Consumer Search, Market Power, and the Distributional Effects of Inflation*

Zachary Bethune
Rice University

Shihan Shen
Rice University

June 2026

Abstract

This paper studies how inflation shapes consumers' search behavior, firms' market power, and the prices different households pay. We develop a model in which households face uninsurable income and expenditure risk and self-insure with liquid assets whose return inflation erodes. Search in the goods market is random and households choose shopping effort to find low price offers, as in Burdett and Judd (1983). In addition, given a set of offers, households have the outside option to continue searching, as in Diamond (1971). Calibrated to the low-inflation US economy of the 2010s, the model features positive sorting between wealth and prices paid: low-wealth households choose higher shopping effort, are more price-sensitive — driven by a high option value of continued search — and, as a result, pay lower prices than high-wealth households. Higher inflation unravels this sorting through two opposing channels: high-wealth households raise shopping effort relatively more, ending up in low-price stores where they compete with poorer households. In addition, a higher inflation tax erodes low-wealth households' option value of continued search, making them *less* price-sensitive. Low price stores then need to adjust prices by less, compared to high price stores, and their markups can even increase. We validate the model's central prediction using US NielsenIQ scanner data from 2010 to 2023: the positive sorting pattern, stable throughout the low-inflation period of 2010–2021, largely disappears during the Covid inflation episode.

JEL Classification: D82, D83, E40, E50
Keywords: Market power, inequality, welfare cost of inflation, consumer search.

*Contact: Bethune, bethune@rice.edu; Shen: sshen@rice.edu. We thank Guido Menzio, Ricardo Lagos, Guillaume Rocheteau, Kjetil Storesletten, Greg Kaplan, Michael Choi, Randall Wright, Lukas Nord, Jose Victor Rios-Rull, and Árpád Ábrahám for useful suggestions. We also thank seminar and conference participants at NYU, UC Irvine, UC Santa Barbara, the SED Conference in Athens, the 2025 Workshop on Money, Banking, Payments, and Finance at the Bank of Canada, the 2025 Mini-workshop on Household Consumption at Rice, and the 2025 Honorary Conference for Finn Kydland in Madrid. Finally, we thank the Center for Computational Insights on Inequality and Society (CIISR) at Rice for their generous support. All errors are the authors. Researcher's own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

Recent elevated inflation in many advanced and emerging economies has renewed interest in the costs of inflation and how they are distributed across households. In surveys, a growing consensus of households report inflation as their primary economic and social concern. Further, the idea that increases in firms' markups and profit margins are the leading cause of high inflation has gained traction, often termed "greedflation".[1] In this paper, we study how heterogeneous households adjust their shopping behavior in response to inflation, how firms' pricing strategies and market power respond in turn, and the consequences for the re-distributive effects of inflation.

We begin by providing empirical evidence on how household shopping behavior has changed during the high inflation episode in the US following Covid-19. Leveraging the NielsenIQ Consumer Panel, we construct a time-varying, within-market ranking of stores by their relative price level for narrowly defined goods—what we term a store's "expensiveness"—and analyze how households at different income levels shift their allocation of shopping trips along this expensiveness distribution over time. In the low-inflation period prior to 2021, we find that higher-income households are more likely to shop at stores with higher average prices compared to lower-income households. When the high-inflation episode began in 2021–2023, all households increased their share of trips to low-price stores, but the gap in store choice between high- and low-income households narrowed notably: higher-income households increased their relative presence at less expensive retailers. Additionally, we document that low-income households experienced a higher effective inflation rate relative to high-income households, even after controlling for differences in their consumption baskets. These findings are consistent with earlier results from the literature (Hobijn and Lagakos, 2005; Kaplan and Schulhofer-Wohl, 2017; Argente and Lee, 2021), but extended to the high inflation period in the US.

In terms of theory, previous studies have focused on two primary mechanisms through which inflation affects market power through consumer search. First, in models featuring price disper-

[1] For instance, Stantcheva (2024) reports that households list "Greed" as the second leading cause of high inflation, a belief more concentrated amongst low income households, see Figures 3 and 20. Figure 3 shows households with income below $40k report greed as the second leading cause while hosueholds with income above $125k report it as only the seventh leading cause, reporting monetary policy as the leading cause. Figure 20 shows that all households report inflation as the highest ecoomic and social issue, regardless of political party.

sion, arising either from nominal rigidities (as in, e.g., Benabou (1988, 1989, 1992) or Sara-Zaror (2025)) or due to search frictions (as in, e.g., Head and Kumar (2005) or Head, Lloyd-Ellis, and Sun (2014)), inflation changes the dispersion of prices and, in turn, the incentives of households to search for low prices. This liteature tends to find that inflation increases dispserion, increases consumer search, and therefore lowers firm markups. Second, a recent literature has focused on how inflation affects the cost of searching over time through the inflation tax (Choi and Rocheteau (2024), Bethune, Choi, Lotz, and Rocheteau (2024)). An increase in inflation, increases the cost of holding liquidity and reduces the outside option of households to search; increasing firm market power.

We develop a general equilibrium model that combines these two mechanisms. Households face uninsurable, idiosyncratic income and expenditure risk and smooth consumption through precautionary savings in liquid assets. Expenditure risk takes the form of *lumpy consumption*: at random times, a household values purchasing a good sold in a decentralized market, and it chooses how much effort to devote to searching for that good.[2] Search is random and in the spirit of Burdett and Judd (1983): with higher shopping effort, households draw more price quotes from a distribution of posted prices. Search also takes time. Price quotes arrive stochastically and, given the best price in hand, a household chooses whether to purchase, how much to purchase, or — as in Diamond (1971) — to decline and continue searching, postponing consumption while it waits for new offers. These search decisions combine the two primary forces studied in the literature. Purchases in the decentralized market are made out of liquid wealth, whose real return is eroded by inflation; inflation therefore affects both the incentive to exert shopping effort and the option value of postponing purchases. Firms post prices and face a mix of customers who differ in their shopping effort and purchase probabilities. In equilibrium, firms posting different markups earn equal expected profits, sustaining a non-degenerate distribution of markups.

In the model, a household's wealth and income determine the returns to search. The marginal value of finding a low price is higher for households with less wealth or lower income, because their opportunity cost of spending is higher and their option value of continued search is more attractive. These forces push poorer households toward higher shopping effort and make them

[2] Our model with no random search but featuring lumpy consumption is identical to Rocheteau, Weill, and Wong (2018).

more willing to walk away from high prices. We make part of this logic precise: in partial equilibrium, holding the distribution of posted prices fixed, we prove that near the borrowing constraint higher inflation raises the probability a household purchases at any given price and lowers its search effort — the inflation tax falls on the money balances a household must carry while it continues to search — while these direct effects vanish at high wealth. On the firm side, we show analytically that the demand elasticity a firm faces at a given price is a sales-weighted average of individual demand elasticities, which can be decomposed into three components: i) the elasticity of demand due to households switching to lower-price stores in their choice set, ii) the elasticity of demand due to households deciding to postpone purchases and continue searching, and iii) the elasticity of demand along the intensive margin. This decomposition provides the lens through which we interpret firms' pricing and market power throughout the paper.

We next turn to a quantitative analysis. We first extend the model to include an illiquid asset and a rich income process following Kaplan, Moll, and Violante (2018). We calibrate the model to the US in a low, 2% inflation steady state, and target the liquidity premium, wealth-to-income moments, and the average retail markup. In addition, we target the empirical (increasing) relationship between income and prices paid and the hump-shaped relationship between income and shopping effort, using evidence from NielsenIQ and the American Time Use Survey. The calibration delivers two key parameters that govern the frequency of search: opportunities for lumpy consumption arrive about once every two months, and searching households draw price quotes about every one and a half weeks. At these frequencies, walking away from a high price delays consumption by weeks, on average, so the option value of continued search is a quantitatively important margin.

In the calibrated steady state, equilibrium features positive sorting between wealth and prices paid. Low wealth households exert more shopping effort than high wealth households and have a higher demand elasticity at a given price, driven largely by their stronger outside option of continued search; they pay lower average prices and markups in equilibrium. We decompose the gradient in prices paid by wealth and find that heterogeneity in the purchase probability — the extensive margin of whether to buy or keep searching — is its main driver, rather than shopping effort alone. On the firm side, the equilibrium distribution of markups across firms is hump-shaped: low-price firms post low markups because they attract a disproportionate share of high-effort, high-elasticity (primarily low wealth) customers, while high-markup firms serve customers with lower elasticities.

We then simulate the model under a moderate, 5%-inflation rate regime and a high, 10%-inflation rate regime. In our baseline, money creation finances wasteful government spending, so inflation acts as a pure tax on liquid wealth, with no offsetting redistribution. Higher inflation moves shopping behavior in opposite directions across the wealth distribution. For households at the bottom of the distribution, the inflation tax makes continued search more costly: the option value of waiting for a better price falls, their purchase probability rises sharply, and their shopping effort declines. Households at the top of the distribution instead increase their shopping effort and sort more often into low-price stores.

These responses unravel the positive sorting between wealth and prices paid. Moving from 2% to 5% inflation, the average real price paid rises by around 8% for households in the lowest wealth percentiles, while it falls modestly for households in the top half of the distribution. A decomposition attributes the increase at the bottom almost entirely to the purchase-probability margin: low wealth households accept higher price draws rather than continue searching. The decline at the top instead reflects a change in market composition — as wealthy households search more and low wealth households become less price sensitive, the equilibrium distribution of posted prices and the customer base of low-price firms shift in favor of the wealthy. Higher inflation thus shifts relative prices paid against low wealth households, consistent with the compression in store-choice sorting we document in Section 2.

Related literature This paper relates to several strands of literature at the intersection of inflation, search, and heterogeneous-agent macroeconomics. It relates first to the literature on the distributional effects of inflation in economies with heterogeneous agents. Early contributions, including Bewley (1980, 1983), Erosa and Ventura (2002), Molico (2006), and Doepke and Schneider (2006), emphasize the balance-sheet channel of inflation, whereby changes in the price level redistribute resources between nominal asset holders and debtors. Subsequent work, such as Albanesi (2007), Chiu and Molico (2010, 2011), Algan, Challe, and Ragot (2011), Rocheteau et al. (2018); Rocheteau, Weill, and Wong (2021), and Auclert (2019), embeds monetary frictions in incomplete-markets frameworks and quantifies welfare effects when households differ in wealth, income risk, or portfolio composition. More recent studies, including Ferreira, Leiva, Nuno, Ortiz, Rodrigo, and Vazquez (2026), combine several of these channels within richer empirical settings. Across this literature, redistribution operates primarily through nominal balance sheets, income dynamics, or differences in expenditure baskets. By contrast, we focus on a goods-market channel in which inflation alters households' search incentives and outside options in decentralized product markets, thereby affecting firms' market power and the distribution of markups.

A second strand of literature links search frictions to price dispersion and market power. Foundational contributions such as Butters (1977a), Varian (1980), and Burdett and Judd (1983) show how consumer search generates endogenous price dispersion and imperfect competition, while sequential search models following McCall (1970) and Diamond (1971) emphasize the role of reservation strategies and outside options in governing demand elasticities and markups. Our framework builds on this tradition: firms randomize over prices, and search intensity is endogenous and varies with household characteristics and inflation.

In monetary environments, this literature identifies two channels through which inflation interacts with search. In models combining menu costs and consumer search, such as Benabou (1988, 1989, 1992) and Sara-Zaror (2025), higher inflation increases price dispersion and search incentives, thereby reducing market power. Burdett and Menzio (2018) combine menu costs and search frictions to study how inflation affects dispersion and markups in equilibrium. In frictional dispersion models without menu costs, including Head and Kumar (2005) and Head et al. (2014), monetary conditions similarly affect equilibrium markups through decentralized trading. A second channel operates through liquidity and outside options: in monetary search frameworks such as Diamond (1993), Choi and Rocheteau (2024), and Bethune et al. (2024), higher inflation lowers the real return to liquid balances and weakens buyers' continuation values, potentially increasing firms' market power. Baqaee, Farhi, and Sangani (2024) study how markup heterogeneity affects the pass-through of nominal shocks and the resulting changes in price dispersion. We incorporate both channels in a heterogeneous-agent general equilibrium setting, allowing inflation to affect search effort ex ante through dispersion and purchase decisions ex post through the real return to liquid wealth.

Third, there is a recent and growing line of work that embeds search frictions in heterogeneous-agent environments. Kaplan and Menzio (2016) document how household heterogeneity generates systematic differences in prices paid in a non-monetary setting. More recent working papers, including Pytka (2024), Nord (2024), Sangani (2023), and Nord and Kaplan (2026), introduce search into incomplete-markets models and show how heterogeneity in preferences or wealth drives price dispersion and markups. These papers abstract from monetary policy and inflation as determinants of outside options and market power. We instead embed search in a monetary incomplete-markets framework in which inflation affects the real return to liquid assets, thereby altering search behavior, sorting, and equilibrium markups.

Finally, a related empirical literature documents heterogeneity in prices paid and shopping behavior across households. Aguiar and Hurst (2007) show that lower-income households devote more time to shopping and pay lower prices for identical goods. Stroebel and Vavra (2019) document how markups respond to wealth-induced demand shifts. Edmond, Midrigan, and Xu (2023) show that markups are positively related to firm size, while Faber and Fally (2022) document that higher-income households disproportionately purchase from larger firms. Consistent with these patterns, Mongey and Waugh (2025) and Sangani (2023) show that markups paid increase with income, and Auer, Burstein, Lein, and Vogel (2024) find that higher-income households exhibit lower demand elasticities. Relatedly, Argente and Lee (2021) and Jaravel (2019) document systematic differences in inflation experiences across income groups. Together, this evidence suggests that heterogeneity in shopping behavior and firm exposure translates into systematic differences in markups and inflation exposure across households.

The remainder of the paper is organized as follows. Section 2 provides empirical evidence about how household shopping responds to aggregate inflation. Section 3 presents the model and characterizes the stationary equilibrium. Section 4 characterizes analytically, in partial equilibrium, how inflation affects household shopping behavior across the wealth distribution. Section 5 extends the model for quantitative analysis and describes the calibration strategy and model fit. Section 6 discusses the cross-sectional relationships in household shopping behavior and markups paid. Section 7 examines the response of household behavior to changes in inflation. Section 8 concludes.

2 Empirics: how do households respond to aggregate inflation?

2.1 Data

In this section, we examine how households adjust shopping behavior in response to changes in the aggregate price environment. We use the NielsenIQ Consumer Panel, accessed through the Kilts Center for Marketing. The panel records household-level purchases from individual shopping trips of consumer packaged goods across a broad set of retail outlets in the United States. Our sample covers the period 2010–2023. The years 2010–2019 were characterized by relatively low and stable inflation, while 2020–2023 include the Covid-19 disruption and the subsequent sustained rise in inflation. This time span allows us to compare shopping behavior across markedly different price environments.

