. Despite the absence of any adjustment cost, the dynamics of the model are not inconsistent with significant price stickiness. The price distribution for each variety on impact of the shock overlaps to about 96% with its steady-state counterpart. Therefore, the model can be consistent with up to 96% of retailers not adjusting their prices, as they are indifferent between any price on the support of the posted distribution. This is in line with the findings of Burdett and Menzio (2018).

response closely tracks the response to only the wealth component, while retailers barely react to the decline in earnings. With all parameters, including marginal cost κ_j, fixed at steady-state levels, these price responses are driven entirely by changes in posted markups. Hence, the model yields procyclical responses of retail prices and markups to the decline in wealth, but acyclical responses to the change in labor earnings during the Great Recession, reconciling the conflicting empirical evidence in Anderson et al. (2020), Coibion et al. (2015), and Stroebel and Vavra (2019).[55]

With constant parameters, posted prices are driven by demand-weighted shopping effort

$$\bar{s}_{jt} = \int\int \frac{\Lambda_t(z,a)\,[p_{jt}(s_t(z,a))]^{\sigma_j}\, c_{jt}(z,a)}{\bar{C}_{jt}}\; s_t(z,a)\; de.$$

Demand weighted effort responds to individual search effort or if shifts in demand composition change how retailers take the effort of different households into account. The first effect is captured by the response of $\bar{s}_{jt}$ to shopping policies $s_t(z,a)$. Shifts in demand composition arise due to changes in households' demand $[p_{jt}(s_t(z,a))]^{\sigma_j}\, c_{jt}(z,a)$ or the distribution of households across the state space $\Lambda_t(z,a)$. I consider the contribution of each of the three separately, fixing the others at steady state level.

Figure 7 plots the decomposition of price responses separately for the full shock and only the wealth and earnings component respectively. Changes in households' shopping policies alone reduce prices in response to both the wealth and earnings component, as each affected household increases her search effort to insure against an income loss. What accounts for the differences in cyclicality are differential responses of demand composition, driven by changes in households' consumption policies. This is due to the incidence of the shocks, earnings losses being concentrated among low-income (high-effort) households and high-income (low-effort) households facing a larger absolute decline in wealth.

The stronger low-income households are affected the more they have to reduce consumption and the lower becomes their share in demand. Retailers now face relatively more

[55] The model also resonates with the literature quantitatively. Figure 16 in Appendix C.3 shows that the decline in aggregate grocery prices (markups) paid in response to a 1% decline in households' wealth is about 0.035%. This elasticity is at the lower end but of similar magnitude as the range of estimates for the elasticity of retail prices to house prices reported in Stroebel and Vavra (2019), who find values of 0.02-0.2. As most household wealth is held in real estate and households' position in housing is levered due to mortgages, a 1% wealth shock is a lower bound for of a 1% decline in house values.

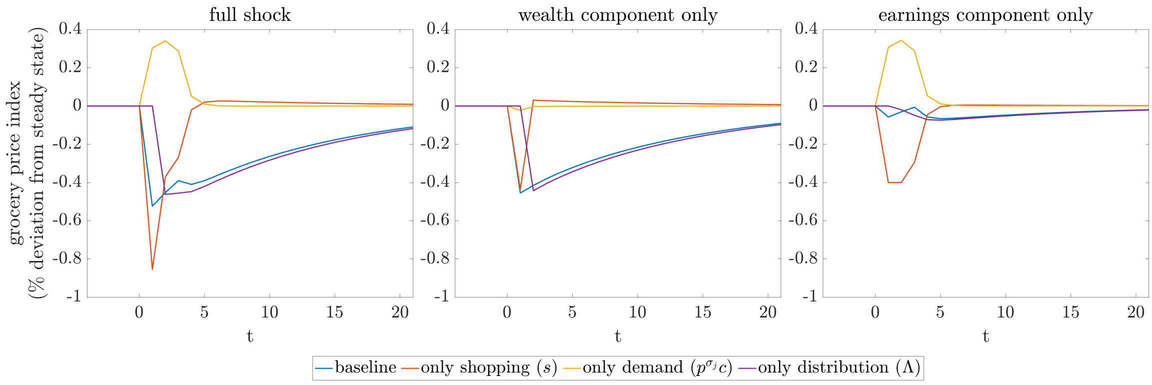


Figure 7: Decomposition of Posted Price Responses

Note: Response of an aggregate Laspeyres index of posted prices $P_t^F = \frac{\sum_{j=1}^J C_j^{SS} \mu_{jt}^F}{\sum_{j=1}^J C_j^{SS}}$ to the Great Recession shock (15% loss in wealth and earnings losses from Heathcote et al. (2020)). Panels report responses to the full shock, only the wealth, and only the earnings component. Complete response as a baseline, for each panel decomposed into the response to changes in individual demand, shopping policies, and the distribution of households, holding the respective others constant at steady state levels.

high-income buyers and respond to this shift in demand composition by attaching more weight to their (lower) shopping effort. In response, they increase prices. For the earnings component, this demand composition effect is strong enough to offset the direct increase in shopping effort. As high-income households are disproportionately affected by the wealth component, the effect of changes in households' consumption policies goes in the opposite direction and reallocates relative demand towards low-income (high-effort) households. In addition, high-income households significantly reduce their savings which increases their future shopping effort as they have become relatively poorer, captured by changes in the distribution $\Lambda_t(z, a)$. In combination of both components the change in households' consumption policies dampens the response of retail prices to the increase in individual shopping effort by about 50%.[56]

6.2 Demand Composition Across Policy Regimes

When posted prices become a function of the distribution of income and wealth in the economy, redistributive policies shaping these distributions affect households' cost of consumption. To show this, I introduce a flat tax on earnings and redistribute the

[56]Figures 17 and 18 in the Appendix provide additional evidence on the effect of demand composition on the cyclicality of retail prices by simulating 3% decline in aggregate income distributed differentially along the earnings distribution, showing that if losses are sufficiently concentrated among the bottom of the distribution retail prices can increase in response to a decline in aggregate earnings.

proceeds lump sum to all households. Their adjusted budget constraint is given by $e + a' \leq (1+r)a + (1-\tau)zw + T$, with transfer $T = \int\int \lambda(z,a)zw \; dzda$.[57]

I solve for the steady state of the model for given τ and compute the ensuing changes in households' earnings post taxes and transfers

$$\Delta\text{earn(z,a)} = \frac{(1-\tau)zw + T - zw}{zw} = \frac{T - \tau zw}{zw}$$

as well as changes in an individual Laspeyres index for total consumption $P^{lasp}(z,a)$

$$\Delta P^{lasp}(z,a) = \frac{\tilde{e}^{\tau}(z,a)}{e^{0}(z,a)} - 1 = \frac{c_O^0 + \sum_{j=1}^{J} p_j^{\tau}(s^0(z,a))c_j^0(z,a)}{c_O^0 + \sum_{j=1}^{J} p_j^{0}(s^0(z,a))c_j^0(z,a)} - 1$$

and grocery consumption $p_G^{lasp}(z,a)$

$$\Delta p_G^{lasp}(z,a) = \frac{\tilde{e}_G^{\tau}(z,a)}{e_G^{0}(z,a)} - 1 = \frac{\sum_{j=1}^{J} p_j^{\tau}(s^0(z,a))c_j^0(z,a)}{\sum_{j=1}^{J} p_j^{0}(s^0(z,a))c_j^0(z,a)} - 1.$$

where $p_j^{\tau}(s)$ is the price paid for grocery variety j by a household exerting effort s in an economy with redistributive tax τ. The index should be interpreted as the counterfactual expenditure level $\tilde{e}^{\tau}(z,a)$ a household in state (z,a) needs to buy the same basket as in the original steady state with the same shopping effort. Again, I focus on changes in a Laspeyres index and keep all policy functions at the original steady state with $\tau = 0$ to isolate changes in posted prices. For a fall in prices, the Laspeyres index provides a lower bound on the welfare impact as households can gain further by adjusting their choices.

The change in households' real income without re-optimizing policy functions is approximated by Δearn(z,a) $-$ $\Delta P^{lasp}(z,a)$. I aggregate changes in earnings and prices by expenditure quintile using the distribution of households in the original steady state. Table 5 presents results for a 5% earnings tax ($\tau = 0.05$). Overall, prices decline in response to the policy change as redistribution increases the share of relatively low-income (high-shopping-effort) households in demand, increasing demand-weighted search effort driving down posted prices. The effect is stronger for varieties with higher quality q_j, yielding larger declines in the price index of high-spending households.

[57] As the process for z is calibrated to households' earnings post taxes and transfers, the introduction of τ should be interpreted as additional redistribution relative to the current US system.

While the transfer dominates changes in real income at the bottom of the expenditure distribution, price changes are relatively more important at the top of the distribution and can compensate net contributors for a significant share of their earnings loss. Due to the assumption of a perfectly competitive outside good market, the reported changes in the aggregate price index P provides a lower bound while the results for p_G provide an upper bound if all consumption was subject to the same frictions. Table 5 shows that households at the top of the expenditure distribution are compensated for up to 11% of the loss in their post-tax earnings due to the response of posted prices.[58] Net contributors to redistribution can partially benefit from lower price levels as retailers place more weight on the higher effort of low-income households.