For each shopping trip, the data report purchased items identified by Universal Product Codes (UPCs), quantities, total expenditure, purchase dates, and retailer identity. UPCs uniquely identify specific packaged goods, such as a particular size of cereal. The panel includes approximately 40,000 to 60,000 households per year and provides detailed demographic information, including household income and geographic location.

Geographic location is measured using Designated Market Areas (DMAs), which partition the United States into local retail markets. Because pricing and competitive conditions operate at the local level, the DMA provides a natural unit for defining market-specific pricing environments. Households are followed over time, allowing repeated observation of shopping behavior across retailers within a given local market.

2.2 Variable Construction

To capture how households reallocate shopping trips across retailers, we construct measures of store-level price positioning, which we refer to as store "expensiveness." We measure store price positioning separately in each quarter over 2010–2023. This time-varying measure allows store-level prices to reflect contemporaneous pricing decisions, promotional activity, and demand conditions in both low- and high-inflation periods.

Because observed prices vary across products and over time, raw comparisons across retailers would conflate store-level price differences with changes in product mix and aggregate price movements. To isolate relative price positioning, we net out product-by-quarter components so that store-level measures reflect differences across retailers within a local market at a given point in time.

Store-level price positioning. The NielsenIQ data report the retailer associated with each purchase and the DMA in which the household resides. Because pricing strategies vary across geographic markets, we define a store as a retailer-DMA pair. This definition captures pricing behavior within a local market while allowing the same retail chain to differ across regions.

We construct quarterly unit prices at the UPC-retailer-DMA level as total expenditure divided by total quantity purchased. For each quarter t, we estimate

$$\log p_{urdt} = \theta_{rdt} + \phi_{ut} + \varepsilon_{urdt}, \tag{1}$$

where p_{urdt} denotes the unit price of UPC u at retailer r in DMA d during quarter t. The fixed effects ϕ_{ut} absorb product-specific price movements in quarter t that are common across retailers, including aggregate inflation and product-level cost shocks. The store-by-quarter effects θ_{rdt} therefore capture the relative price level of retailer r within DMA d at time t.

The regression is weighted by total quantities sold at the UPC-retailer-DMA-quarter level so that store positioning reflects high-volume products more heavily than infrequently purchased items.

Because overall price levels may differ across local markets in a given quarter, we measure store positioning relative to other retailers operating in the same DMA at time t. Specifically, we remove DMA-by-quarter means from θ_{rdt} and rank stores within each DMA and quarter. To ensure meaningful cross-sectional comparisons, we restrict attention to DMA-by-quarter cells with at least five active retailers.

Within each DMA and quarter, we construct the percentile

$$\pi_{rdt} = F_{dt}(\theta_{rdt}), \tag{2}$$

where $F_{dt}(\cdot)$ denotes the empirical distribution of store price positioning in DMA d at time t. The percentile π_{rdt} summarizes a store's location in the local price distribution in that quarter.

Income groups and store-choice measures. Households are grouped into within-year income quartiles based on reported annual household income and sampling weights.[3] We use relative income ranks within each year rather than the nominal income bins provided by NielsenIQ, which are defined in fixed dollar ranges and do not adjust for inflation over time. This approach ensures comparability of income groups across periods despite changes in the aggregate price level.

For each income group g and quarter t, we summarize store choice using the mean expensiveness percentile of visited stores,

$$\bar{\pi}_{g,t} = E[\pi_{rdt} \mid g, t] \,, \tag{3}$$

which captures the average position of the stores visited by group g in the local price distribution at time t.

As a complementary threshold-based measure, we compute the share of trips to stores below a given percentile cutoff of the expensiveness distribution,

$$s_{g,t}^{\text{bot}X} = \Pr(\pi_{rdt} \leq X \mid g, t) \,, \tag{4}$$

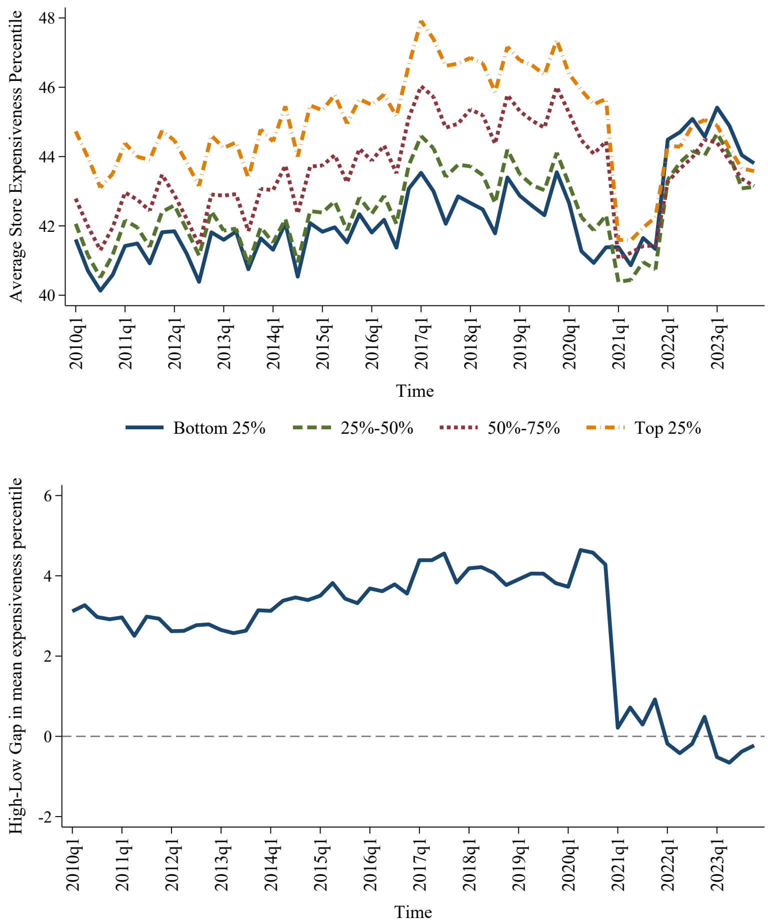
where $X \in \{0.20, 0.25, 0.30, 0.35\}$ denotes alternative definitions of the low-price segment.

Because threshold-based measures are mechanically sensitive to the choice of cutoff and exhibit greater sampling variability, we report these results in the appendix and focus in the main text on the continuous percentile-based measure, which uses information from the entire distribution.

2.3 Main Empirical Fact

Figure 1 presents the evolution of store choice across income groups over time. Panel A plots the mean expensiveness percentile of visited stores separately for each income quartile, while Panel B summarizes the dispersion by reporting the difference between the top and bottom quartiles.

[3] NielsenIQ reports income in discrete nominal brackets. Because these brackets do not adjust for inflation and may span multiple quartile cutoffs within a given year, households within the same nominal bracket can be assigned to adjacent quartiles based on their weighted rank. Results are robust to alternative tie-breaking rules.



Notes: Panel A plots the mean expensiveness percentile of visited stores separately for each income quartile. Panel B plots the difference between the top and bottom income quartiles. Percentiles are computed within each DMA and quarter. All measures are trip-weighted.

Figure 1: Income Differences in Store Choice

Two patterns stand out. First, during the low-inflation period of the 2010s, store choice displays a clear and persistent gradient across the income distribution. Higher-income households consistently shop at stores that rank higher in the contemporaneous local price distribution. The ordering across quartiles is monotonic and stable: by the late 2010s, bottom-quartile households shop at stores around the 42nd percentile on average, while top-quartile households shop near the 46th percentile. The income gradient is economically meaningful and shows little evidence of compression prior to the high-inflation episode.

Panel A also reveals broader movements that affect all groups. Average store percentiles trend upward over much of the 2010s, and the Covid period is associated with sizable shifts for every quartile. These common movements likely reflect changes in the retail environment, entry and exit of stores, promotional intensity, or pandemic-related disruptions. Our focus, however, is on how store choice differs across income groups rather than on aggregate shifts in store positioning.

Against this background, the central fact is a sharp compression in dispersion once inflation accelerates. Beginning in 2021, the distance between the quartile lines in Panel A visibly narrows. The adjustment is driven primarily by a downward shift in the store percentiles of higher-income households, while the bottom quartile changes comparatively little. As a result, dispersion in store positioning across income groups narrows during the high-inflation period.

Panel B isolates this compression directly by plotting the difference between the top and bottom income quartiles. Prior to the inflation surge, this gap fluctuates around roughly four percentile points. Beginning in 2021, it declines sharply and approaches zero during 2021-2023. The magnitude of this reduction implies that the income gradient in store choice, which had remained stable for nearly a decade, largely disappears once inflation rises.

The compression in the income gradient is robust to alternative measurement choices and ranking procedures. First, the result remains when store choice is measured using expenditure shares rather than trip counts, as is shown in Appendix B, Figure 16. Because expenditure shares mechanically reflect within-store price movements during inflation, the persistence of the compression indicates that the pattern is not driven by arithmetic effects of rising prices, but by changes in cross-store allocation. Second, in the main text we focus on trip-based measures rather than expenditure shares in order to isolate changes in store choice from mechanical effects driven by within-store price movements during inflation. the findings are not sensitive to the precise definition of the low-price segment. In Appendix B, Figure 17 indicates that alternative thresholds for defining the lower tail of the store distribution yield similar convergence across income groups .

Finally, we consider a time-invariant store ranking based on pre-inflation price dispersion. While the baseline allows store percentiles to adjust within each DMA and quarter, fixing the ranking isolates reallocation from contemporaneous shifts in store positioning. Figure 18 in Appendix B shows that the high-low gap continues to narrow under this specification, indicating that the compression reflects genuine re-sorting across the retail price distribution.

3 A model of the shoopping and consumption response to inflation

3.1 Environment

Time is continuous and infinite, $t \in \mathbb{R}_+$. There exists a continuum of infinitely-lived households that discount the future at rate $\rho > 0$. There are two types of goods in the economy. The first is a homogeneous good produced and sold in a competitive market and consumed in flows, c_t. The second is a homogeneous good produced and sold in a decentralized goods market and consumed in lumps, y_t. There is a large measure of firms producing the competitive good and a measure $N > 0$ of firms producing the decentralized good.

A household's lifetime expected utility is given by

$$\mathbb{E}\left[\int_0^{+\infty} e^{-\rho t} u(c_t^i)dt + \sum_{n=1}^{\infty} e^{-\rho T_n}\left[U(y_{T_n}^i) + \zeta_{T_n}^i\right]\right]. \tag{5}$$

Households value flow consumption, c_t^i, according to an instantaneous utility function $u(c_t^i)$, where u is twice continuously differentiable with $u' > 0$, $u'' < 0$, and $\lim_{c\to 0} u'(c) = +\infty$. In addition, households have preferences to consume, y_t^i that generate lumps of utility, $U(y_t^i) + \zeta_t^i$, at random points in time $\{T_n\}_{n=1}^{\infty}$. There are two components of a household's preference for lumpy consumption. The first depends on the intensive margin; households receive a lump of utility $U(y)$. We assume U is twice continuously differentiable with $U(0) = 0$, $U' > 0$, $U'' < 0$, and $\lim_{c\to 0} u'(c) = +\infty$. The second component depends on the extensive margin and is a random shock. Specifically, we assume $\zeta_t^i = \zeta_{0,t}^i \mathbf{1}\{y_t = 0\} + \zeta_{1,t}^i \mathbf{1}\{y_t > 0\}$ and that $\boldsymbol{\zeta}_t^i = \{\zeta_{0,t}^i, \zeta_{1,t}^i\}$ are drawn *iid* across time and households and are distributed according to the Type-I Extreme Value distribution

$$\mathcal{Z}(\boldsymbol{\zeta}_t^i) = exp\left\{-\left(\sum_{j=0}^{1} exp\{-\eta\zeta_{j,t}^i\}\right)\right\}, \tag{6}$$

where $\eta > 0$ represents the dispersion in preferences between purchasing and not purchasing. The random component of utility captures idiosyncratic preferences for purchasing a good and leads to a convenient characterization of the probability of consuming.

Households supply labor inelastically and continuously in a competitive labor market with wage w_t. Labor is perfectly substitutable across firms. A household's idiosyncratic labor productivity, e_t^i, is stochastic and evolves according to an independent two-state Markov chain with transitions given by Poison rates $\mathcal{P}_{e,e'}$. The good c_t is produced according to a linear technology and sold in a competitive market. A representative firm has production technology $Y_t^c = L_t^c$, which yields profits $\Pi_t^c = P_t^c Y_t^c - w_t L_t^c$. Optimal labor demand implies $P_t = w_t$ and setting the competitive market good as numeraire gives $w_t = 1$ for all t. The good y_t is produced according to linear technology $Y_t = (1/\kappa)L_t$. Since labor is perfectly substitutable across goods, profits are $\Pi_t = (p_t - \kappa)Y_t$.

The idiosyncratic stochastic process $\{T_n^i\}_{n=1}^{\infty}$ in (5) is generated by two independent, idiosyncratic stochastic processes for preferences and search. First, households can be in one of two states with respect to their preference to consume the lumpy good y_t, $\epsilon_t \in \{0, 1\}$, where a *searching* household, $\epsilon_t = 1$, has a preference to consume y and an *idle* household, $\epsilon_t = 0$, does not. Idle households do not search and transition to being active at Poisson rate $\lambda > 0$.

Searching households have a preference for lumpy consumption and look for goods to consume. At each point in time, all firms post a price, p_t, with commitment, denominated in units of numeraire. Households search for goods by randomly contacting firms for their prices through a time-consuming process. Specifically, the time it takes to collect prices is random, distributed exponentially with parameter $\alpha > 0$, and identically and independently distributed across all households. However, households have a choice of the effort they put into shopping. Specifically, with probability $s \in [0, 1]$, the searching household is *active* and visits a k number of firms to observe their price, where k is drawn from a Poisson distribution with mean $\mu > 1$. With complement probability $1 - s$, the household is *passive* and observes exactly $k = 1$ price offer from firms. The probability s is a choice of the searching household, subject to a flow utility cost $\delta(s)$ with $\delta' > 0$, $\delta'(0) = 0$, and $\lim_{s\to 1} \delta'(s)$ is a large number. The search process is similar to the noisy random search framework in Butters (1977b), Varian (1980), and Burdett and Judd (1983). After a household collects their price draws, they choose one firm in which to purchase y. Once a household consumes the good, they transition back to being idle.[4] The timing of a household's states and decisions for lumpy consumption is depicted in Figure 2.

[4]Generally, we can allow active agents to transition back to being idle either after consuming or at some exogenous Poisson rate $\gamma > 0$ without consuming. This captures the fact that preferences might disappear over time. Here and in the calibration, we assume $\gamma = 0$.