Table 5: Earnings and Price Changes under Redistributive Policies ($\tau = 0.05$)

		quintile of expenditures				
		Q1	Q2	Q3	Q4	Q5
income	Δearn	16.65%	5.22%	1.92%	-0.15%	-2.05%
prices	ΔP	-0.06%	-0.06%	-0.07%	-0.07%	-0.08%
	Δp_G	-0.16%	-0.18%	-0.20%	-0.21%	-0.23%
share ΔP	$\frac{-\Delta P}{\lvert\Delta \text{earn}\rvert}$	0.34%	1.24%	3.61%	48.47%	3.87%
share Δp_G	$\frac{-\Delta p_G}{\lvert\Delta \text{earn}\rvert}$	0.98%	3.54%	10.32%	138.49%	11.05%

Note: Average change in post-tax earnings (Δearn), grocery (Δp_G), and aggregate Laspeyres price index (ΔP) within each expenditure quintile in response to a 5% earnings tax and budget neutral transfer.

7 Conclusion

This paper develops an equilibrium theory of expenditure inequality and price dispersion, featuring incomplete markets, search for prices, non-homothetic preferences, and endogenous price distributions for multiple varieties. Average posted prices decline in the share of demand stemming from high-search-effort households. Theoretical predictions on the skewness of posted price distributions are in line with empirical evidence from the Nielsen Consumer Panel. The response of posted prices across varieties doubles the contribution of shopping effort to the difference between inequality in expenditure and consumption. After a shock similar to the Great Recession, posted prices respond to losses in wealth but not to losses in earnings, reconciling conflicting evidence on the cyclicality of retail

[58] While both income and price changes increase in the degree of redistribution (with higher values of τ), their relative contributions to the change in real income is similar for alternative values of τ.

markups. Endogenous price changes in response to redistributive taxes reduce the cost of contributing at the top by up to 11%. The results highlight the importance of equilibrium effects of shopping effort and demand composition when thinking about retail prices.

References

Aguiar, Mark and Mark Bils (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, and Arlene Wong (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.

Auer, Raphael, Ariel Burstein, Sarah M Lein, and Jonathan Vogel (2022). "Unequal Expenditure Switching: Evidence from Switzerland."

Bewley, Truman (1977). "The Permanent Income Hypothesis: A Theoretical Formulation." *Journal of Economic Theory* 16.2, pp. 252–292.

Bils, 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, 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.

Broda, Christian and Jonathan A Parker (2014). "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption." *Journal of Monetary Economics* 68, pp. 20–36.

Burdett, Kenneth and Kenneth L Judd (1983). “Equilibrium Price Dispersion.” *Econometrica* 51.4, 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* 39.2, pp. 257–273.

Butters, Gerard R. (1977). “Equilibrium Distributions of Sales and Advertising Prices.” *The Review of Economic Studies* 44.3, pp. 465–491.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong (2015). “The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications.” *American Economic Review* 105.3, pp. 993–1029.

Comin, Diego, Danial Lashkari, and Martí Mestieri (2021). “Structural Change With Long-Run Income and Price Effects.” *Econometrica* 89.1, pp. 311–374.

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.

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.”

Ferriere, Axelle, Philipp Grübener, and Dominik Sachs (2023). “Optimal Redistribution: Rising Inequality vs. Rising Living Standards.”

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.

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.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L Violante (2020). “The Rise of US Earnings Inequality: Does the Cycle drive the Trend?” *Review of Economic Dynamics* 37, S181–S204.

Huo, Zhen and José-Víctor Ríos-Rull (2015). “Tightening Financial Frictions on Households, Recessions, and Price Reallocations.” *Review of Economic Dynamics* 18.1, pp. 118–139.

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."

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.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri (2016). "Macroeconomics and Household Heterogeneity." *Handbook of Macroeconomics*. Vol. 2, pp. 843–921.

Menzio, Guido (2021). "Optimal Product Design: Implications for Competition and Growth under Declining Search Frictions."

Michelacci, Claudio, Luigi Paciello, and Andrea Pozzi (2022). "The Extensive Margin of Aggregate Consumption Demand." *The Review of Economic Studies* 89.2, pp. 909–947.

Mongey, Simon and Michael Waugh (2022). "Pricing Inequality."

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.

Olivi, Alan, Vincent Sterk, and Dajana Xhani (2023). "Optimal Monetary Policy during a Cost-of-Living Crisis."

Pisano, Luigi, Luca Riva, and Andrea Stella (2022). "Price Heterogeneity and Consumption Inequality."

Pytka, Krzysztof (2018). "Shopping Effort in Self-Insurance Economies."

Sangani, Kunal (2023). "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.

Waugh, Michael E (2023). "Heterogeneous Agent Trade."

ONLINE APPENDIX

(Intended for Online Publication Only)

Shopping, Demand Composition, and Equilibrium Prices

Lukas Nord – Minneapolis Fed[1]

A Extensions to the Retailer Problem

A.1 Retailer's Problem without an Intensive Margin

This section outlines retailers' price posting problem for a single variety j without an intensive margin of demand. As in Section 2.2, retailers take the distribution of households over expenditure $\lambda(e)$, their shopping effort $s(e)$ and their consumption $c_j(e)$ as given. I abstract from retailers' entry decision and fix the mass of active retailers to one.[2] Retailers face constant marginal cost κ_j, households make a unit continuum of transactions and observe two draws from the distribution of posted prices $F_j(p)$ for a share $s(e)$ of those transactions. Contrary to before, I shut down the intensive margin of demand. Households are assumed to purchase an equal amount of the product during each transaction, i.e. a type e household purchases $c_j(e)$ units in each transaction. Purchases with two simultaneous price observations are made at the lowest observed price as long as this price is below an exogenous willingness to pay $\bar{p}_j$. These assumptions are identical to those in Pytka (2018) and Sangani (2023).

[1]The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. 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 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.

[2]This is without loss of generality. Free entry with fixed cost of operating and an endogenous mass of retailers yields the identical solution for the distribution of posted prices.

Retailers' Problem. Retailers commit to a price before meeting any buyers. The total profits of a retailer posting price p are given by

$$\pi_j(p) = \left[\int \lambda(e) c_j(e) \left[(1 - s(e)) + s(e) 2(1 - F_j(p)) \right] de \right] (p - \kappa_j)$$
$$= C_j \left[(1 - \bar{s}_j) + \bar{s}_j 2(1 - F_j(p)) \right] (p - \kappa_j),$$

where $C_j = \int \lambda(e) c_j(e) de$ is total demand for variety j and $\bar{s}_j = \int \frac{\lambda(e) c_j(e)}{C_j} s(e) de$ demand weighted average shopping effort. With constant quantities purchased across transactions, retailers trade off profits per sale $(p - \kappa_j)$ only against the extensive margin of demand.

Posted Price Distribution. For given κ_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 posted prices $F_j(p)$ exists with compact support $[\underline{p}_j, \bar{p}_j]$, where

$$F_j(p) = \begin{cases} 0 & \text{if } p < \underline{p}_j \\ 1 - \frac{1-\bar{s}_j}{2\bar{s}_j} \frac{\bar{p}_j - p}{p - \kappa_j} & \text{if } p \in [\underline{p}_j, \bar{p}_j] \\ 1 & \text{if } p > \bar{p}_j \end{cases} \tag{15}$$

and

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

Retailers play a mixed strategy, randomizing prices over the interval $[\underline{p}_j, \bar{p}_j]$ according to the density $f_j(p)$ associated with $F_j(p)$. The distribution of posted prices depends on the marginal cost κ_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.

Moments of the Posted Price Distribution. Retailers' problem without an intensive margin of demand yields a fully analytical characterization of $F_j(p)$ and its support. This allows to solve for the moments of the distribution and their relation with $\bar{s}_j$ in closed form. Expressions for the mean, standard deviation, as well as Pearson's and Kelly's measure of skewness are presented in Proposition 2.

Proposition 2 *The mean μ_j^F, standard deviation σ_j^F, Pearson's moment coefficient of skewness γ_j^F, and Kelly's measure of skewness $\tilde{\gamma}_j^F$ of the posted price distribution $F_j(p)$ for given κ_j, $\bar{p}_j$, and $0 < \bar{s}_j < 1$ can be derived as*

(i)

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

(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)

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

(iv)

$$\tilde{\gamma}_j^F = 0.8\bar{s}_j$$

Proof. *Follows from equation* (15) *and the standard formulas for mean, standard deviation, Pearson's and Kelly's measure of skewness of any continuous distribution.* ■

Proposition 3 implies that the average price posted is increasing in marginal cost κ_j and maximum willingness to pay $\bar{p}_j$, but decreasing in demand-weighted shopping effort $\bar{s}_j$. The effect of $\bar{s}_j$ operates through a decrease in profit margins $\left(\frac{\partial(\mu_j^F - \kappa_j)}{\partial \bar{s}_j} < 0\right)$. An increase in $\bar{s}_j$ makes it more likely that the average buyer observes a second price, and hence tilts sellers' tradeoff between higher margins and retaining market share at the extensive margin in favor of the latter.