There is a single real risk-free asset, a, with exogenous return $r_t > 0$. Households face an exogenous borrowing constraint $a_t^i \geq \underline{a}$, with $\underline{a} \in (-\infty, 0]$. We will focus on the case in which the risk-free asset is money. In this case, we assume the supply of money, M_t, is exogenous and grows at the rate $\pi_t \geq 0$. The price of money in units of numeriare is given by ϕ_t, which implies the real return on money is $r_t = -\dot{\phi}_t/\phi_t$. We will consider two cases in terms of how newly-created money is used: i) either to finance wasteful government spending (in units of numeraire) or ii) distributed lump-sum to all households. Generally let τ_t represent the transfer (tax if negative) of real resources to the household.

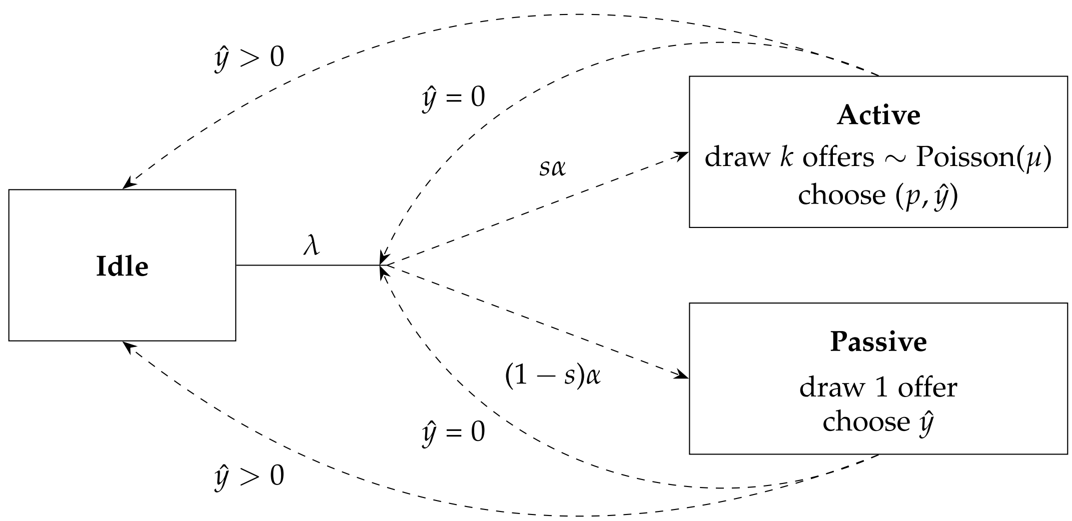


Figure 2: Flow of a household across states for lumpy consumption

3.2 Stationary Equilibrium

This section characterizes the stationary equilibrium of the economy under a constant money growth rate $\pi_t = \pi$. The problem of a household is characterized first, followed by the price setting problem of firms producing the decentralized good. Then, we characterize the distribution of posted prices, state aggregation conditions, and define the equilibrium.

Household value functions and decision rules Let $W(a, e)$ represent the expected lifetime discounted utility of a searching household, with current wealth a and labor productivity e that satisfies the following Hamilton-Jacobi-Bellman (HJB) equation,

$$\rho W(a,e) = \max_{c,s\in[0,1]} \; u(c) - \delta(s) + \alpha V(a,e,s) + \mathcal{P}_{e,e'}\left[W(a,e') - W(a,e)\right] + W_a(a,e)\left[ra + e + \tau - c\right]. \quad (7)$$

In (7), the second term represents the expected utility change from receiving an opportunity to consume the decentralized market good y, the third term represents the change in utility from idiosyncratic productivity shocks, and the last term represents the drift in utility caused by changes in a household's wealth given by the budget constraint, $\dot{a} = ra + e + \tau - c$. The searching household satisfies their Euler equation that $u'(c^1) = W_a(a,e)$ and the first-order condition on shopping effort $\delta'(s) = \partial V(a,e,s)/\partial s$.

The function $V(a,e,s)$ represents the expected change in lifetime utility of a searching household with shopping effort s. Let $F(p)$ denote the equilibrium distribution of posted prices. Let $G(p;s)$ represent the distribution of the lowest price a household draws, given by

$$G(p;s) = (1-s)F(p) + s\sum_{k=0}^{\infty} \frac{\mu^k e^{-\mu}}{k!}\left[1 - (1-F(p))^k\right]. \quad (8)$$

In (8), with probability $1-s$ the household is passive and draws a single draw from $F(p)$. With probability s, the household is active and draws k offers with probability $\mu^k exp\{-\mu\}/k!$. The term $\left[1-(1-F(p))^k\right]$ represents the probability that the lowest price quote from k draws is below p. Given $G(p;s)$, the expected utility change of a household can be written as

$$V(a,e,s) = \int_p \mathbb{E}_\zeta \max\left\{\max_{y>0}\; U(y) + \zeta_1 + X(a-py,e) - W(a,e),\; \zeta_0\right\} dG(p;s) \quad (9)$$
$$\text{subject to} \quad a - py \geq \underline{a},$$

where $X(a,e)$ represents the lifetime expected utility of being idle. In (9), given a lowest price offer p, a household chooses between i) purchasing a strictly positive amount of consumption, $y > 0$, at that price, receiving utility $U(y) + \zeta_1$, and transitioning to being idle with wealth $a - py$, versus ii) not purchasing, $y = 0$, receiving utility flow ζ_0, and continuing being active in search. Given a household chooses a strictly positive amount of lumpy consumption, the solution is given by

$$U'(y) \geq pX_a(a - py, e), \quad \text{“="} \quad \text{if} \quad py < a - \underline{a}. \tag{10}$$

Households equate the marginal utility gain from lumpy consumption with the marginal lifetime utility loss, equal to the marginal value of wealth times the price. If the borrowing constraint is binding, then $y = (a - \underline{a})/p$. Let $\hat{y}(p; a, e)$ stand for the solution to (10) and $\hat{V}(p; a, e) \equiv U(\hat{y}) + X(a - p\hat{y}, e) - W(a, e)$. Notice, we can write $\hat{V}(a, e)$ as

$$\hat{V}(p; a, e) + \zeta_1 = \underbrace{U(\hat{y}) + \zeta_1}_{\text{flow payoff}} - \underbrace{\left[X(a, e) - X(a - p\hat{y}, e)\right]}_{\text{option value of wealth}} - \underbrace{\left[W(a, e) - X(a, e)\right]}_{\text{option value of search}}. \tag{11}$$

Consuming y leads to a flow payoff of $U(\hat{y})$, but the household gives up two option values. The second term in (11) is the option value of keeping $p\hat{y}$ units of wealth, which is lower for wealthier households. The last term in (11) represents the option value of search — the change in lifetime utility from transitioning from being an active searcher to being idle.

Given any realization of $\boldsymbol{\zeta}$ and a household's lowest price p, it is possible that $\hat{V}(p; a, e) + \zeta_1 < \zeta_0$ and the household exercises their outside option of not purchasing any variety and continuing to search. This possibility stems from households' two outside options. The extensive margin problem is then determined by the probability of making a purchase conditional on p, denoted by $\varrho(p; a, e)$. Given (6), it's solution is given by

$$\varrho(p; a, e) = \frac{exp\left\{\eta \hat{V}(p; a, e)\right\}}{1 + exp\left\{\eta \hat{V}(p; a, e)\right\}}. \tag{12}$$

A household's expected purchases at a firm with price p (if the price is the lowest of a household's choice set) is $\varrho(p; a, e)\hat{y}(p; a, e)$. It is easy to show that $\partial\varrho/\partial p < 0$ and $\partial\hat{y}/\partial p < 0$. Hence, a household has both an extensive and intensive margin elasticity of demand.

Next, we compute the optimal shopping effort, $s(a, e)$. Given policy functions $\hat{y}(p; a, e)$ and $\varrho(p; a, e)$, value function $\hat{V}(p; a, e)$, and the distribution of $\boldsymbol{\zeta}$ in (6), the expectation inside the integral in (9) is equal to

$$EV(p;a,e) \equiv \mathbb{E}_{\zeta} \max\left\{\hat{V}(p;a,e)+\zeta_1, \zeta_0\right\} = \frac{1}{\eta} log\left[1+e^{\eta \hat{V}(p;a,e)}\right] = \hat{V}(p;a,e) - \frac{1}{\eta} log\, \varrho(p;a,e). \quad (13)$$

The second equality in (13) follows from (12).[5] Under the assumption $F(p)$ has no mass points (which we later show holds), the probability density function of $G(p;s)$ in (8) is given by $g(p;s) = \left[(1-s)+s\mu e^{-\mu F(p)}\right]$. We can write the first-order condition for shopping effort as

$$\delta'(s) = \alpha \int_p EV(p;a,e)\left[\mu e^{-\mu F(p)} - 1\right] dF(p). \quad (14)$$

Households optimally set the marginal utility cost of shopping effort, $\delta'(s)$, equal to the expected marginal utility gain, given by the right-hand side of (14). For any price in the support of $F(p)$, the marginal gain in searching with more effort is the change in the probability of finding that price, equal to $\mu e^{-\mu F(p)} - 1$. Notice, at the lowest price $F(\underline{p}) = 0$ and the marginal gain in probability is $\mu - 1 > 0$. For the highest price $F(\bar{p}) = 1$ and the marginal gain in probability is $\mu e^{-\mu} - 1 < 0$. In words, an increase in search intensity increases the probability the household receives low price offers. Note, however, that the marginal gains integrate to $\int_p [\mu e^{-\mu F(p)} - 1] dF(p) = -e^{-\mu} < 0$: an active household draws no offer at all with probability $e^{-\mu}$, so greater effort also raises the risk of coming away empty-handed. This property links the effort margin to the purchase margin and plays a central role in the analytical results of Section 4.

How valuable more search is for a household depends on the shape of the function $EV(p;a,e)$ in (13), which depends on their state (a,e). Figure 3 illustrates $EV(p;a,e)$ for two households with different wealth levels. There are a few important features to point out that drive cross-sectional heterogeneity in shopping effort. First, any household's expected value is decreasing in price, $\partial_p EV(p;a,e)$, which guarantees that there is always a solution to (14) for any household given $\lim_{s\to 1} \delta'(s)$ a large enough. Second, a household's expected value is increasing in wealth at any price, $\partial_a EV(p;a,e) > 0$. Wealthier households demand more goods along both the extensive and intensive margin that implies they have a higher gain from searching. However, the marginal

[5] Throughout, we center the taste shocks ζ to have mean zero, which removes the Euler–Mascheroni constant from (13). The location normalization is not innocuous here: because an active searcher draws no offer at all with probability $e^{-\mu}$, additive constants in EV do not cancel out of the shopping effort choice in (14).

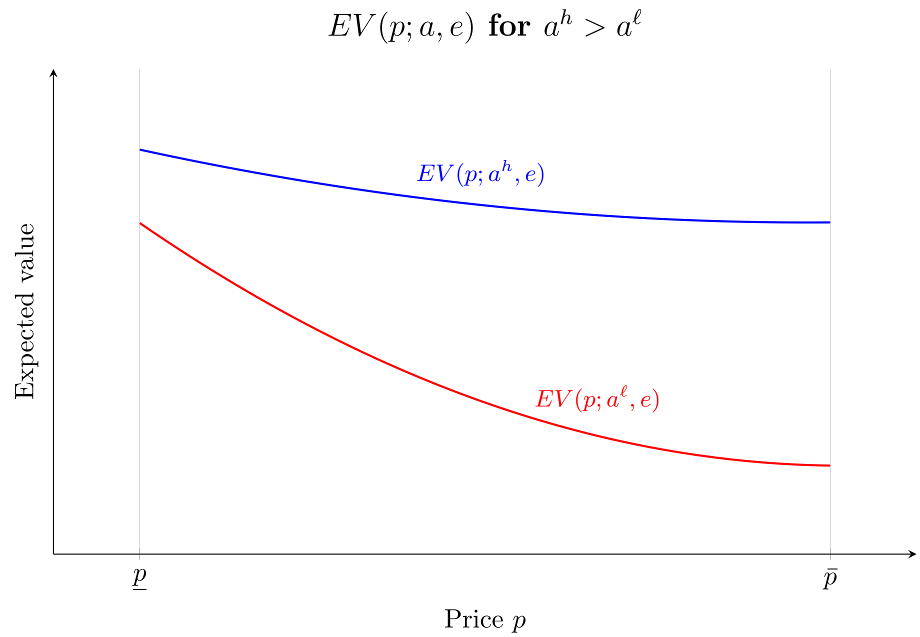


Figure 3: Illustration of the expected value function $EV(p; a, e)$.

utility cost of paying a higher price is decreasing in wealth, $\partial_a \partial_p EV(p; a, e) < 0$, implying a larger marginal gain from searching for lower wealth households. Similar patterns hold across income levels.

The value of being idle is given by the HJB equation

$$\rho X(a, e) = \max_{c \geq 0} \; u(c) + \lambda \left[W(a, e) - X(a, e)\right] + \mathcal{P}(e, e') \left[X(a, e') - X(a, e)\right] + X_a(a, e) \left[ra + e + \tau - c\right]. \tag{15}$$

Equation (15) follows closely to (7) with the exception of the second term that captures the change in lifetime utility from receiving a preference shock. Consumption of an inactive agent satisfies the Euler equation, $u'(c^0) = X_a(a, e)$.

Stationary distribution of wealth and income Let $\Gamma(a, e, \epsilon)$ stand for the stationary distribution of active, $\epsilon = 1$, and inactive, $\epsilon = 0$, households and $\gamma(a, e, \epsilon)$ stand for the density, potentially with Dirac mass points. Appendix **??** shows that the system $\gamma(a, e, \epsilon)$ must solve the following

Kolmogorov Forward Equations (KFEs),

$$
\begin{aligned}
0 = &- \frac{\partial}{\partial a}\left[s(a,e,1)\gamma(a,e,1)\right] - \mathcal{P}(e,e')\gamma(a,e,1) + \mathcal{P}(e',e)\gamma(a,e',1) \qquad (16)\\
&-\alpha\gamma(a,e,1)\left[\int_p \varrho(p;a,e)dG(p;s(a,e))\right] + \lambda\gamma(a,e,0), \text{ and}\\
0 = &- \frac{\partial}{\partial a}\left[s(a,e,0)\gamma(a,e,0)\right] - \mathcal{P}(e,e')\gamma(a,e,0) + \mathcal{P}(e',e)\gamma(a,e',0) \qquad (17)\\
&+\frac{\partial}{\partial a}\left[\int_x \alpha\gamma(x,e,1)\left[\int_p \mathbf{1}\{x - p\hat{y}(p;x,e) \leq a\}\varrho(p;x,e)dG(p;s(x,e))\right]dx\right] - \lambda\gamma(a,e,0).
\end{aligned}
$$

The first three terms in (16) and (17) are common in the benchmark incomplete markets model (e.g. Achdou, Han, Lasry, Moll, and Lions, 2022). For instance in (16), the first term represents changes in the density due to continuous movements of wealth while the last two terms capture the outflows and inflows due to changes in labor productivity. The last two terms in both (16) and (17) are new. Lumpy consumption opportunities and search generate transitions between active search and inactivity. In (16), a measure of $\alpha\gamma(a,e,1)$ of active agents receive an opportunity to consume the lumpy good. They draw a lowest price from distribution $G(p;s)$ and, given p, purchase a strictly positive quantity with probability $\varrho(p;a,e)$. The last term in (16) represents the inflow of idle households to being active. The second-to-last term in (17) represents the inflow into inactivity from active households making a purchase. The term captures that active households with various amounts of incoming wealth, x, decide to purchase lumpy consumption and whose optimal purchases lead to an after-consumption amount of wealth at approximately a. Finally, the last term in (17) represents the outflow of inactive households to being active.