Proposition 3 *The mean of the posted price distribution μ_j^F is strictly increasing in marginal cost κ_j and maximum willingness to pay $\bar{p}_j$, but strictly decreasing in demand-weighted search effort $\bar{s}_j$ for $0 < \bar{s}_j < 1$, 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 from taking first derivatives of μ_j^F.* ■

Proposition 4 shows that both measures of skewness of the posted price distribution are a function only of demand-weighted shopping effort $\bar{s}_j$ and independent of parameters. Furthermore, they are strictly increasing in $\bar{s}_j$.

Proposition 4 *Pearson's and Kelly's measure of skewness of the posted price distribution (γ_j^F and $\tilde{\gamma}_j^F$) are strictly increasing in demand-weighted search effort $\bar{s}_j$ for $0 < \bar{s}_j < 1$, but independent of marginal cost κ_j and maximum willingness to pay $\bar{p}_j$, i.e.*

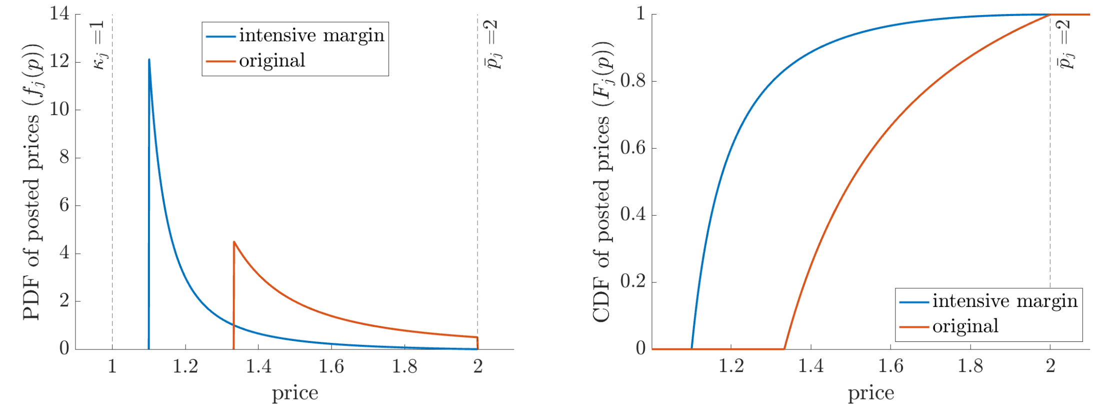
$$(i)\ \frac{\partial \gamma_j^F}{\partial \kappa_j} = \frac{\partial \tilde{\gamma}_j^F}{\partial \kappa_j} = 0, \quad (ii)\ \frac{\partial \gamma_j^F}{\partial \bar{p}_j} = \frac{\partial \tilde{\gamma}_j^F}{\partial \bar{p}_j} = 0, \quad (iii)\ \frac{\partial \gamma_j^F}{\partial \bar{s}_j} > 0 \text{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

Proof. *Follows from taking first derivatives of γ_j^F and $\tilde{\gamma}_j^F$.* ■

The analytical results confirm the main mechanism of how search effort affects retailers' price posting decisions in the simplified model without an intensive margin of demand.

Price Distribution With and Without an Intensive Margin. It is natural to a compare price dispersion in the original Burdett-Judd model without an intensive margin of demand to the setup with an intensive margin across transactions as outlined in Section 2.2. The original Burdett-Judd framework does not feature an intensive margin price elasticity but requires to parameterize the upper bound of the price distribution $\bar{p}_j$ directly as a maximum willingness to pay. I assume $\sigma_j = 2$, identical marginal cost $\kappa_j = 1$ and demand weighted search effort $\bar{s}_j = 0.5$ and set $\bar{p}_j = \frac{\sigma_j}{\sigma_j - 1}\kappa_j = 2$ exogenously for the model without an intensive margin to obtain the same upper bound of the support of $F_j(p)$ and make both models comparable.

Figure 8 compares the price distributions of the original BJ model and the version with an intensive margin of demand as outlined above. Everything else equal, the presence of an intensive margin incentives retailers to post relatively lower prices as compared to the original Burdett-Judd economy with only an extensive margin of demand. The intensive margin increases the overall price elasticity of demand beyond undercutting competitors for households with a second price observation, making it optimal for retailers to post lower markups as compared to the economy without an intensive margin.



(a) PDF **(b)** CDF

Figure 8: Comparison with Original Burdett-Judd

Note: Model solution with and without an intensive margin of demand for a calibration with $\bar{p}_j = \frac{\sigma_j}{\sigma_j - 1}\kappa_j = 2$, $\underline{\kappa}_j = 1$, $\bar{s}_j = 0.5$.

A.2 Heterogeneous Marginal Cost

Take the setup from Section A.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 retailers for which the support has to end at the maximum willingness to pay. This assumption imposes no restriction on the solution 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 κ of posting price p for variety j are given by

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

To solve for the equilibrium distribution of posted prices $F_j(p)$ I follow closely the steps of Burdett and Mortensen (1998) and Mortensen (2003) for an equivalent model of wage posting. As in the baseline model, $F_j(p)$ has no mass points, has a connected support, and the upper bound of the support of $F_j(p)$ is $\bar{p}_j$. $p(\kappa)$, the set of prices maximizing $\pi_j(p, \kappa)$ for given $F_j(p)$, is unique, single valued, and maps the distribution of marginal costs into posted prices such that $F_j(p(\kappa)) = \Gamma_j(\kappa)$.

Uniform Marginal Costs. An analytical characterization of moments for the case with heterogeneous marginal cost is beyond the scope of this paper. I simulate the model for 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

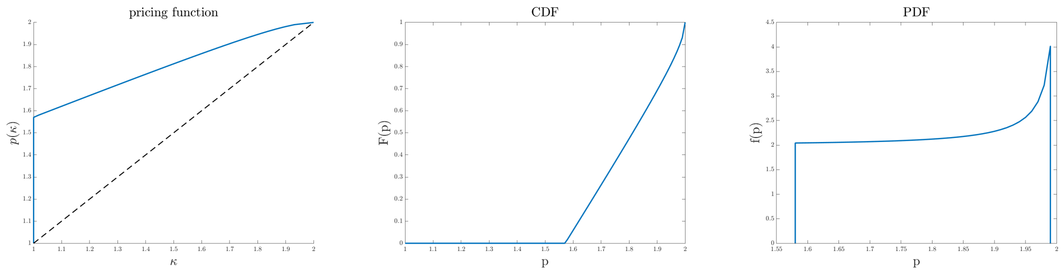


Figure 9: Uniform Distribution - Example

Note: Model solution for a uniform distribution of marginal costs, $\bar{p} = 2$, $\underline{\kappa} = 1$, $\bar{s} = 0.75$.

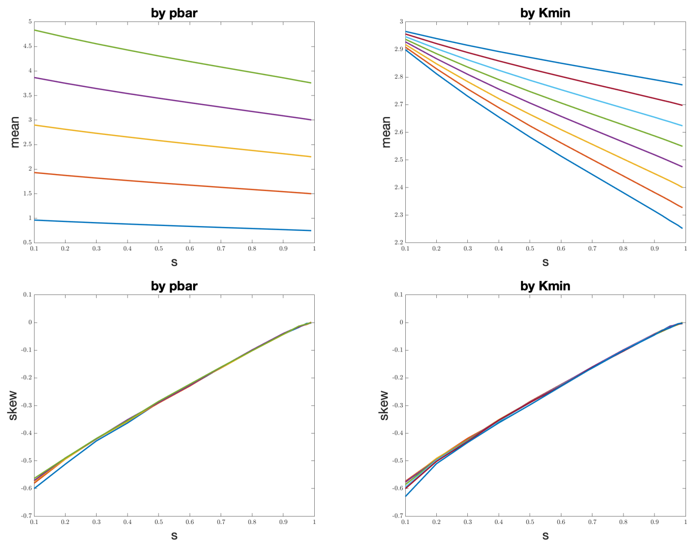


Figure 10: Uniform Distribution - Simulations

Note: Mean and Pearson's moment coefficient of skewness of simulated price distributions over $\bar{s} \in [0, 1]$ for a uniform distribution by $\bar{p}$ and $\underline{\kappa}$.

$\bar{\kappa}_j = \bar{p}_j$. I take $\bar{p}_j = 3$ and $\kappa_{min} = 0$ as the baseline and change one parameter at a time, simulating 1,000,000 price draws for each combination of parameters and computing the mean and skewness of the posted price distribution. To highlight the properties of a solution to the model with heterogeneous κ, Figure 9 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 10 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 considered, 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 calibrations 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. 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.

B Empirical Appendix

B.1 The Nielsen Dataset

The main empirical results presented in this paper 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 grocery 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. The dataset is constructed as an annual panel and the median household remains in the sample for about 3 consecutive waves. Nielsen applies several quality checks such as minimum reporting requirements to the sample before making data available. 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 35% of spending excluding durable goods.[3]

Prices and quantities are reported at the barcode level. Nielsen organizes all barcodes into 10 departments (e.g. dry groceries or fresh foods), which are further divided into 125 product groups (e.g. snacks vs. pasta within dry groceries), and 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.

[3] For further details on the dataset and its application in Macroeconomic research see e.g. Argente and Lee (2021), Kaplan and Menzio (2015), Pisano et al. (2022), Broda and Parker (2014) or Michelacci et al. (2022).

Data is provided in annual waves and I use the 2007-2019 waves of the Consumer Panel. 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. 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.

No data on wealth is available in the Nielsen panel and income data is only available as categorical variable and reported as the tax base for the previous calendar year, i.e. refer to households' taxable income two years prior to the sample. In contrast, expenditures on the consumption categories covered in Nielsen are well measured. 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 adjusted to 2019 USD using the CPI for all Urban Consumers.

In addition to the Consumer Panel, Nielsen makes a Retail Panel available. The Retail panel contains information on quantities sold and average prices charged at the store-week-barcode level and is constructed based on data from cash registers. The sample contains up to 50,000 participating grocery, drug, mass merchandiser, and other stores per year, covering about one out of six stores in the relevant categories in the US. Due to the considerable amount of data contained in each wave, I am limited to using the 2019 wave which I employ to conduct robustness exercises for some of the key results obtained from the Consumer Panel.

B.2 Local vs. National Average Prices

For the baseline analysis in Section 3.1 I measure the prices households pay relative to the national, annual average price across all households. The literature often defines price distributions and the relative price a household pays for a given barcode more narrowly, i.e. over Scantrack Market regions and by quarter (see e.g. Kaplan and Menzio 2015). This way of measuring price differences within barcodes can be subject to a small sample bias. In this appendix, I define the bias formally and show that it can be alleviated by focusing on products with sufficiently many transactions to compute average prices. I further show that this comes at the cost of introducing a bias in price differences across products. I use the results to discuss robustness of the main specification based on national average prices.

To measure differences in prices paid, the literature generally benchmarks household i's average price p_{ij} for barcode j against the average price paid $\bar{p}_j$ for the barcode across all households. This measure is subject to a potential downward bias if household i accounts for a large share of transactions of barcode j. More formally, the price p_{ij} is defined as the quantity-weighted average over all transactions T_i of household i

$$p_{ij} = \frac{\sum_{\tau=1}^{T_i} p_{\tau ij} q_{\tau ij}}{\sum_{\tau=1}^{T_i} q_{\tau ij}},$$

where $p_{\tau ij}$ and $q_{\tau ij}$ are respectively the price paid and quantity purchased of barcode j by household i in transaction τ. The average price $\bar{p}_j$ is defined accordingly as

$$\bar{p}_j = \frac{\sum_i \sum_{\tau=1}^{T_i} p_{\tau ij} q_{\tau ij}}{\sum_i \sum_{\tau=1}^{T_i} q_{\tau ij}}.$$

One can rewrite the average price paid as

$$\bar{p}_j = \nu_{ij} p_{ij} + (1 - \nu_{ij}) p_{-ij},$$

where $\nu_{ij} = \frac{\sum_{\tau=1}^{T_i} q_{\tau ij}}{\sum_i \sum_{\tau=1}^{T_i} q_{\tau ij}}$ is household i's share in demand for variety j and

$$p_{-ij} = \frac{\sum_h \sum_{\tau=1}^{T_h} p_{\tau hj} q_{\tau hj} - \sum_{\tau=1}^{T_i} p_{\tau ij} q_{\tau ij}}{\sum_h \sum_{\tau=1}^{T_h} q_{\tau hj} - \sum_{\tau=1}^{T_i} q_{\tau ij}}$$

is the average price paid by all households except household i. The difference of household i's price relative to the average yields

$$\Delta p_{ij} = p_{ij} - \bar{p}_j = (1 - \nu_{ij})(p_{ij} - p_{-ij}).$$

While p_{-ij} is an unbiased measure of the true average price paid, the price difference will be biased towards zero by a factor $(1 - \nu_{ij})$, i.e. will be biased more the larger the demand share of household i for good j.

A way to alleviate the bias is to increase the number of transactions considered to compute $\bar{p}_j$, thereby decreasing ν_{ij}. This can be done by either computing the average price for barcode j at the national, annual level or defining it at the local, quarterly level but only considering transactions for barcodes with a minimum number of transactions in the region and quarter. Alternatively, one could also drop a household's own transactions when computing the average price. However, as for many goods there are only few households consuming it in a narrow region this increases the noise in average prices and often effectively implies dropping the good if a household accounts for a significant share of local purchases of this barcode.

Table 6 reproduces the results on price differences within barcodes, based on the decomposition in section 3.1, for local average prices and barcodes with a minimum of 1, 25, 50 and 100 of local transactions respectively. Results are qualitatively similar across all definitions of the average price and become quantitatively similar (or even stronger) to the baseline of national average prices the larger the minimum number of transactions.

Table 7 shows that concentrating on barcodes with a high number of transactions comes at the cost of a bias in estimates for price differences across products. When defining both the average price paid for a barcode and across barcodes within a product module at the quarter and Scantrack Market level results are quantitatively similar to the baseline when including all products. Reducing the set of considered products to those with a higher number of transactions yields qualitatively similar results but reduces the magnitudes of all coefficients. This is a mechanical result when reducing the set of products considered. It is amplified by potential selection of products that are popular across all households (for which households' consumption baskets overlap significantly) into having a higher number of transactions.

Considering the tradeoff between introducing a bias into estimates of price differences within vs. across barcodes by defining average prices at the local level, I apply national average prices for all baseline results. This choice is justified as the least biased local specifications (imposing restrictions on the minimum number of transactions for price differences within barcodes vs. not restricting the set of products for price differences across barcodes) yield qualitatively and quantitatively similar results to the national baseline.

Table 6: Price Differences Within Barcodes (Local Average Prices)

	within barcodes (national) ($N^{min}=1$)	within barcodes (local) ($N^{min}=1$)	within barcodes (local) ($N^{min}=25$)	within barcodes (local) ($N^{min}=50$)	within barcodes (local) ($N^{min}=100$)
	(1)	(2)	(3)	(4)	(5)
log(expenditure)	0.961***	0.476***	0.925***	1.104***	1.325***
	(0.069)	(0.044)	(0.077)	(0.094)	(0.124)
income 30k-60k	0.059	0.008	0.006	−0.034	0.005
	(0.050)	(0.026)	(0.065)	(0.088)	(0.142)
income 60k-100k	0.186***	0.065**	0.169*	0.144	0.187
	(0.063)	(0.033)	(0.086)	(0.125)	(0.178)
income >100k	0.363***	0.145***	0.237**	0.209	0.102
	(0.080)	(0.041)	(0.106)	(0.149)	(0.222)
1 non-employed household head	−0.284***	−0.110***	−0.207***	−0.262***	−0.417***
	(0.041)	(0.021)	(0.054)	(0.076)	(0.118)
2 non-employed household heads	−0.456***	−0.212***	−0.370***	−0.335***	−0.754***
	(0.072)	(0.040)	(0.094)	(0.128)	(0.184)
head's age 25-65	0.023	0.025	0.121	0.229**	0.412**
	(0.054)	(0.029)	(0.074)	(0.106)	(0.166)
sqrt(HH size)	0.544***	0.282***	0.575***	0.658***	0.914***
	(0.073)	(0.040)	(0.093)	(0.124)	(0.185)
FE year-state	X	X	X	X	X
FE household	X	X	X	X	X
Observations	801,398	801,398	800,229	797,418	778,492

Note: Regression of price differences within barcode as a share of expenditures. Column (1) average price defined based on national annual average price. Column (2) average price defined based on local quarterly average price. Column (3) average price defined based on local quarterly average price, restricted to products with at least $N = 25$ local quarterly observations. Column (4) average price defined based on local quarterly average price, restricted to products with at least $N = 50$ local quarterly observations. Column (5) average price defined based on local quarterly average price, restricted to products with at least $N = 100$ local quarterly observations. Data obtained from Nielsen Consumer Panel waves 2007-2019. Observation weighted with Nielsen provided sample weights. Standard errors clustered at the household level. *p<0.1; **p<0.05; ***p<0.01.

Table 7: Price Differences Across Barcodes (Local Average Prices)

	across barcodes (national) ($N^{min} = 1$)	across barcodes (local) ($N^{min} = 1$)	across barcodes (local) ($N^{min} = 25$)	across barcodes (local) ($N^{min} = 50$)	across barcodes (local) ($N^{min} = 100$)
	(1)	(2)	(3)	(4)	(5)
log(expenditure)	3.426***	3.674***	1.109***	0.564***	0.424***
	(0.147)	(0.125)	(0.085)	(0.089)	(0.099)
income 30k-60k	0.544***	0.416***	0.211**	0.288***	0.098
	(0.100)	(0.094)	(0.084)	(0.096)	(0.103)
income 60k-100k	1.048***	1.051***	0.436***	0.373***	0.027
	(0.127)	(0.119)	(0.109)	(0.122)	(0.135)
income >100k	1.451***	1.849***	0.737***	0.367**	0.096
	(0.160)	(0.150)	(0.140)	(0.151)	(0.175)
1 non-employed household head	−0.665***	−0.678***	−0.270***	−0.050	−0.072
	(0.083)	(0.077)	(0.070)	(0.084)	(0.095)
2 non-employed household heads	−1.471***	−1.613***	−0.462***	−0.239*	−0.214
	(0.160)	(0.147)	(0.119)	(0.132)	(0.149)
head's age 25-65	0.155	0.030	−0.034	0.008	−0.042
	(0.113)	(0.107)	(0.101)	(0.109)	(0.121)
sqrt(HH size)	0.645***	1.104***	−0.239**	−0.390***	−0.389***
	(0.148)	(0.140)	(0.120)	(0.130)	(0.143)
FE year-state	X	X	X	X	X
FE household	X	X	X	X	X
Observations	801,398	801,398	800,229	797,418	778,492

Note: Regression of price differences across barcodes as a share of expenditures. Column (1) average price defined based on national annual average price. Column (2) average price defined based on local quarterly average price. Column (3) average price defined based on local quarterly average price, restricted to products with at least $N = 25$ local quarterly observations. Column (4) average price defined based on local quarterly average price, restricted to products with at least $N = 50$ local quarterly observations. Column (5) average price defined based on local quarterly average price, restricted to products with at least $N = 100$ local quarterly observations. Data obtained from Nielsen Consumer Panel waves 2007-2019. Observation weighted with Nielsen provided sample weights. Standard errors clustered at the household level. *p<0.1; **p<0.05; ***p<0.01.