Firm price setting and the distribution of posted prices An individual firm's price setting problem is static; since they are infinitesimal, they do not affect the law of motion of the aggregate state. Consider the expected profits of a firm posting price p at any point in time, denoted by $\Pi(p)$, given the equilibrium distribution of posted prices. First, conditional on a household of type (a,e) observing the firm's price, the probability they are passive is $1 - s(a,e)$ while the probability they are active is $s(a,e)$.

The expected demand from a passive household with price draw p is $\varrho(p;a,e)\hat{y}(p;a,e)$ since they only have one offer in-hand. The expected profit from a active household with price draw p

is

$$\sum_{k=0}^{\infty}(k+1)\mu^{k+1}\frac{e^{-\mu}}{(k+1)!}[1-F(p)]^{k}\varrho(p;a,e)\hat{y}(p;a,e)=\mu e^{-\mu F(p)}\varrho(p;a,e)\hat{y}(p;a,e). \tag{18}$$

In (18), an active household can draw $k=\mathbb{N}^{+}$ offers in addition to the firm's, with probability $(k+1)\mu^{k+1}e^{-\mu}/(k+1)!$. The probability that the firm has the lowest price amongst all the offers is $[1-F(p)]^{k}$. Conditional on being the lowest price, the firm expects to sell $\varrho(p;a,e)\hat{y}(p;a,e)$ units of the good. Taking expectations across households, the probability a firm posting price p is the lowest price for a household of type (a,e) is

$$h(p;a,e)=1-s(a,e)+s(a,e)\mu e^{-\mu F(p)}. \tag{19}$$

Then, we can write a firm's expected profits $\Pi(p)$ as

$$\Pi(p)=\underbrace{\sum_{e}\int_{a}\frac{\alpha\gamma(a,e,1)}{N}h(p;a,e)\varrho(p;a,e)\hat{y}(p;a,e)da}_{\equiv Y(p)}(p-\kappa) \tag{20}$$

In (20), a firm posting p expects to face customers with heterogeneous wealth and income (a,e). For each household type, there is a measure $\alpha\gamma(a,e,1)/N$ of shoppers with price draws per firm. Importantly, we can write profits as a standard expression of the aggregate demand times the price margin: $\Pi(p)=\mathcal{Y}(p)(p-\kappa)$.

Let $\mathbb{P}$ stand for the equilibrium support of the posted price distribution $F(p)$. Our first Lemma shows that any price in the support can be written as a profit maximizing price. All proofs are provided in Appendix A.

Lemma 1. *A price $p\in\mathbb{P}$ if an only if $p\in argmax_{x}\,\Pi(x)$. Further, all prices can be written as a markup over marginal cost,*

$$p=\frac{-\epsilon^{d}(p)}{1-\epsilon^{d}(p)}\kappa, \tag{21}$$

where $\epsilon^{d}(p)=-\frac{\partial ln(\mathcal{Y}(p))}{\partial ln(p)}$ is the elasticity of demand at p.

Intuitively, firms must be indifferent ex-ante to posting any price in the equilibrium support. If a firm was posting a price p that does not maximize profits, given the strategies of other firms,

$F(p)$, then they would deviate to the profit maximizing price. The second result relies on showing that a maximum exists for any price in the support so the first-order condition is necessary.

We can decompose the elasticity of demand $\epsilon^d(p)$ as a sales-weighted average of *individual* price elasticities, $\epsilon^d(p; a, e)$,

$$\epsilon^d(p) = \sum_e \int_a \epsilon^d(p; a, e)\omega(p; a, e)da, \tag{22}$$

where $\omega(p; a, e) \equiv (\alpha/N)[1 - s(a, e) + s(a, e)\mu e^{-\mu F(p)}]\varrho(p; a, e)\hat{y}(p; a, e)\gamma(a, e, 1)/Y(p)$ is the share of sales from active households in state (a, e). The individual elasticity of demand at price p can be further decomposed into three terms,

$$\epsilon^d(p; a, e) = \underbrace{-\frac{\partial h(p; a, e)}{\partial p}\frac{p}{h(p; a, e)}}_{\epsilon^h(p;a,e)} + \underbrace{-\frac{\partial \varrho(p; a, e)}{\partial p}\frac{p}{\varrho(p; a, e)}}_{\epsilon^\varrho(p;a,e)} + \underbrace{-\frac{\partial \hat{y}(p; a, e)}{\partial p}\frac{p}{\hat{y}(p, a, e)}}_{\epsilon^y(p;a,e)}. \tag{23}$$

The first term in (23) captures that, in response to a price increase, the household may be induced to switch to a firm with a lower price. It can be written as

$$\epsilon^h(p; a, e) = \frac{s(a, e)\mu^2 e^{-\mu F(p)} f(p) p}{[1 - s(a, e) + s(a, e)\mu e^{-\mu F(p)}]}. \tag{24}$$

Notice any cross-sectional heterogeneity in ϵ^h is entirely driven by shopping effort $s(a, e)$. Further, $\partial\epsilon^h/\partial s > 0$ since more shopping effort leads to an ex-ante larger choice set; it is more likely the household 'knows' of a lower price firm in which to purchase the good.

The second term in (23) captures that an increase in price may cause the household to not purchase the good entirely and continue searching. It can be written as

$$\epsilon^\varrho(p; a, e) = \eta\,[1 - \varrho(p; a, e)]\,\frac{-\partial \hat{V}(p; a, e)}{\partial p} p. \tag{25}$$

The firm term in (25), η, captures the dispersion in tastes to consume good y at a point in time; as dispersion increases (high η) a consumer is more likely to substitute towards searching in the future their option value of search is larger. The second term captures how captive a particular household is to the firm's price — at price p how likely is it that the household chooses to continue searching given by $1 - \varrho(p; a, e)$. Since ϱ is monotone in $\hat{V}$, the larger a household's valuation for the good, either because they experience a high payoff $U(y)$ or because they have low outside options of maintaining wealth or searching (discussed after (11)), their demand is less elastic. The last term in (25) captures how sensitive a given household's valuation is to the price. Using the definition of $\hat{V}$ and the first order condition (10), it equals $-\frac{\partial \hat{V}}{\partial p} p = U'(\hat{y})\hat{y}$. For constant relative risk aversion preferences with $U(c) = c^{1-\sigma}/(1-\sigma)$ and $\sigma > 1$, the price effect is larger for wealthier households, since $\partial \hat{y}/\partial a > 0$.

Finally, the last term in (23), ϵ^y, represents an intensive margin elasticity. It can be computed directly under CRRA preferences, $U(y) = y^{1-\sigma}/(1-\sigma)$ and $u(c) = y^{1-\varsigma}/(1-\varsigma)$:

$$\epsilon_j^y(a, e) = \begin{cases} \frac{1-\varsigma p\hat{y}\frac{\partial c^0/c^0}{\partial a}}{\sigma-\varsigma p\hat{y}\frac{\partial c^0/c^0}{\partial a}} \in [\frac{1}{\sigma}, 1) & \text{if unconstrained} \\ 1 & \text{if constrained.} \end{cases} \tag{26}$$

The intensive margin elasticity is bounded between $1/\sigma$ and 1. Further, it is a function of the household's marginal propensity to consume in terms of the general good, $\partial ln(c^0)/\partial a$. Since $\sigma > 1$, households with larger marginal propensities to consume are more price elastic regardless of search activity.

We can now return to the determination of the posted price distribution $F(p)$. Since firms are ex-ante identical at each point in time, it must be true that they are indifferent to posting any price in the support of $F(p)$. Following the logic of Butters (1977b), Varian (1980), and Burdett and Judd (1983), we can show the following properties of $F(p)$.

Lemma 2. *There exists an equilibrium posted price distribution $F(p)$ that has the following properties:*

i) F(p) has no mass points,

ii) F(p) has a finite and connected support $[\underline{p}, \bar{p}]$, where $\kappa < \underline{p} < \bar{p}$, and

iii) profits are equalized at any $p \in [\underline{p}, \bar{p}]$.

Our statement in Lemma 2 is weaker than is typical in frictional price dispersion models; we only claim existence instead of showing any equilibrium price distribution satisfies these properties. This arises because demand, conditional on being the lowest price draw of a household, $\varrho(p)\hat{y}(p)$, does not have a constant elasticity and its shape depends itself on the posted price distribution $F(p)$.[6] However, we posit and verify that there exists an equilibrium price distribution that induces a demand system is elasticity such that these properties hold.

To solve for the posted price distribution, first let $\bar{p}$ denote a solution to (21) in which $F(\bar{p}) = 1$. Then, we can write firm profits as

$$\Pi(p) = \bar{Y}(p)\left[1 - \bar{s}(p) + \bar{s}(p)\mu e^{-\mu F(p)}\right](p - \kappa), \tag{27}$$

where $\bar{Y}(p) = \sum_e \int_a \frac{\alpha}{N} \varrho(p; a, e)\hat{y}(p; a, e) d\Gamma(a, e, 1)$ is the aggregated demand of a firm posting price p, conditional on having the lowest price draw, and $\bar{s}(p) = \sum_e \int_a \frac{\alpha}{N}\left[\frac{\varrho(p;a,e)\hat{y}(p;a,e)}{\bar{Y}(p)}\right] s(a, e) d\Gamma(a, e)$ is the average, demand-weighted shopping effort of a firm posting price p. Since firms must make identical ex-ante profits, for any $p \in [\underline{p}, \bar{p}]$, profits must satisfy

$$Y(\bar{p})(\bar{p} - \kappa) = \bar{Y}(p)\left[1 - \bar{s}(p) + \bar{s}(p)\mu e^{-\mu F(p)}\right](p - \kappa) \tag{28}$$

Which implies $F(p)$ can be written as,

$$F(p) = -\frac{1}{\mu} ln\left(\frac{1}{\mu} - \frac{1}{\mu\bar{s}(p)}\left[1 - \frac{Y(\bar{p})(\bar{p} - \kappa)}{\hat{Y}(p)(p - \kappa)}\right]\right). \tag{29}$$

Finally, the lower bound for prices must satisfy $F(\bar{p}) = 0$, or

$$Y(\bar{p})(\bar{p} - \kappa) = \bar{Y}(\underline{p})\left[1 - (1 - \mu)\bar{s}(\underline{p})\right](\underline{p} - \kappa). \tag{30}$$

Equilibrium Definition A stationary equilibrium of this economy consists of value functions, $X(a, e)$, $W(a, e)$, and $V(a, e, s)$, household decision rules and choice probabilities, $c^0(a, e)$, $c^1(a, e)$, $s(a, e)$, $\hat{y}(p; a, e)$, $\varrho(p; a, e)$, firm pricing strategies, $F(p)$, and distributions of households over wealth, income, and idle or searching states, $\gamma(a, e, 0)$ and $\gamma(a, e, 1)$, such that

i) given the distribution $F(p)$, value functions and decision rules solve (7)-(15),

ii) given decision rules and the distribution of households, $F(p)$ solves (21) and (27)-(30), and

[6] For instance, in Burdett and Judd (1983) demand is indivisible; in Nord (2024) demand is continuous but the elasticity of demand is constant.

iii) the distribution of households solves (16) and (17).

4 Analytical results: inflation and shopping behavior

Before turning to the quantitative analysis, we characterize analytically how inflation affects the two margins of shopping behavior — the purchase probability ϱ and search effort s — and how the effects differ across the wealth distribution. To do so, we work in partial equilibrium, holding fixed the posted price distribution $F(p)$, and consider the problem of an individual household when the money growth rate rises permanently, so that the real return on liquid wealth falls, $r = -\pi$ (we set $\tau = 0$, the case in which money creation finances wasteful spending). To streamline the exposition, we also take $\lambda \to \infty$, so that the idle and searching states merge into a single value function $W(a, e)$; Appendix A shows that at the borrowing constraint the distinction between the two states has exactly no effect on the results below, so the simplification is for clarity rather than substance. The net purchase surplus is then

$$\hat{V}(p; a, e) = U(\hat{y}) + W(a - p\hat{y}, e) - W(a, e). \tag{31}$$

As shown in equations (12) and (14), both shopping margins respond to inflation through $\hat{V}$. Define $\Psi(a, e) \equiv \partial W(a, e)/\partial \pi$, the inflation sensitivity of the value function, and $\Psi_a \equiv \partial \Psi / \partial a$. Applying the envelope theorem to $\hat{y}$ (Lemma 3 in Appendix A),

$$\frac{\partial \hat{V}(p; a, e)}{\partial \pi} = \Psi(a - p\hat{y}, e) - \Psi(a, e) = -\int_0^{p\hat{y}} \Psi_a(a - t, e)\, dt, \tag{32}$$

so that $\operatorname{sign}(\partial \hat{V}/\partial \pi) = -\operatorname{sign}(\Psi_a)$. If the value function is more sensitive to inflation at higher wealth ($\Psi_a < 0$), then purchasing — which sheds $p\hat{y}$ units of money — becomes relatively more attractive as inflation rises. Both margins inherit this sign: differentiating (12) and (14),

$$\frac{\partial \varrho(p; a, e)}{\partial \pi} = \eta\, \varrho(1 - \varrho) \frac{\partial \hat{V}(p; a, e)}{\partial \pi}, \qquad \delta''(s) \frac{\partial s(a, e)}{\partial \pi} = \alpha \int_p \varrho(p; a, e) \frac{\partial \hat{V}(p; a, e)}{\partial \pi}\, \omega(p)\, dF(p), \tag{33}$$

where $\omega(p) \equiv \mu e^{-\mu F(p)} - 1$ is the marginal effect of effort on the probability of drawing price p.