B.3 Price Differences and Shopping

Table 8: Prices and Shopping by Expenditure Quintile

		expenditure quintiles				
		1	2	3	4	5
	within	-1.256	-0.698	-0.411	-0.054	0.735
		(0.051)	(0.039)	(0.034)	(0.032)	(0.030)
by	across	-3.712	-2.574	-1.495	0.167	3.517
barcode		(0.094)	(0.070)	(0.069)	(0.062)	(0.059)
	both	-4.968	-3.272	-1.906	0.113	4.253
		(0.112)	(0.086)	(0.081)	(0.075)	(0.072)
	within	-2.421	-1.521	-0.889	-0.088	1.733
		(0.080)	(0.055)	(0.050)	(0.050)	(0.047)
by	across	-2.547	-1.751	-1.016	0.202	2.520
brand		(0.079)	(0.056)	(0.053)	(0.053)	(0.053)
	both	-4.968	-3.272	-1.906	0.113	4.253
		(0.115)	(0.086)	(0.079)	(0.074)	(0.071)
trips per purchase		0.189	0.145	0.136	0.131	0.130
		(0.001)	(0.0004)	(0.0003)	(0.0003)	(0.0003)

Note: First two panels show results of the decomposition in (12) defining a product at the barcode level and at the brand-module level respectively. Results reported in percent of households' total grocery spending and averaged by quintile. Third panel reports trips per purchase as a measure of shopping effort. Trips per purchases are defined as the number of store-day pairs with transactions divided by the number of store-day-barcode pairs with transactions. Households are sorted into quintiles within the year of observation based on annual spending divided by the square root of household size, applying sampling weights provided by Nielsen. Sample restricted to waves 2007-2019 and households with heads age 25-65. Aggregated with Nielsen sampling weights within waves, weighting each wave equally. Standard errors in brakets bootstrapped with 1000 replications.

Table 9: Price and Shopping by Household Demographics (full results)

	within products	across products	trips per purchase	price index
	(1)	(2)	(3)	(4)
log(expenditure)	0.961***	3.426***	−0.042***	0.706***
	(0.069)	(0.147)	(0.001)	(0.078)
income 30k-60k	0.059	0.544***	−0.001*	0.080*
	(0.050)	(0.100)	(0.001)	(0.046)
income 60k-100k	0.186***	1.048***	−0.002**	0.178***
	(0.063)	(0.127)	(0.001)	(0.057)
income >100k	0.363***	1.451***	−0.002**	0.326***
	(0.080)	(0.160)	(0.001)	(0.070)
1 non-employed household head	−0.284***	−0.665***	0.002***	−0.236***
	(0.041)	(0.083)	(0.000)	(0.037)
2 non-employed household heads	−0.456***	−1.471***	0.004***	−0.422***
	(0.072)	(0.160)	(0.001)	(0.068)
head's age 25-65	0.023	0.155	−0.001	−0.013
	(0.054)	(0.113)	(0.001)	(0.050)
sqrt(HH size)	0.544***	0.645***	−0.032***	0.399***
	(0.073)	(0.148)	(0.001)	(0.066)
mean			0.15	
FE year-state	X	X	X	X
FE household	X	X	X	X
Observations	801,398	801,398	801,398	801,398

Note: Regression of the contribution of differences between prices paid and the average price (1) within product or (2) across products to expenditure inequality on household characteristics. Contributions defined as a share of households' grocery expenditures. The price within products is computed by barcode and the price across products is computed across barcodes within a product module by unit of measurement. Column (3) number of annual shopping trips (stores visited) divided by number of annual purchases (transactions for a barcode-store-day pair). Column (4) price index defined as households grocery spending divided by spending at average price of a barcode. Data from the Nielsen Consumer Panel waves 2007-2019. Standard errors clustered at the household level. *p<0.1; **p<0.05; ***p<0.01.

Table 10: Price and Shopping by Household Demographics (spending quintiles)

	within products	across products	trips per purchase	price index
	(1)	(2)	(3)	(4)
spending quintile 2	0.369***	1.188***	−0.020***	0.249***
	(0.046)	(0.084)	(0.000)	(0.039)
spending quintile 3	0.644***	2.062***	−0.028***	0.477***
	(0.057)	(0.098)	(0.001)	(0.044)
spending quintile 4	0.919***	3.048***	−0.034***	0.738***
	(0.058)	(0.109)	(0.001)	(0.048)
spending quintile 5	1.274***	4.493***	−0.039***	1.059***
	(0.067)	(0.127)	(0.001)	(0.056)
income 30k-60k	0.062	0.554***	−0.001***	0.079*
	(0.050)	(0.100)	(0.001)	(0.046)
income 60k-100k	0.190***	1.061***	−0.002***	0.176***
	(0.063)	(0.127)	(0.001)	(0.057)
income >100k	0.374***	1.488***	−0.003***	0.328***
	(0.080)	(0.159)	(0.001)	(0.070)
1 non-employed household head	−0.279***	−0.646***	0.001***	−0.235***
	(0.041)	(0.083)	(0.000)	(0.037)
2 non-employed household heads	−0.443***	−1.471***	0.002***	−0.418***
	(0.072)	(0.159)	(0.001)	(0.067)
head's age 25-65	0.016	0.129	−0.000	−0.015
	(0.053)	(0.113)	(0.001)	(0.050)
sqrt(HH size)	0.439***	0.256*	−0.022***	0.359***
	(0.067)	(0.138)	(0.001)	(0.066)
mean			0.15	
FE year-state	X	X	X	X
FE household	X	X	X	X
Observations	801,398	801,398	801,398	801,398

Note: Regression of the contribution of differences between prices paid and the average price (1) within product or (2) across products to expenditure inequality on household characteristics. Contributions defined as a share of households' grocery expenditures. The price within products is computed by barcode and the price across products is computed across barcodes within a product module by unit of measurement. Column (3) number of annual shopping trips (stores visited) divided by number of annual purchases (transactions for a barcode-store-day pair). Column (4) price index defined as households grocery spending divided by spending at average price of a barcode. Data from the Nielsen Consumer Panel waves 2007-2019. Standard errors clustered at the household level. *p<0.1; **p<0.05; ***p<0.01.

B.4 Evidence on Demand Composition and Price Distributions

Table 11: Demand Composition and the Skewness of Price Distributions (Pearson)

	by expenditures		by shopping			by income		by employment
	all	working age	prices	trips		working age		all
	(1)	(2)	(3)	(4)		(5)		(6)
quintile 2	-1.642^{***}	-1.483^{***}	-0.084	0.856^{***}	income	-0.131	1 non-employed	0.850^{***}
	(0.198)	(0.182)	(0.152)	(0.132)	30k-60k	(0.114)	household head	(0.088)
quintile 3	-2.309^{***}	-2.083^{***}	-1.471^{***}	1.380^{***}	income	-0.805^{***}	2 non-employed	0.994^{***}
	(0.190)	(0.173)	(0.128)	(0.141)	60k-100k	(0.116)	household heads	(0.152)
quintile 4	-3.062^{***}	-2.587^{***}	-2.733^{***}	1.910^{***}	income	-0.805^{***}		
	(0.168)	(0.158)	(0.118)	(0.127)	>100k	(0.084)		
quintile 5	-3.414^{***}	-3.014^{***}	-2.083^{***}	1.856^{***}				
	(0.152)	(0.145)	(0.098)	(0.113)				
FE product module	X	X	X	X		X		X
FE quarter-SMA	X	X	X	X		X		X
Observations	3,026,551	3,026,404	3,026,551	3,026,551		3,026,404		3,026,551

Note: Regression of Pearson' measure of skewness on demand shares by household groups. Price distributions defined as all transactions of a barcode within a Scantrack Market Region and quarter. Demand shares defined as the share of national annual spending on a barcode by each group of households. Data from Nielsen Consumer Panel waves 2007-2019. Observations weighted by total sales in given price distribution. Standard errors clustered at the barcode-year level. *p<0.1; **p<0.05; ***p<0.01.

Table 12: Demand Composition and the Skewness of Price Distributions (UPC FE)

	by expenditures		by shopping			by income		by employment
	all	working age	prices	trips		working age		all
	(1)	(2)	(3)	(4)		(5)		(6)
quintile 2	-0.037	-0.023	-0.029	0.049^{**}	income	-0.005	1 non-employed	-0.035^{*}
	(0.035)	(0.030)	(0.026)	(0.022)	30k-60k	(0.022)	household head	(0.018)
quintile 3	-0.073^{**}	-0.063^{**}	-0.108^{***}	0.071^{***}	income	0.012	2 non-employed	-0.000
	(0.034)	(0.030)	(0.027)	(0.024)	60k-100k	(0.023)	household heads	(0.032)
quintile 4	-0.088^{***}	-0.073^{**}	-0.136^{***}	0.138^{***}	income	-0.022		
	(0.032)	(0.028)	(0.027)	(0.026)	>100k	(0.022)		
quintile 5	-0.137^{***}	-0.103^{***}	-0.158^{***}	0.133^{***}				
	(0.032)	(0.028)	(0.025)	(0.027)				
FE barcode	X	X	X	X		X		X
FE quarter-SMA	X	X	X	X		X		X
Observations	3,026,551	3,026,404	3,026,551	3,026,551		3,026,404		3,026,551

Note: Regression of Kelly's measure of skewness on demand shares by household groups. Price distributions defined as all transactions of a barcode within a Scantrack Market Region and quarter. Demand shares defined as the share of national annual spending on a barcode by each group of households. Data from Nielsen Consumer Panel waves 2007-2019. Observations weighted by total sales in given price distribution. Standard errors clustered at the barcode-year level. *p<0.1; **p<0.05; ***p<0.01.

Table 13: Demand Composition and the Skewness of Price Distributions (Robustness)

	baseline	unweighted regression	unweighted skewness	only HH weights skewness	retail panel
	(1)	(2)	(3)	(4)	(5)
expenditure quintile 2	-0.125^{***}	-0.115^{***}	-0.012	-0.071^{**}	-0.013
	(0.035)	(0.022)	(0.036)	(0.035)	(0.021)
expenditure quintile 3	-0.195^{***}	-0.172^{***}	-0.087^{***}	-0.146^{***}	-0.015
	(0.032)	(0.020)	(0.033)	(0.033)	(0.020)
expenditure quintile 4	-0.244^{***}	-0.235^{***}	-0.117^{***}	-0.173^{***}	-0.031^{*}
	(0.029)	(0.019)	(0.031)	(0.030)	(0.018)
expenditure quintile 5	-0.330^{***}	-0.338^{***}	-0.213^{***}	-0.280^{***}	-0.095^{***}
	(0.027)	(0.017)	(0.028)	(0.028)	(0.018)
FE module	X	X	X	X	X
FE quarter-SMC	X	X	X	X	X
Observations	3,026,551	3,026,551	3,026,551	3,026,551	44,587,153

Note: Regression of the skewness of price distributions on demand shares by expenditure quintile. Price distributions defined as all transactions of a barcode within a Scantrack Market Region and quarter. Demand shares defined as the share of national annual spending on a barcode by quintile. Column (1): Baseline result, observations weighted within distribution by household weights and quantities purchased and across distributions by total spending on given price distribution. Column (2): No weighting of price distributions in regressions. Column (3): No weighting of price observations within distributions. Column (4): Price observations within distributions weighted by household weights but not quantities. Column (5): Baseline specification but price distributions constructed from Nielsen Retail Panel wave 2019, demand shares from Consumer Panel wave 2019. Data for columns (1)-(4) from Nielsen Consumer Panel waves 2007-2019. Standard errors clustered at the barcode-year level. *p<0.1; **p<0.05; ***p<0.01.

Table 14: Demand Composition and Price Distributions (Number of Transactions)

	Nmin = 25 (1)	Nmin = 50 (2)	Nmin = 100 (3)
expenditure quintile 2	-0.125^{***} (0.035)	-0.148^{**} (0.068)	-0.022 (0.133)
expenditure quintile 3	-0.195^{***} (0.032)	-0.206^{***} (0.062)	-0.274^{**} (0.121)
expenditure quintile 4	-0.244^{***} (0.029)	-0.229^{***} (0.056)	-0.263^{**} (0.107)
expenditure quintile 5	-0.330^{***} (0.027)	-0.315^{***} (0.052)	-0.287^{***} (0.100)
FE module	X	X	X
FE quarter-SMC	X	X	X
Observations	3,026,551	803,604	202,067

Note: Regression of the skewness of price distributions on demand shares by expenditure quintile. Price distributions defined as all transactions of a barcode within a Scantrack Market Region and quarter. Demand shares defined as the share of national annual spending on a barcode by quintile. Column (1): Only price distributions with at least $N = 25$ transactions. Column (2): Only price distributions with at least $N = 50$ transactions. Column (3): Only price distributions with at least $N = 100$ transactions. Data obtained from Nielsen Consumer Panel waves 2007-2019. Observations weighted by total expenditures included on given price distribution. Standard errors clustered at the barcode-year level. *p<0.1; **p<0.05; ***p<0.01.

B.5 Consumption Baskets and Separation in the Goods Market

Quantifying non-homotheticities in the data requires a measure for the similarity of consumption baskets. Define the consumption basket of any group g of households i via the share of their annual total expenditures allocated to each good ω_j^g. The expenditure share of good j for group g in a given year is given as

$$\omega_j^g = \frac{\sum\limits_{i \in g} e_j^i}{\sum\limits_{j \in J} \sum\limits_{i \in g} e_j^i}.$$

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

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

Note that under homothetic preferences and the law of one price $\omega_j^g = \omega_j^h \; \forall j, g, h$ and hence $\Omega^{gh} = 1$, so any deviation of the overlap from one can be interpreted as a deviation from these assumptions. Conducting the analysis by groups of households accounts for variation in taste within groups and computing statistics at the annual frequency averages out seasonal fluctuations.

Figure 11 reports the histogram overlap between the first and fifth quintile of the distribution of annual expenditures, defining a good 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 barcode level) the overlap decreases substantially to 63%. I.e. consumption baskets of high and low expenditure households exhibit a significant mismatch driven by variation in purchases of closely substitutable goods within Nielsen-defined product modules. For the empirical decomposition in Section 3.1 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. 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.

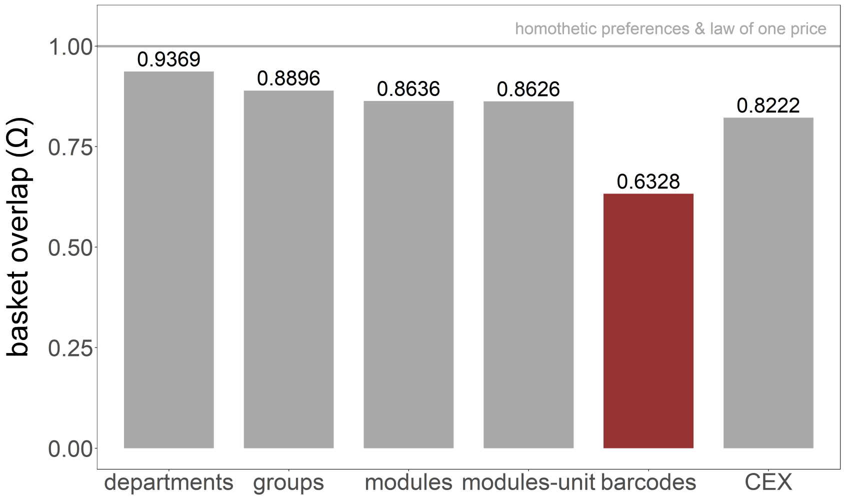


Figure 11: Consumption Basket Overlap - Top vs. Bottom Expenditure Quintile
Note: Histogram overlap in the vector of expenditure shares for the bottom and top quintile of expenditures, by different definitions of a product. First five columns derived from Nielsen Consumer Panel, final column from Consumer Expenditure Survey (CEX). CEX column considers 14 spending categories.

To complement the analysis based on Nielsen data for even broader categories of consumption goods, the final bar in Figure 11 produces the overlap between the bottom and top quintile of the income distribution in the Consumer Expenditure Survey (CEX) defining goods at the 14 most aggregated categories.[4] The non-homotheticity in CEX categories is roughly at the level of Nielsen defined product modules, while the barcode level overlap measured in Nielsen is approximately 25% lower.

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

[4] I use aggregated series for consumption by category and income quintile reported by the Bureau of Labor Statistics (BLS). The 14 expenditure categories considered include: *food at home, food away from home, alcoholic beverages, housing, apparel and services, transportation, healthcare, entertainment personal care products and services, reading, education, tobacco products and smoking supplies, miscellaneous expenditures, cash contributions, personal insurance and pensions.*

is for a given good j, we define the *demand share* (DS) of group h for good j as

$$DS_j^h = \frac{\sum\limits_{i \in h} e_j^i}{\sum\limits_{g \in G} \sum\limits_{i \in g} e_j^i}.$$

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

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

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

Figure 12 plots the cross market shares by quintile of the expenditure distribution at the barcode 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.