What determines the sign of Ψ_a? Differentiating the household's HJB equation with respect to π and applying the envelope theorem yields

$$\rho\,\Psi(a,e) = \underbrace{-a\,W_a(a,e)}_{\text{direct erosion}} + \underbrace{\alpha\int_p \varrho(p;a,e)\left[\Psi(a-p\hat{y},e)-\Psi(a,e)\right]dG(p;s)}_{\text{search option value}} + \underbrace{\Psi_a(a,e)\,\dot{a}(a,e)}_{\text{wealth drift}} + \mathcal{P}_{e,e'}\Psi. \quad (34)$$

Three forces govern Ψ and its slope. The *direct erosion* term reduces lifetime utility in proportion to the money balances held, valued at the marginal value of wealth. The *search option value* term captures how inflation is capitalized differently before and after a purchase: buying moves the household to a lower wealth state whose value is differentially exposed to inflation. The *wealth drift* term reflects the household's saving dynamics. In what follows, we show that these forces operate differentially across the wealth distribution, generating heterogenous responses to inflation.

Low wealth and income households Consider a household at the borrowing constraint and with the lowest income, $(a,e) = (\underline{a},\underline{e})$. At the borrowing constraint, households have no resources for lumpy purchases ($\hat{y} \to 0$), so the search option value force shuts down exactly. The wealth drift force need not shut down: the state constraint delivers only $\dot{a}(\underline{a},e) \geq 0$, and whether households at the constraint consume their entire inflow ($\dot{a} = 0$) or accumulate wealth ($\dot{a} > 0$) is an equilibrium outcome.[7] In the lowest income state, however, the drift is bounded above: positive consumption implies $\dot{a}(\underline{a},\underline{e}) \leq \underline{e} - \pi\underline{a}$. The direct erosion term therefore dominates provided the lowest income state is sufficiently low and persistent.

Proposition 1. *Fix $F(p)$ and suppose $\underline{a} \leq 0$. Let $\underline{e}$ denote the lowest income state, and suppose $\underline{e}$ is low enough and sufficiently persistent, in the precise sense of condition (44) in Appendix A. Then $\Psi_a(\underline{a},\underline{e}) < 0$, and there exists $\bar{\varepsilon} > 0$ such that for all $a \in (\underline{a}, \underline{a}+\bar{\varepsilon})$:*

i) $\partial\varrho(p;a,\underline{e})/\partial\pi > 0$ for every p in the support of F, and

ii) $\partial s(a,\underline{e})/\partial\pi < 0$.

The proof is in Appendix A. The economics is the following. A low income household near the borrowing constraint holds essentially all of its spending capacity in money, and it must

[7] Contrary to a standard continuous-time incomplete markets model as in Achdou et al. (2022), households with the lowest income state in this model may accumulate wealth to finance lumpy purchases.

carry those balances for as long as it keeps searching; completing a purchase sheds the exposure. Higher inflation therefore erodes the value of continued search by more than it erodes the post-purchase value, so the household becomes less picky at every price: ϱ rises. In turn, a household that is willing to buy at almost any price it encounters gains less from sampling more prices, so search effort falls. Both responses push the average price paid by low wealth households up. This is the sense in which our model inherits the logic of Diamond (1971): the inflation tax on money balances acts like a cost of continued search, and it bears most heavily on households whose wealth is mostly liquid. The lowest income state needs to be low in order to keep the drift $\dot{a}$ low, and persistent enough since otherwise the wealth drift force may overturn the effect of direct erosion at the constraint. The precise bound, condition (44), is stated in the appendix.

Higher wealth and income households Away from the borrowing constraint and lowest productivity state, all three forces in (34) are active and the sign of Ψ_a is theoretically ambiguous: the direct erosion effect weakens as the marginal value of wealth falls, while the search option value effect — which is absent at the constraint — pushes in the opposite direction. As $a \to \infty$ under CRRA preferences with $\sigma > 1$, both the erosion and option value effects vanish and the partial-equilibrium effects of inflation on shopping behavior disappear altogether. Table 1 summarizes.

Wealth level	Ψ_a	$\partial \hat{V}/\partial \pi$	$\partial \varrho/\partial \pi$	$\partial s/\partial \pi$
$a \approx \underline{a}$, $e = \underline{e}$ (Proposition 1)	< 0	> 0	> 0	< 0
a intermediate	ambiguous	ambiguous	ambiguous	ambiguous
$a \to \infty$	driven by wealth drift	$\to 0$	$\to 0$	$\to 0$

Table 1: Partial-equilibrium effects of inflation on shopping behavior, by wealth.

The analysis thus pins down the bottom of the wealth distribution and bounds the top: strong responses of ϱ and s to inflation are a low-wealth phenomenon, and since the partial-equilibrium effect vanishes at high wealth, the increase in search effort at the very top of the distribution that we document in Section 7 largely reflects the general-equilibrium adjustment of the price distribution $F(p)$ rather than a direct effect of the inflation tax. We now turn to a quantitative evaluation of the model.

The combined effect of these forces is that the sensitivity of $\hat{V}$ to inflation — and thus the response of ϱ and s — diminishes with wealth. The analytical characterization does not yield a determinate sign for Ψ_a at general wealth levels, as the balance among the three forces depends on parameters. What the analysis does establish is a clear ordering: the inflation-induced change in $\hat{V}$ is largest in magnitude for low-wealth households, where the direct erosion effects dominates. The general equilibrium effects, which incorporate the endogenous responses of the posted price distribution $F(p)$ and the equilibrium distribution of households $\gamma(a, e, \cdot)$ to changes in aggregate shopping behavior, further shape the distributional pattern, which we now turn to analyzing quantitatively.

5 Quantitative Results

For the quantitative analysis, we extend the model in two important dimensions to capture realistic dispersion in income and wealth following Kaplan et al. (2018). We detail the equilibrium equations in Appendix **??**. First, we assume in addition to money that households can hold an illiquid asset. We denote holdings of the illiquid assets as b and the holdings of the liquid asset (money) as a. Their steady-state returns are r_b and r_a, respectively where $r_a = -\dot{\phi}_t/\phi_t$ as before. The real supply of illiquid assets is constant and denoted B^s. Households face a transaction cost $\chi(d, b)$ for depositing a flow d into their illiquid account, given by

$$\chi(d, b) = \chi_0|d| + \frac{\chi_1}{2}\left(\frac{d}{b}\right)^2 b,$$

Further, in the event of a lumpy expenditure shock we assume that households can only access their illiquid assets.[8] The market clearing condition for illiquid assets is $\int b\gamma(a, b, e)db = B^s$. As our baseline, we assume that money creation finances wasteful government spending, so the budget constraint of the consolidated government is simply $g_t = \dot{M}_t$.[9]

Second, we assume a households labor productivity is determined by the sum of two inde-

[8]This simplifies the computation of decision rules substantially and captures that in a very short period of time, there may be an infinite transaction cost of converting illiquid assets into liquid assets, e.g. it is difficult to sell or collateralize a home quickly.

[9]Our illiquid asset is then explicitly not a government bond but represents private real assets (e.g. corporate debt and equity, residential real estate, etc.).

pendent jump-diffusion processes.

$$ln(e_{it}) = z^{\rho}_{it} + z^{\tau}_{it}, \tag{35}$$

where z^j_{it} is given by the process

$$dz^j_{it} = -\beta^j z^j_{it} dt + dJ^j_{it}, \tag{36}$$

for each $j \in \{\rho, \tau\}$. The jump component dJ^j_{it} is a Poisson process with rate γ^j. The innovation after a jump is normally distributed mean zero with standard deviation σ^j. As we discuss in the next section, allowing for two independent components gives the calibration the flexibility to match moments from the empirical earnings distribution and aligns well with the notion of having both a persistent and transitory component in idiosyncratic earnings risk.

5.1 Calibration procedure

We set a unit of time in the model to represent one month. Preferences for flow and lumpy consumption are given by

$$u(c) = \frac{c^{1-\varsigma}}{1-\varsigma}, \text{ and } \qquad U(y) = \Sigma\left[\frac{(y+\underline{y})^{1-\sigma}}{1-\sigma} - \frac{(\underline{y})^{1-\sigma}}{1-\sigma}\right], \tag{37}$$

where $\varsigma, \sigma > 1$, $\Sigma > 0$, and $\underline{y} \approx 0^{+}$, which guarantees that $U(0) = 0$. We set $\varsigma = 2$ but allow both the level and curvature of preferences for lumpy consumption to be determined by the calibration. The shopping effort function is $\delta(s) = \Delta_0 \frac{s^{\Delta_1}}{1-s}$, where $\Delta_0 > 0$ and $\Delta_1 > 1$. We set the measure of firms $N = 1$.

Collectively, our functional form assumptions imply we have 21 parameters to determine: those for the labor productivity process $(\beta^j, \gamma^j, \sigma^j)_{j\in\{\rho,\tau\}}$, preference parameters $(\rho, \sigma, \Sigma, \lambda, \eta)$, search cost parameters (Δ_0, Δ_1), search frictions α, transaction cost parameters (χ_0, χ_1), illiquid asset supply B^s, the borrowing constraint for liquid wealth $\underline{a}$, the mean of active searchers price draws μ, marginal cost κ, and the money growth (or, steady-state inflation) rate π^m.

Calibration of the income process We calibrate the income process separately to match moments from the distribution of annual log earnings and 1-year log earnings changes from the US in 2015 from the Global Repository of Income Dynamics (GRID) database.[10] For annual log earnings, we target the variance, as well as six percentiles that coincide with our measures of income percentiles in the Kilts-NielsenIQ data. In addition, we target the mean, variance, and kurtosis of the distribution of 1-year log earnings changes. We choose seven grid points for the persistent shock and three grid points for the transitory shock.

Moment	**Model**	**Data**	**Diff.**
Annual log earnings			
Variance	0.8908	0.9369	−0.0462
Skewness	−0.0981	−0.393	0.2949
Kurtosis	1.639	0.6387	1.0003
Percentile 9.2	−1.5017	−1.3984	−0.1032
Percentile 26.8	−0.3485	−0.4868	0.1382
Percentile 42.2	−0.1514	−0.0835	−0.0678
Percentile 58.0	−0.0343	0.2473	−0.2817
Percentile 73.0	0.1598	0.5585	−0.3987
Percentile 89.4	1.2559	1.0591	0.1968
1-year change in annual log earnings			
Mean	0.0093	0.0585	−0.0492
Variance	0.2708	0.294	−0.0232
Skewness	−0.039	−0.9278	0.8888
Kurtosis	9.6797	9.8369	−0.1572
Income Parameters	γ^j	β^j	σ^j
Persistent ($j = \rho$)	0.0130	0.0076	1.1987
Transitory ($j = \tau$)	0.1810	0.7676	1.3231

Table 2: Calibration results for the income process.

Table 2 reports the calibration results for the income process, including the parameters and both targeted and untargeted moments. Figure 4 illustrates the fit of the model compared to the empirical distributions. First in terms of parameters, the process e^{ρ}_{it} has a low rate of mean reversion ($\beta^{\rho} = 0.0076$) and a low rate of jumps ($\gamma^{\rho} = 0.0130$) while the process e^{τ}_{it} has a relatively high rate of mean reversion ($\beta^{\tau} = 0.7676$) and a high rate of jumps ($\gamma^{\tau} = 0.1810$). As in Kaplan et al. (2018), this process allows for a relatively persistent and transitory components in idiosyncratic earnings. The procedure matches the targeted empirical moments well.

[10]See Guvenen, Pistaferri, and Violante (2022a), or their website `https://grid-database.org`, in order to access the data and Guvenen, Pistaferri, and Violante (2022b) for details about the variable construction and methodology.

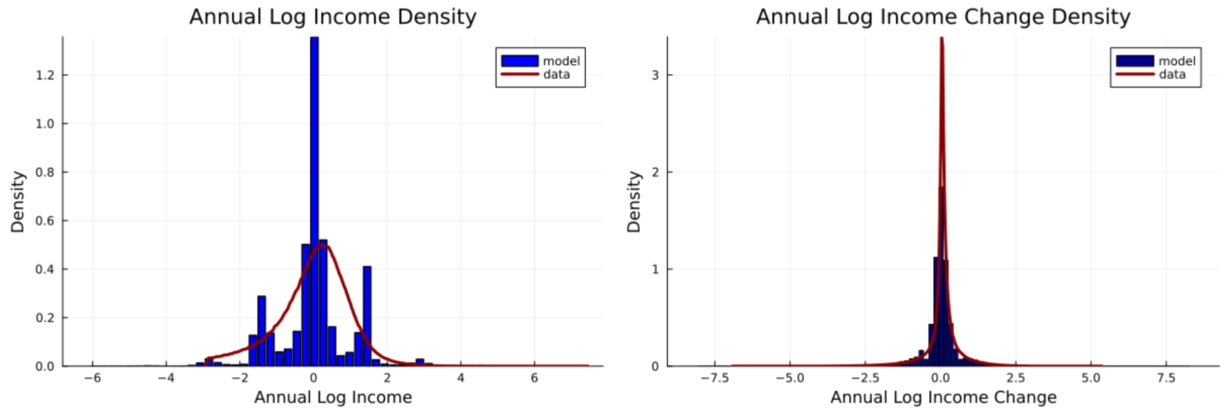


Figure 4: Fit of the model compared to the empirical distributions for the income process.

In terms of untargeted earnings moments, the model underpredicts the skewness and overpredicts the kurtosis of annual log earnings, and also underpredicts the skewness of 1-year log earnings changes. However, as shown in Figure 4, this is largely driven by the limitation on the number of grid points used in the numerical procedure.

Externally calibrated parameters We normalize firm's marginal cost to $\kappa = 0.75$ and set the average price draw of active searchers to $\mu = 3$. While the latter is not a normalization, its choice does not greatly affect the outcome of the calibration procedure or the numerical results. We set the money growth rate to 2 percent per year, $\pi^m = 0.02/12$, and set the supply of illiquid assets to target an annualized liquidity premium of 6.2%. Since our model does not feature aggregate risk, we measure the liquidity premium following Bethune and Rocheteau (2023) by equating to the real user cost of holding zero maturity assets (MZM as defined by Anderson and Jones (2011).[11] The real user cost equals the interest rate spread between the real return on MZM and the maximum real return on a set of short-term money market accounts and the Baa corporate bond yield. The average user cost over 1998-2013 was 6.2%.

We calibrate the remaining parameters jointly to capture three sets of moments: the relationship between markups paid and annual income, and the relationship between measures of shopping effort and annual income, and moments from the distribution of wealth to annual income.

[11]See FRED series OCMZMA.