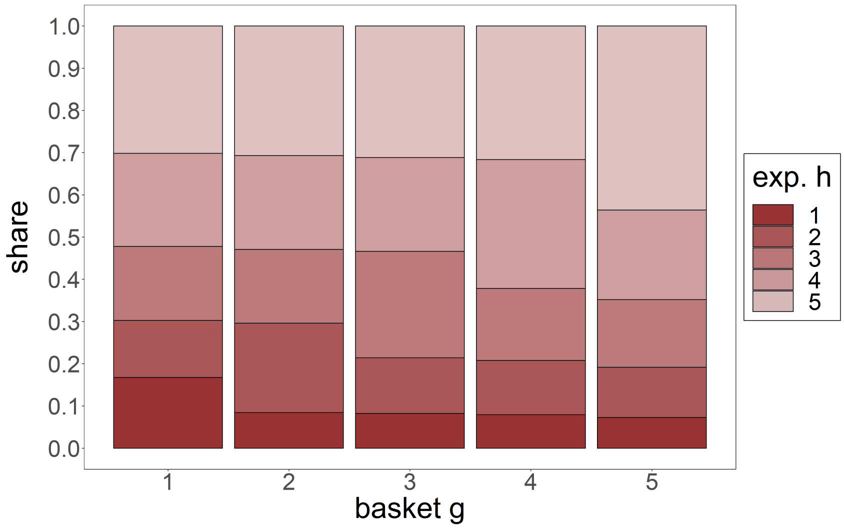


Figure 12: Cross Market Shares

Note: Barcode-level cross market shares of expenditure quintile h for the basket of quintile g. Cross market shares are constructed weighting the share of demand for a product j coming from quintile h by the expenditure share ω_j^g of product j in the basket of quintile g. Data from Nielsen Consumer Panel.

C Model Appendix

C.1 Definition of Equilibrium

A stationary equilibrium in the economy consists of households' value function $V(z,a)$, consumption policy functions $\{c_O(e), \{c_j(e), c_{j\tau}(p,e)\}_{j=1}^J\}$, shopping policy $s(e)$, expenditure policy $e(z,a)$ and savings policy $a'(z,a)$, the induced distributions of households across states $\Lambda(z,a)$ and expenditure $\lambda(e)$, aggregated demand $\{\bar{C}_j\}_{j=1}^J$ and demand-weighted shopping effort $\{\bar{s}_j\}_{j=1}^J$ for each variety, posted price distributions $\{F_j(p)\}_{j=1}^J$ and implied pricing functions $\{p_j(s)\}_{j=1}^J$, where

(i) Given $\{p_j(s)\}_{j=1}^J$, households' value and policy functions solve (1), (2), and (5).

(ii) The distribution of households is a stationary solution to the law of motion

$$\Lambda(z',a') = \int\int \Lambda(z,a) Pr(z'|z) \mathbb{1}_{a'=a(z,a)} \, dz \, da.$$

(iii) The distribution of expenditures satisfies

$$\lambda(e) = \int\int \Lambda(z,a) \mathbb{1}_{e=e(z,a)} \, dz \, da.$$

(iv) Aggregated demand for variety j is given by

$$\bar{C}_j = \int \lambda(e) \left[p_j(s(e))\right]^{\sigma_j} c_j(e) \, de.$$

(v) Demand weighted shopping effort for variety j is given by

$$\bar{s}_j = \int \frac{\lambda(e) \left[p_j(s(e))\right]^{\sigma_j} c_j(e)}{\bar{C}_j} s(e) \, de.$$

(vi) Given $\{\bar{s}_j\}_{j=1}^J$, the posted price distributions $\{F_j(p)\}_{j=1}^J$ solve (9).

(vii) Given $\{F_j(p)\}_{j=1}^J$, the pricing functions $\{p_j(s)\}_{j=1}^J$ satisfy (6).

C.2 Income Process

For households' idiosyncratic labor productivity z, I assume an AR(1) process with innovations from a Gaussian mixture, formally defined as

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

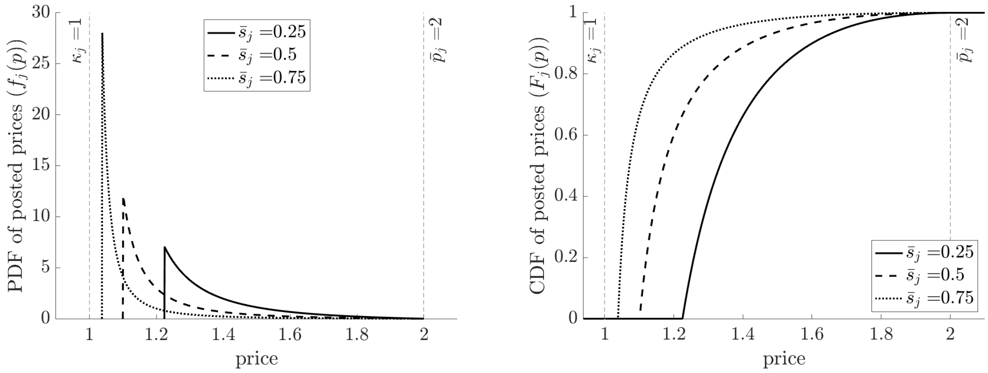
I discretize the process with 16 states for z following the method of Farmer and Toda (2017). The income process requires calibrating 6 parameters $(\rho, p, \mu_{\varepsilon,1}, \sigma^2_{\varepsilon,1}, \mu_{\varepsilon,2}, \sigma^2_{\varepsilon,2})$. I impose $\mu_{\varepsilon,2} = -\frac{\chi}{1-\chi}\mu_{\varepsilon,1}$ 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). For more information on how the target values are constructed see their Appendix A.3. All targets are reported in Table 15 along with the model counterparts. The associated parameter values are $\rho = 0.91$, $\sigma_1 = 0.59$, $\sigma_2 = 0.23$, $\chi = 0.082$, and $\mu_1 = -0.57$.

Table 15: Calibration Targets – 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

Note: Results of the calibration of an AR(1) income process with Gaussian-mixture shocks. The process is discretized with 16 states following Farmer and Toda (2017). Data moments for the PSID obtained from De Nardi et al. (2020).

C.3 Additional Model Results



(a) Density **(b)** Cumulative Density

Figure 13: Posted Price Distributions

Note: Price distributions derived from retailers' optimal price posting problem for marginal cost $\kappa_j = 1$, intensive margin elasticity $\sigma_j = 2$, and three levels of demand-weighted shopping effort $\bar{s}_j$.

Table 16: Demand Composition of Grocery Varieties

	quintile of expenditures Q1	Q2	Q3	Q4	Q5
low quality (q_1)	0.1132	0.1761	0.2101	0.2373	0.2634
medium quality (q_2)	0.0549	0.1234	0.1834	0.2537	0.3846
high quality (q_3)	0.0246	0.0774	0.1431	0.2424	0.5126

Note: Model implied demand shares by varieties of the grocery good and expenditure quintile. Demand shares are computed as total sales of variety j to quintile g divided by total sales of variety j.

Table 17: Cross Market Shares – Model vs. Data

		model market share of exp. quintile					**data** market share of exp. quintile				
		Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
	Q1	0.09	0.15	0.19	0.24	0.33	0.17	0.14	0.18	0.22	0.30
by basket	Q2	0.08	0.14	0.19	0.24	0.35	0.08	0.21	0.17	0.22	0.31
of expenditure	Q3	0.07	0.13	0.18	0.24	0.37	0.08	0.13	0.25	0.22	0.31
quintile	Q4	0.06	0.13	0.18	0.24	0.39	0.08	0.13	0.17	0.30	0.32
	Q5	0.06	0.12	0.17	0.24	0.41	0.07	0.12	0.16	0.21	0.44

Note: Cross market shares are computed as the market share of a quintile for each variety averaged by the expenditure shares in the consumption basket of another quintile. Data obtained from Nielsen Consumer Panel waves 2007-2019, consumption baskets in the data defined at the barcode level.

Table 18: Price Distributions of Grocery Varieties

		quality of grocery variety		
		low (q_1)	medium (q_2)	high (q_3)
demand-weighted shopping effort	$\bar{s}_j$	0.42	0.39	0.36
average price posted	μ_j^F	1.48	1.64	1.80
average price paid	μ_j^G	1.38	1.53	1.68
coefficient of variation	CoV_j^F	0.20	0.19	0.18
return to search (at $s = 0.5$)	$\frac{\partial p_j(0.5)}{\partial s}$	-0.12	-0.13	-0.14
Kelly's skewness of posted prices	$\tilde{\gamma}_j^F$	0.75	0.74	0.72
Pearson's skewness of posted prices	γ_j^F	1.98	1.86	1.77

Note: Summary statistics of model implied price distributions for each grocery variety j.

Table 19: Income Distribution – Model vs. Data

	quintile of post-tax income				
	Q1	Q2	Q3	Q4	Q5
Model	5.01%	10.49%	15.78%	23.89%	44.82%
Data	6.44%	10.91%	14.70%	20.31%	47.65%

Note: Fit of the model implied income distribution. In the model, income is measured as labor and financial income ($zw + ra$). Data moments for household income after taxes and transfers from Congressional Budget Office (CBO) for 2007-2018.

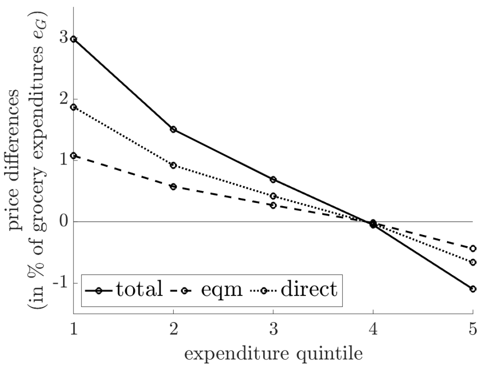


Figure 14

Figure 15: Homogeneous Shopping Effort and Prices

Note: Change in households' grocery spending under constant consumption bundles and flat shopping effort $\hat{s}$ across households. $\hat{s}$ defined by holding consumption bundles and aggregate expenditures constant and assuming $s(e) = \bar{s}_j = \hat{s}\ \forall e, j$. Counterfactuals adjust only individual effort $s(e)$ (direct), only aggregate effort $\bar{s}_j$ and posted prices (eqm), or both jointly (total). Baseline is the calibrated steady-state.

Table 20: Counterfactuals: Welfare (Dynamic)

	counterfactual		**quintile of expenditures**					**difference**
	p_j	$v(\cdot)$	Q1	Q2	Q3	Q4	Q5	**Q5-Q1**
no margin difference	$\kappa_j + (\tilde{p}_k - \tilde{\kappa}_k)$	0	-0.05	0.20	0.36	0.52	0.78	0.83
average price paid	$\hat{p}_{jk}$	0	0.33	0.46	0.54	0.61	0.73	0.40
price discrimination	$\bar{s}_j = s(e)$	> 0	2.63	1.83	1.30	0.73	-0.31	-2.95
no cost of search	κ_j	0	8.62	9.08	9.38	9.68	10.19	1.57
no cost – direct effect	$\mathbb{E}\left[p^{1-\sigma_j} \mid F_j^m(p)\right]$	0	1.21	1.38	1.47	1.57	1.72	0.52
no disutility	$p_j(s(e))$	0	0.34	0.35	0.34	0.34	0.32	-0.02

Note: Change in infinite horizon welfare under counterfactual prices and disutility of effort, allowing households to re-optimize their consumption bundles and spending-savings decisions. p_j: Assumed price faced by households for variety j. $v(\cdot)$: Assumed disutility of shopping effort.

Table 21: Earnings Losses for Great Recession Shock

		$z_1 - z_6$	z_7	z_8	z_9	z_{10}	z_{11}	$z_{12} - z_{16}$
cumulative share of households		24%	34%	47%	60%	74%	84%	100%
linked percentile in Heathcote et al. (2020)		$P20$	$P30$	$P40$	$\frac{P50+P60}{2}$	$P70$	$P80$	$P90$
$\frac{z_t}{z^{SS}} - 1$	$t = 1$ (2008)	-17.3%	-7.4%	-3.0%	-4.3%	-3%	-3%	-0.9%
	$t = 2$ (2009)	-43.4%	-16.8%	-13.3%	-6.8%	-6.6%	-2.9%	-2.9%
	$t = 3$ (2010)	-55.6%	-23.7%	-15.1%	-8.5%	-6.4%	-4.6%	-2.5%

Note: Calibration of earnings losses by productivity state in the Great Recession. Data moments obtained from Heathcote et al. (2020).

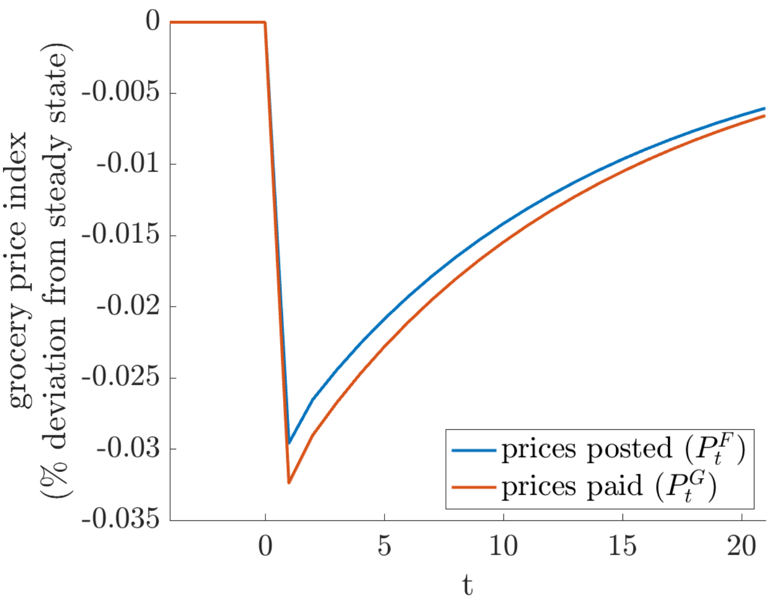


Figure 16: Prices Posted and Prices Paid in Response to a 1% Loss in Wealth

Note: Model implied response of an aggregate Laspeyres index $P_t^l = \frac{\sum_{j=1}^{J} C_j^{SS} \mu_{jt}^l}{\sum_{j=1}^{J} C_j^{SS}}$ of prices posted (P_t^F) and prices paid (P_t^G) to a proportionate 1% decrease in beginning of period assets a for each household.

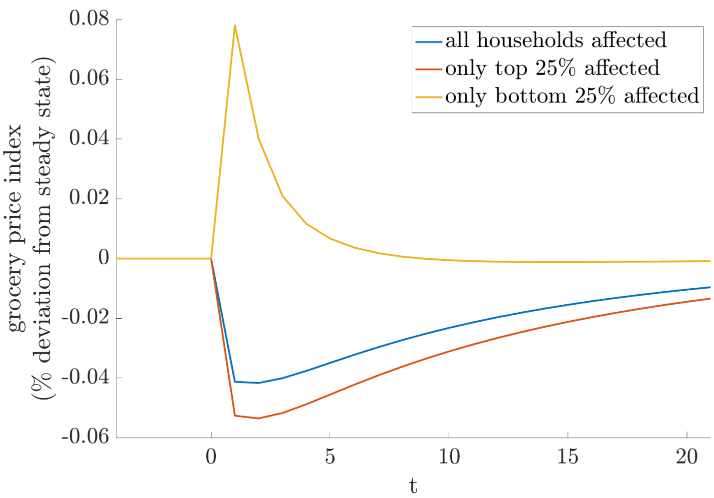


Figure 17: Aggregate Prices under Varying Incidence

Note: Response of an aggregate Laspeyres index of posted prices $P_t^F = \frac{\sum_{j=1}^{J} C_j^{SS} \mu_{jt}^F}{\sum_{j=1}^{J} C_j^{SS}}$ to a 3% initial loss in aggregate labor earnings relative to the steady state with persistence 0.5, affecting (i) all households proportionately to their labor earnings, (ii) only the bottom quartile of labor earnings (proportionately to their earnings), or (iii) only the top quartile (proportionately to their earnings).

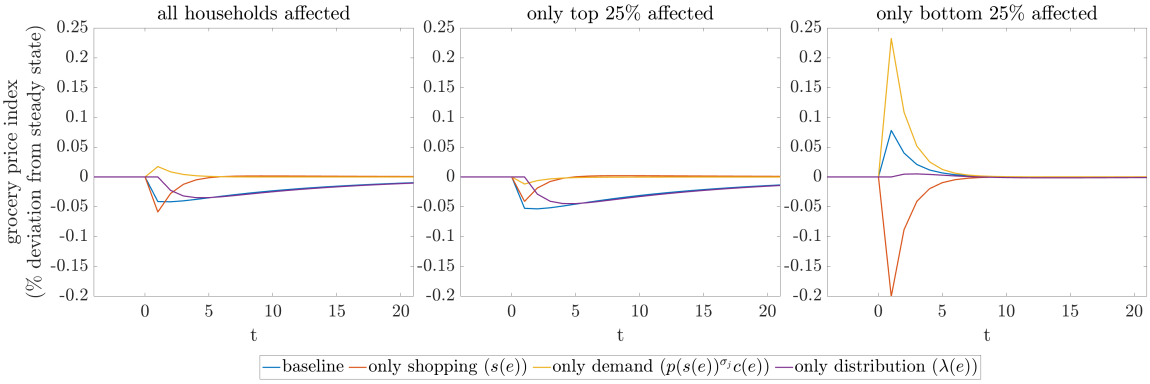


Figure 18: Aggregate Prices under Varying Incidence (Decomposition)

Note: Response of an aggregate Laspeyres index of posted prices $P_t^F = \frac{\sum_{j=1}^J C_j^{SS}\mu_{jt}^F}{\sum_{j=1}^J C_j^{SS}}$ to a 3% loss in aggregate labor earnings relative to the steady state with persistence 0.5, affecting (i) all households proportionately to their labor earnings, (ii) only the bottom quartile of labor earnings (proportionately to their earnings), or (iii) only the top quartile (proportionately to their earnings). Full response as baseline. Decomposed into response to changes in consumption policies (only c), shopping policies (only s), and the distribution of households (only dist), holding the respective others constant at steady state levels.