Evidence on income and prices/markups paid We use the NielsenIQ Consumer Panel, as described in Section 2. The combination of consistent product identifiers and household-level income information allows us to compare prices paid for identical goods (at the barcode level) across income groups. In what follows, we construct income-specific unit prices and estimate conditional price differences across income bins. Nielsen reports household income in discrete nominal brackets. To align with the model and with the income groups used to compute shopping effort from the American Time Use Survey (ATUS) as described below, we aggregate these brackets into six income bins that partition the household income distribution into broad percentile intervals.[12]

For each year, we aggregate purchases to the UPC $\times$ income-bin $\times$ DMA $\times$ demographic-cell level. Within each cell, we compute the expenditure-weighted unit price as total expenditure divided by total quantity purchased and take logs. This aggregation ensures that comparisons are made across households purchasing the same product within the same geographic market.

We estimate the following regression separately for each year:

$$\log p_{ugdt} = \sum_{k=1}^{6} \beta_{k,t} \, \mathbb{1}\{g = k\} + X_g' \gamma_t + \alpha_u + \delta_d + \varepsilon_{ugdt},$$

where $\log p_{ugdt}$ denotes the log unit price of UPC u purchased by households in income bin g in DMA d during year t. The vector X_g includes indicators for household age group, household size, and race. The fixed effects α_u absorb permanent product-level price differences, while δ_d capture geographic price-level variation across DMAs. Standard errors are clustered at the DMA level.

This specification compares prices paid by different income groups for the same product within the same market, controlling for observable demographic characteristics. The coefficients $\beta_{k,t}$ therefore measure conditional price premia associated with income. We normalize the third income bin as the reference group, so that each coefficient $\beta_{k,t}$ can be interpreted as the log price difference relative to the middle-income group. To construct calibration targets, we average these estimated gradients across 2010–2019, using total expenditure as weights across years. The resulting vector of log price differences across income bins constitutes the income–price moment

[12] The mapping from Nielsen's nominal income brackets to percentile-based bins is constructed using sampling weights so that each bin corresponds to a fixed share of the weighted income distribution in the pre-period.

targeted by the model.

We also compare the relationship of income to prices paid to that of Sangani (2023) who also uses the NielsenIQ data combined with measures firms variable costs to construct the relationship between income and markups paid (see Sangani, 2023, Figure 1). They find a stronger income gradient in markups relative to price paid, suggesting that higher income households also shop at retailers with lower marginal costs. Since our model does not feature any heterogeneity in firms' marginal costs, we choose to target the relationship from Sangani (2023) but our results are qualitatively similar when we target the relationship between income and prices paid.

Evidence on income and shopping effort To discipline heterogeneity in shopping effort, we use a measure of the time a household spends shopping reported in the The American Time Use Survey (ATUS), conducted by the U.S. Census Bureau for the Bureau of Labor Statistics. The ATUS provides nationally representative 24-hour time diaries in which respondents report all activities undertaken on the previous day in detailed time intervals. Reported activities are coded into a hierarchical classification system comprising several hundred time-use categories, and the survey is linked to demographic and income information from the Current Population Survey (CPS), allowing us to relate time allocation to household characteristics.[13] ATUS data have been used in the consumer search literature to study time allocation and price-related behavior, for example in Aguiar and Hurst (2007), Aguiar, Hurst, and Karabarbounis (2013), and Kaplan and Menzio (2016).

We focus on the period 2010-2019 and construct shopping time from two detailed ATUS activity categories. Specifically, we sum minutes coded as *Grocery Shopping* (a subcategory under the major category *Consumer Purchases*) and minutes coded as travel to/from grocery stores (a subcategory under *Travel Related to Consumer Purchases*). We exclude other shopping- and travel-related categories such as non-grocery retail and service-related purchases so that the time measure corresponds as closely as possible to the goods covered in the NielsenIQ purchase data.

Income groups are constructed to match the six-bin structure used from the Nielsen moments.[14] Income heterogeneity is expressed in percentile terms using weighted population shares

[13]The ATUS diary is linked to CPS background information collected two to five months prior to the diary interview. See the ATUS data dictionary for details.

[14]The income brackets in ATUS and NielsenIQ are nearly identical after aggregation to six bins. The only difference

in the pre-period, mirroring the construction used for the price moments. For respondent i in income bin g, residing in state s in year t, we estimate

$$\text{Time}_{igst} = \sum_{k=1}^{6} \eta_k \, \mathbb{1}\{g = k\} + Z_i' \lambda + \delta_s + \kappa_t + \varepsilon_{igst}, \tag{38}$$

where Time_{igst} denotes grocery shopping minutes on the diary day. Because the ATUS identifies respondents at the state level rather than by local retail markets, geographic variation is absorbed using state fixed effects δ_s based on the respondent's state FIPS code. The control vector Z_i includes indicators for respondent age group, household size, race, and sex. Including sex is particularly important in time-use data, as household shopping time may reflect gender differences in the division of labor within households. We also include year fixed effects κ_t and estimate the regression using ATUS sampling weights.

This specification identifies income-related differences in shopping time from within-state, within-year variation, conditional on observable demographics. The coefficients η_k measure income gradients in shopping time. After re-centering the estimates relative to the middle-income group, we obtain a vector of income-specific shopping-time deviations that serves as the shopping-effort moment in the calibration.

Wealth-to-income and average markup targets We also target moments from the distribution of wealth, following Kaplan et al. (2018). Specifically, we target four moments: i) an average wealth to income ratio of 3.43, ii) an average liquid wealth to income ratio of 0.36, iii) a fraction of households with negative liquid wealth of 0.15, and iv) a fraction of "wealthy hand-to-mouth" households — those with $a < 0$ and $b > 0$ — of 0.20. In addition, we target an average markup of 50%. Hall (2018) finds a typical markup of about 1.3 (prices roughly 30% above marginal cost) using U.S. KLEMS productivity data, and Loecker, Eeckhout, and Unger (2020) document a rise in average firm-level markups to about 67% above marginal cost in recent years. Generally, retail and consumer-facing markups tend to be larger than economy-wide averages. For instance, Atalay, Frost, Sorensen, Sullivan, and Zhu (2023) document substantial and rising markups in U.S. grocery markets using scanner data, with the median product markup rising by about 10

concerns the boundary between the fourth and fifth bins, which is $75,000 in ATUS and $70,000 in NielsenIQ; all other bin thresholds coincide. We adopt the closest feasible alignment across datasets to maintain comparability.

percentage points between 2006 and 2018.

5.2 Model fit of targeted moments

Tables 3 and 4 report the fit of the model to the targeted moments and the resulting parameters, respectively. In Figure 5, we illustrate the relationship between annual income and prices paid (left panel) and annual income and shopping effort (right panel). The data shows that as household income increases the average markup paid also increases. For instance, households in the highest income percentile bin ($79^{th} - 100^{th}$ percentile) pay on average 8.2% higher markups relative to households in the lowest income percentile bin ($0^{th} - 18^{th}$ percentile). The model captures this increasing relationship well. It slightly underpredicts markups paid for households in the lowest income percentile bin and slightly overpredicts markups paid for households in the highest income percentile bin.

The empirical relationship between income and shopping effort (right panel of Figure 5) is hump-shaped; low- and high-income households have a lower shopping effort compared to middle income households. Our calibrated model can also fit a hump-shaped relationship, although the gradient is flatter in our model compared to the data.

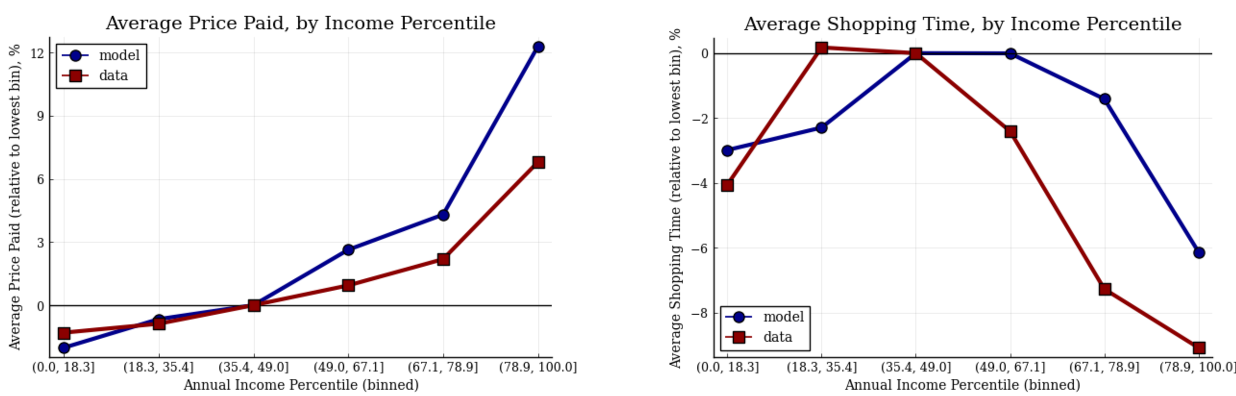


Figure 5: Relationship between annual income and markups paid and shopping effort.

Finally, the model slightly overpredicts total and liquid wealth to income, but does well in terms of matching the fraction of wealthy hand-to-mouth households. It overpredicts the fraction of households with negative liquid wealth as well as the markup.

In terms of parameters, the calibration procedure results in a rate of lumpy expenditure

Moment	**Model**	**Data**	**Error (% squared)**
Average markups paid (% deviation from middle-income), by income percentile			
Income percentile (0.0, 18.3]	−2.99	−1.30	0.03
Income percentile (18.3, 35.4]	−0.14	−0.88	0.01
Income percentile (49.0, 67.1]	2.59	0.94	0.03
Income percentile (67.1, 78.9]	4.48	2.19	0.05
Income percentile (78.9, 100.0]	11.53	6.80	0.22
Shopping effort (% deviation from middle-income), income percentile			
Income percentile (0.0, 18.3]	−0.54	−4.06	0.12
Income percentile (18.3, 35.4]	0.03	0.18	0.00
Income percentile (49.0, 67.1]	−0.43	−2.41	0.04
Income percentile (67.1, 78.9]	−1.89	−7.28	0.29
Income percentile (78.9, 100.0]	−7.42	−9.09	0.03
Wealth-to-income and markups			
Average wealth to income	5.21	3.18	0.41
Average liquid wealth to income	0.48	0.26	0.71
Fraction wealthy HtM	0.16	0.20	0.03
Fraction negative liquid wealth	0.35	0.15	1.77
Average markup	2.72	1.30	1.19

Table 3: Model vs. data: targeted moments comparison.

shocks of $\lambda = 0.54$, or shocks arriving about every two months. In the event of an expenditure shock the arrival rate of price draws occurs relatively quickly, $\alpha = 2.5$, or about every 1.6 weeks. Households have slightly higher risk-aversion for lumpy expenditures relative to flow expenditures, $\sigma = 2.3$ compared to $\sigma = 2.0$, however the level of utility is higher, $\Sigma = 2.4$.

6 Cross-sectional heterogeneity in shopping behavior and markups

In this section, we analyze the calibrated (low-inflation) steady state. On the household side, we show that heterogeneity in wealth and income incudes heterogeneity in shopping behavior and markups paid; households sort across retailers charging different markups. We decompose this sorting behavior into the effects that are driven by differences in shopping effort, s, purchase probabilities, ϱ, and the intensive margin of demand, $\hat{y}$. On the firm side, we illustrate the equilibrium distribution of markups and how the elasticity of demand at each is composed of the three effects shown in equation (23).

Description	Parameter	Value
Preference Parameters		
Discount rate (annual)	ρ^{annual}	0.05
Level of $U(y)$	Σ	2.36
Curvature of $U(y)$	σ	2.27
Lumpy expenditure arrival rate	λ	0.54
Variance of ζ	η	4.35
Technology Parameters		
Search frictions	α	2.57
Search cost level	Δ_0	0.02
Search cost curvature	Δ_1	1.57
Adjustment cost parameter	χ_0	0.014
Adjustment cost parameter	χ_1	2.00
Mean of active searchers price draws	μ	3.00
Borrowing constraint	$a_{\min}$	-5.04
Firms's marginal cost	κ	0.75
Policy Parameters		
Money growth rate (annual)	$\pi^{m,annual}$	0.02
Supply of illiquid assets	B^s	35.22

Table 4: Calibrated parameters

Household heterogeneity in shopping behavior and the elasticity of demand Figure 6 shows the relationship between household wealth (measured in percentiles) and shopping effort. The left panel does so for total wealth while the right panel focuses on liquid wealth. The solid lines in each panel show the average for that percentile of wealth (averaging over portfolios and income) while the dashed lines represent the minimum and maximum within the wealth percentile.

As a household's total or liquid wealth increase, their shopping effort tends to decrease. An average household in the bottom of the wealth distribution has a shopping effort about 33% higher than an average household at the top of the wealth distribution. Given the equilibrium distribution of wealth this maps into about 25% more price draws for low wealth compared to the high wealth households. For some households, shopping effort decreases rapidly as wealth accumulates, in particular for those with enough liquid wealth. The figure in the right panel shows that some households in the top of the liquid wealth distribution have shopping effort about 60% lower than those at the bottom of the distribution — or about 60% fewer price draws.

Shopping effort, however, is only one aspect that determines the expected price a household eventually pays. Conditional on drawing a price, a household still has the choice of whether to

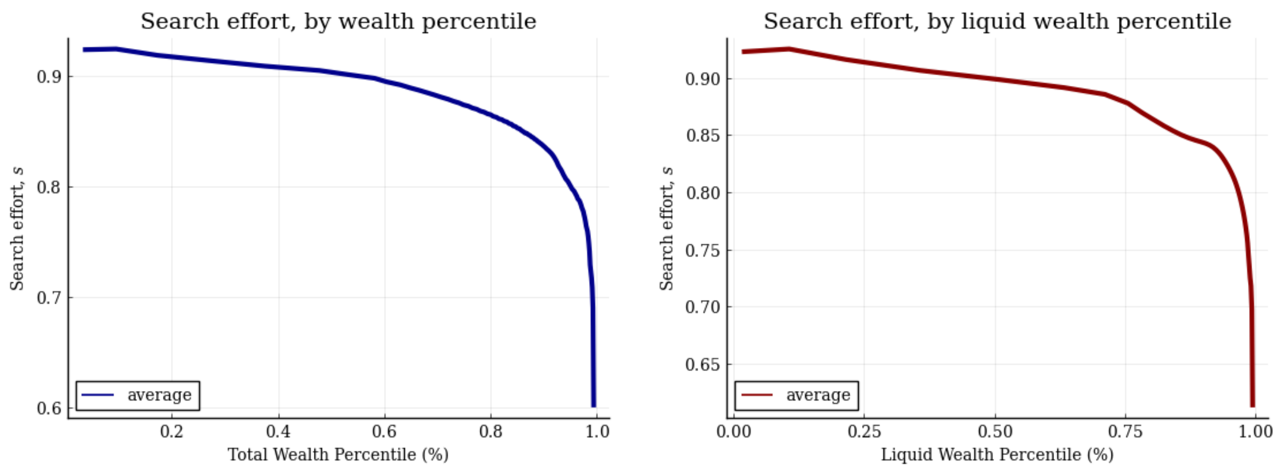


Figure 6: Shopping effort, by wealth.

make a purchase or to keep shopping, as captured by their purchase probability, ϱ. Figure 7 illustrates the relationship between total and liquid wealth and a household's expected probability of making a purchase.

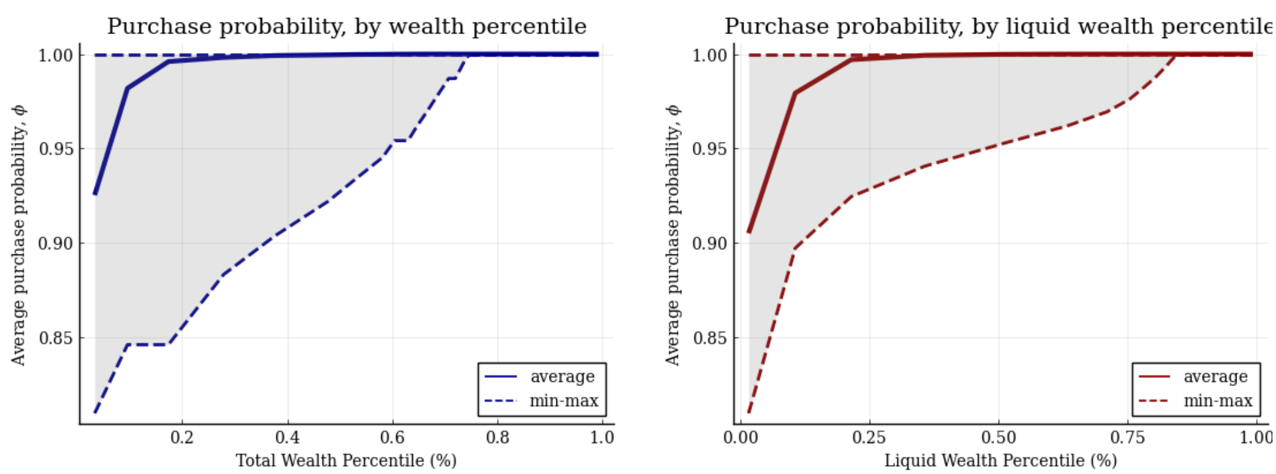


Figure 7: Purchase probabilities, by wealth.

The relationship is upward sloping for both total and illiquid wealth. Note, there are wide min-max bands, in particular for households that have low wealth. This implies that income is an important determinant in a household's choice of making a purchase. Further, the relationship in Figure 7 already incorporates the fact that lower wealth households already put in more shopping effort, implying the best price they draw is lower than for higher wealth households. Ceteris paribus, the lower price a household faces leads to a *higher* probability of making a purchase.

Despite this, the equilibrium relationship between wealth and shopping effort is upward sloping.

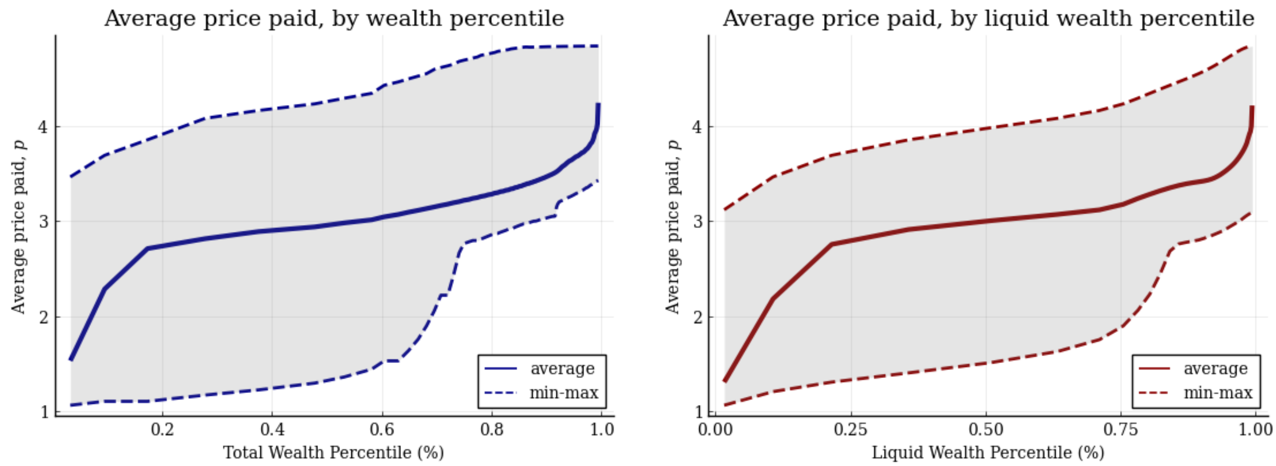


Figure 8: Average prices paid, by wealth.

The patterns in Figures 6 and 7 directly imply the equilibrium relationship between wealth and the average price (or markup) paid, shown in Figure 8. As either total or liquid wealth increase, the expected price a household pays for an identical good increases. Households at the top of the distribution of wealth pay on average 60% higher prices compared to those at the bottom of the distribution. Compared to the gradient with respect to income, in Figure 5, the gradient with respect to wealth is significantly higher.

In Figure 9, we decompose the relationship between wealth and prices paid into two components: i) that being driven by heterogeneity in shopping effort s, and ii) that being driven by heterogeneity in the probability of making a purchase, ϱ. We reproduce the equilibrium relationship as the blue, solid curve. The shopping effort effect is shown as the dashed, red curve and the purchase probability effect is shown as the dash-dotted, green curve. If we only considered heterogeneity in shopping effort, the gradient of prices paid to wealth is only weakly positive — the highest wealth household pays only about 10% more than the lowest wealth household, on average. However, allowing for heterogeneity in purchase probabilities drives almost entirely the equilibrium relationship between wealth and prices.

Why do we find such a strong effect of purchase probability on prices paid? In models following Burdett and Judd (1983) that do not allow for an option value of continued search, households that have high shopping effort may still get unlucky and draw only prices at the top

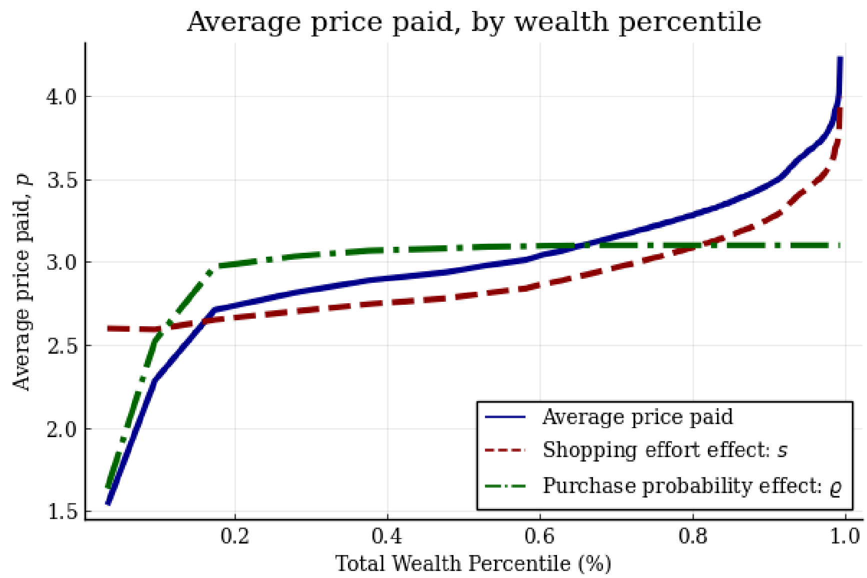


Figure 9: Average prices paid, by wealth.

of the equilibrium price distribution. Since these households have no option value of continued search, they always make a purchase at these higher prices. Households in our model, however, do not have to — if a high shopping effort household gets unlucky and draws only high prices they decide not to purchase and keep shopping. Figure 9 illustrates that removing the option value of search, greatly reduces household sorting by price.

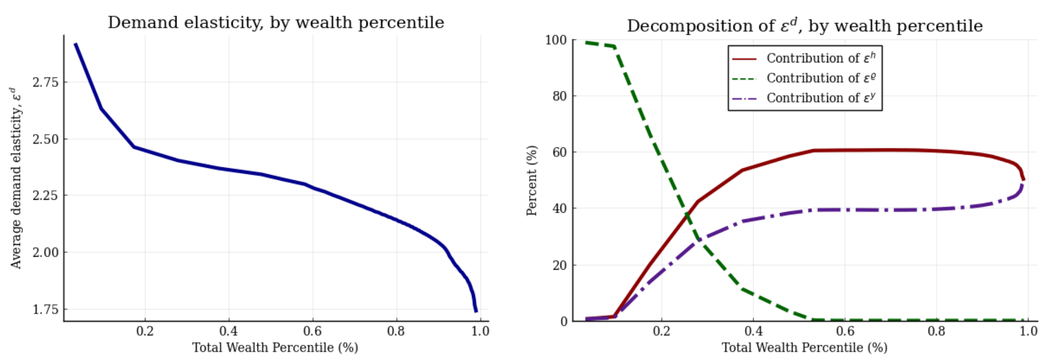


Figure 10: The elasticity of demand at the average price, by wealth.

In addition to prices paid, heterogeneity in shopping behavior also drives differences in the shape of households' demand functions. To illustrate this, the left panel of Figure 10 plots the average elasticity of demand for lumpy consumption by total wealth, for the average price. Low wealth households tend to have a high elasticity of demand at a given price, compared to high wealth households. The right panel of 10 decomposes household demand elasticities as in equation (23) into three components: i) the change in demand due to households switching to lower price firms in their choice set (solid-red line) ii) the change in demand due to households deciding to continue searching (dashed-green line), and iii) the change in demand along the intensive margin (dash-dotted violet line). The figure shows that the high elasticity of demand of low wealth households at a given price is primarily driven by their outside option of continued search. At the average price, these households respond strongly to price changes by taking longer to search. For high wealth households, the elasticity of demand is mostly driven by their intensive margin (they respond to price increases by purchasing fewer goods), and to a lesser extent by switching to lower price stores.

The equilibrium distribution of markups and its determinants We now turn to the firm side. Figure 11 illustrates the equilibrium density of posted prices, $f(p)$ where the scale is converted into a markup, p/κ. As is usual in models following Burdett and Judd (1983), there is a large mass of retailers posting low markups in order to take advantage of selling high volume. However, a unique attribute of our equilibrium distribution is that it is hump-shaped. This is driven by the heterogeneity in customers that firms face. Since there are few households with elasticity high enough to warrant markups at the bottom of the support, there are fewer firms posting these prices. In other words, the total sales volume at this price is relatively low, given the customer base. There is a large mass of retailers posting just above the lowest markup. The density of retailers decreases as the markup increases. There exist some retailers that charge 'exorbitant' markups in equilibrium — above 400% — but these have a low measure.

In Figure 12, we illustrate the importance of which factors determine a retailers markup across the support of the distribution. The left panel maps the markup to the elasticity of demand while the right panel decomposes the elasticity of demand according to equations (22) and (23). The solid-blue curve shows the contribution of ϵ^h, the dashed-green curve shows the contribution of ϵ^ϱ, while the dash-dotted-violet curve shows the contribution of ϵ^y.

First, consider low-markup retailers. The largest contributor to their low markup strategy is competition with other low markup retailers, of which there is a large measure. Customers at these retailers have many other options of finding similar retailers with low markups. The

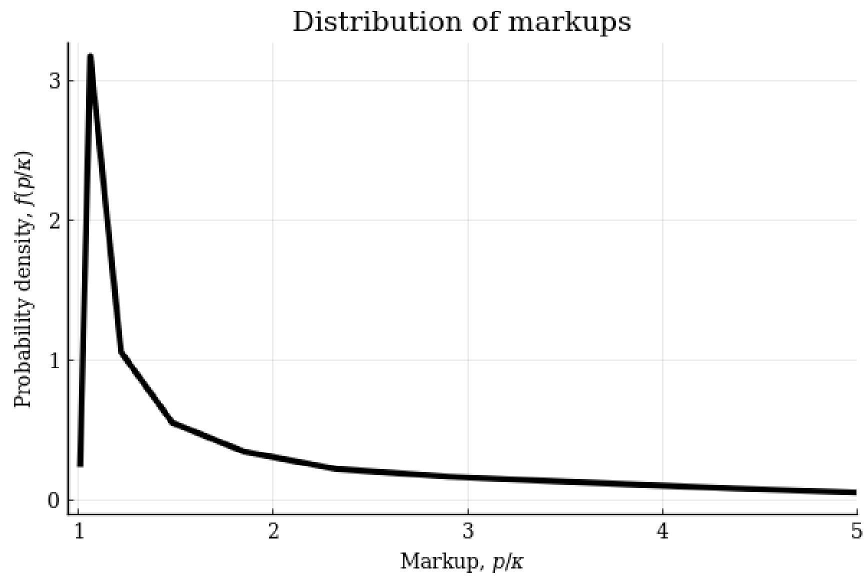


Figure 11: The equilibrium distribution of markups, in low-inflation steady state.

contribution of ϵ^{ϱ} is low since these customer have a high purchase probability and there is little gain in searching more in the future for lower prices. Further, the intensive margin of demand is relatively high at low markups so ϵ^{y} contributes little as well.

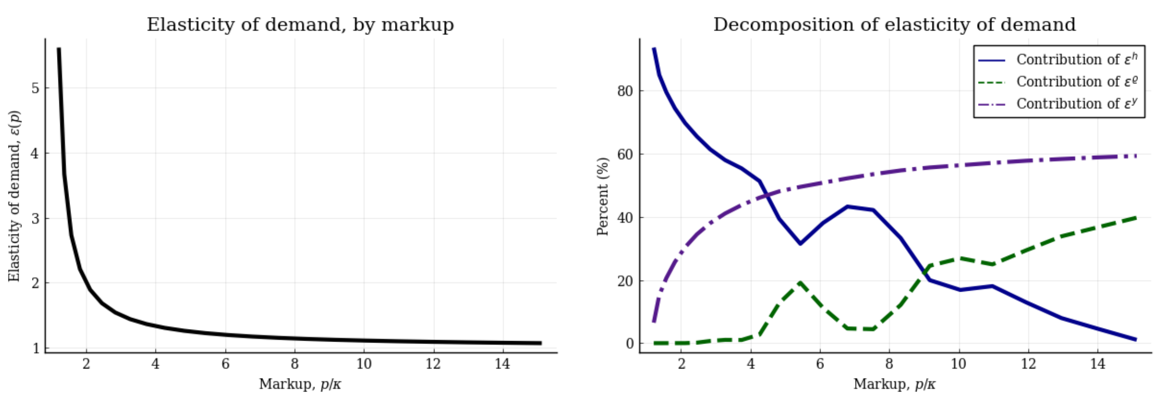


Figure 12: The equilibrium determinants of markups.

Markups at the top of the distribution, however, are largely determined by the elasticity along the intensive margin and the outside option of customers waiting to purchase in the future. Consider a retailer even at the top of the distribution, posting $\bar{p}$. They understand that their customers can still purchase fewer goods in response to a price increase, or wait to purchase and continue searching for lower prices in the future.

7 The response of shopping behavior and markups to inflation

Our empirical results in Section 2 provided evidence that around the inflationary episode in the US, customers systematically adjusted their shopping behaviors. The share of trips to low price stores of all customers increased nearly 30%. Further, high-income households increased their shopping in low-priced stores more, relative to low-income customers. In this section, we investigate how the equilibrium sorting of households across retailers is affected by an increase in the aggregate inflation rate.

We consider two (steady-state) inflation regimes: a 5% inflation rate, $\pi^{m,med} = 0.05$ and a 10% inflation rate, $\pi^{m,high} = 0.10$. We compare the steady-state equilibrium outcomes to the low, 2% inflation benchmark described above, $\pi^{m,low} = 0.02$. For now, we maintain the assumption that money creation is used to solely finance wasteful spending.

Inflation and prices paid, by wealth Figure 13 illustrates how inflation changes shopping behavior by the value of total wealth. The left panel shows the response of shopping effort, s, whlie the right panel shows the response in the probability of making a purchase, ϱ. We compute the average change, conditional on a household being in a particular percentile of the total wealth distribution in the low inflation steady state.

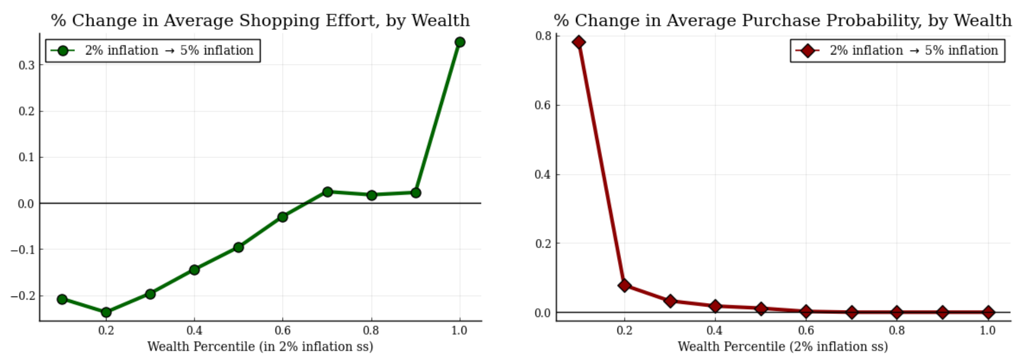


Figure 13: The effect of inflation on shopping effort and purchase probability, by total wealth.

The left panel shows that moderate inflation has opposite effects on shopping effort across the wealth distribution. Households in the bottom two-thirds of the distribution *reduce* their shopping effort, while households at the top increase it, with the largest increase concentrated at the very top of the distribution. The right panel shows that inflation also increases the probability of making a purchase. By increasing the cost of continued search, inflation lowers households' outside options. This response is heavily concentrated among the lowest wealth households, whose purchase probability rises the most, while households in the top half of the distribution hardly respond. These responses line up with the partial-equilibrium results of Section 4: consistent with Proposition 1, the rise in the purchase probability and the fall in search effort are concentrated among low wealth households, while the increase in search effort at the top of the distribution reflects the general-equilibrium adjustment of the price distribution.

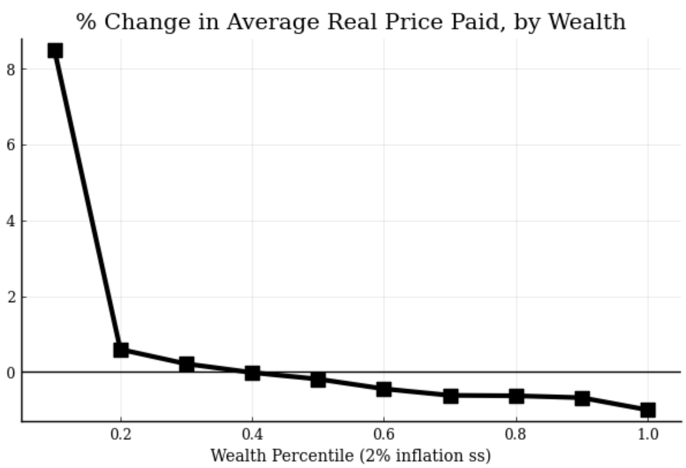


Figure 14: Inflation and prices paid, by total wealth.

Figure 14 shows the change in the average real price paid across the wealth distribution (again keeping fixed wealth percentile in the low inflation steady state). As inflation increases from 2% to 5%, the average real price paid rises sharply for households at the bottom of the wealth distribution — around 8% for the lowest wealth percentiles — while it falls modestly for households in the top half of the distribution. Higher inflation therefore compresses, and at the bottom of the distribution reverses, the positive sorting between wealth and prices paid.

In Figure 15, we decompose the change in average prices paid into three components: the change in households' search effort, the change in their purchase probabilities, and the change in market composition — the equilibrium adjustment of the distribution of posted prices and of the customers each firm serves. The rise in prices paid by low wealth households is driven almost entirely by their higher purchase probability: as the inflation tax raises the cost of continued search, they become less price sensitive and accept higher price draws. In contrast, the decline in prices paid by high wealth households is driven by the change in market composition. [CHECK: exact wording of the market-composition channel.]

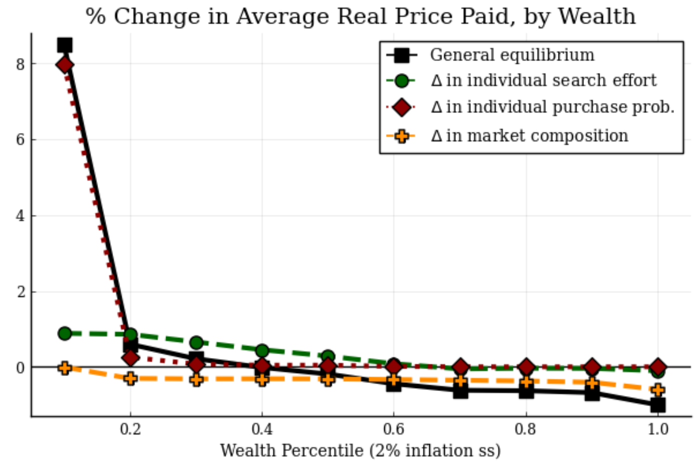


Figure 15: Decomposition of the change in prices paid, by total wealth.

8 Conclusion

This paper studies how inflation affects market power and the distribution of prices paid across households through endogenous search behavior. Using household-level purchase data, we document that during the high-inflation period of 2021–2023 higher-income households increasingly shifted their shopping toward lower-price retailers. As a result, differences in store choice across income groups became much smaller than in the preceding low-inflation period, when higher-income households consistently purchased from more expensive retailers.

To interpret this fact, we develop a heterogeneous-agent monetary search model in which households differ in wealth and income, choose shopping effort, and face an outside option of continued search. Firms post prices and compete for customers whose demand elasticities depend on both search intensity and the value of postponing purchases. The central mechanism is that inflation lowers the real return to liquid balances and changes households' outside options. High-wealth households respond more strongly along the search margin, shift toward lower-price firms, and compete more directly with lower-wealth households. This reallocation alters the composition of demand faced by firms and unravels the positive sorting between wealth and prices paid.

Our findings highlight a goods-market channel of redistribution that operates through endogenous demand composition rather than balance-sheet differences alone. Inflation changes not only real returns and purchasing power, but also who shops where and how elastic demand is at different prices. Market power and markup dispersion therefore vary with macroeconomic conditions.

More generally, our paper suggests that evaluating the distributional effects of inflation requires accounting for decentralized trade and heterogeneous search behavior. Monetary policy influences not only aggregate prices and real allocations, but also the competitive environment in which firms operate and the distribution of markups across households.

References

Achdou, Y., J. Han, J.-M. Lasry, B. Moll, and P.-L. Lions (2022). Income and wealth distribution in macroeconomics: A continuous-time approach. *Review of Economic Studies* (89), 45–86.

Aguiar, M. and E. Hurst (2007). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013). Time use during the great recession. *American Economic Review 103*(5), 1664–1696.

Albanesi, S. (2007). Inflation and inequality. *Journal of Monetary Economics 54*(4), 1088–1114.

Algan, Y., E. Challe, and X. Ragot (2011). Incomplete markets and the output-inflation tradeoff. *Economic Theory 46*, 55–84.

Anderson, R. G. and B. E. Jones (2011). A comprehensive revision of the monetary services (divisia) indexes for the united states. *Federal Reserve Bank of St. Louis Review 93*(5), 325–359.

Argente, D. and M. Lee (2021). Cost of living inequality during the great recession. *Journal of the European Economic Association 19*(2), 913–952.

Atalay, E., E. Frost, A. Sorensen, C. Sullivan, and W. Zhu (2023). Scalable demand and markups.

Auclert, A. (2019, June). Monetary polity and the redistribution channel. *American Economic Review 109*(6), 2333–67.

Auer, R., A. Burstein, S. Lein, and J. Vogel (2024). Unequal expenditure switching: Evidence from switzerland. *Review of Economic Studies 91*(5), 2572–2603.

Baqaee, D. R., E. Farhi, and K. Sangani (2024). The supply-side effects of monetary policy. *Journal of Political Economy 132*(4), 1065–1112.

Benabou, R. (1988). Search, price setting, and inflation. *Review of Economic Studies 55*(3), 353–376.

Benabou, R. (1989). Optimal price dynamics and speculation with a storable good. *Econometrica 57*, 41–80.

Benabou, R. (1992). Inflation and efficiency in search markets. *Review of Economic Studies 59*(2), 299–329.

Bethune, Z., M. Choi, S. Lotz, and G. Rocheteau (2024). Inflation and unemployment in the long run revisited. Technical report.

Bethune, Z. and G. Rocheteau (2023). Unemployment and the distribution of liquidity. *Journal of Political Economy: Macroeconomics 1*(4), 742–787.

Bewley, T. (1980). The optimum quantity of money. In J. Kereken and N. Wallace (Eds.), *Models of Monetary Economics*. Federal Reserve Bank of Minneapolis.

Bewley, T. (1983). A difficulty with the optimum quantity of money. *Econometrica 51*, 1485–1504.

Burdett, K. and K. Judd (1983). Equilibrium price dispersion. *Econometrica 51*(4), 955–969.

Burdett, K. and G. Menzio (2018). The (q, s, s) pricing rule. *The Review of Economic Studies 85*(2), 892–928.

Butters, G. R. (1977a). Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies 44*(3), 465–491.

Butters, G. R. (1977b). Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies 44*(3), 465–491.

Chiu, J. and M. Molico (2010). Liquidity, redistribution, and the welfare cost of inflation. *Journal of Monetary Economics 57*(4), 428–438.

Chiu, J. and M. Molico (2011). Uncertainty, inflation, and welfare. *Journal of Money, Credit, and Banking 43*(2), 487–512.

Choi, M. and G. Rocheteau (2024). Foundations of market power in monetary economies. *Journal of Economic Theory 222*, 105931.

Diamond, P. (1971). A model of price adjustment. *Journal of Economic Theory 3*, 156–168.

Diamond, P. A. (1993). Search, sticky prices, and inflation. *Review of Economic Studies 60*(1), 53–68.

Doepke, M. and M. Schneider (2006, December). Inflation and the redistribution of nominal wealth. *Journal of Political Economy 114*(6), 1069–1097.

Edmond, C., V. Midrigan, and D. Y. Xu (2023). How costly are markups? *Journal of Political Economy 131*(7), 1619–1675.

Erosa, A. and G. Ventura (2002). On inflation as a regressive consumption tax. *Journal of Monetary Economics 49*(4), 761–795.

Faber, B. and T. Fally (2022). Firm heterogeneity in consumption baskets: Evidence from home and store scanner data. *Review of Economic Studies 89*(3), 1420–1459.

Ferreira, C., J. M. Leiva, G. Nuno, A. Ortiz, T. Rodrigo, and S. Vazquez (2026). The heterogeneous impact of inflation on households' balance sheets. *SERIEs*, 1–23.

Guvenen, F., L. Pistaferri, and G. L. Violante (2022a). The global repository of income dynamics. Accessed 11/14/25.

Guvenen, F., L. Pistaferri, and G. L. Violante (2022b). Global trends in income inequality and income dynamics: New insights from GRID. *Quantitative Economics 13*, 1321–1360.

Hall, R. E. (2018). New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy.

Head, A. and A. Kumar (2005). Price dispersion, inflation, and welfare. *International Economic Review 46*(2), 533–572.

Head, A., H. Lloyd-Ellis, and H. Sun (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review 104*(4), 1172–1210.

Hobijn, B. and D. Lagakos (2005). Inflation inequality in the united states. *Review of Income and Wealth 51*(4), 581–606.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the u.s. retail sector. *The Quarterly Journal of Economics 134*(2), 715–783.

Kaplan, G. and G. Menzio (2016). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy 124*(3), 771–825.

Kaplan, G., B. Moll, and G. L. Violante (2018). Monetary policy according to hank. *American Economic Review 108*(3), 697–743.

Kaplan, G. and S. Schulhofer-Wohl (2017). Inflation at the household level. *Journal of Monetary Economics 91*, 19–38.

Loecker, J. D., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics 135*(2), 561–644.

McCall, J. (1970). Economics of information and job search. *The Quarterly Journal of Economics 84*(1), 113–126.

Molico, M. (2006). The distribution of money and prices in search equilibrium. *International Economic Review 47*(3), 701–722.

Mongey, S. and M. Waugh (2025, January). Pricing inequality. in mimeo.

Nord, L. (2024). Shopping, demand composition, and equilibrium prices. Technical report, SSRN. Available at SSRN: https://ssrn.com/abstract=4178271.

Nord, L. and G. Kaplan (2026). How does household spending affect retail prices? Technical report.

Pytka, K. (2024). Shopping frictions with household heterogeneity: Theory and empirics. Technical report, SSRN.

Rocheteau, G., P.-O. Weill, and T.-N. Wong (2018). A tractable model of monetary exchange with ex-post heterogeneity. *Theoretical Economics 13*, 1369–1423.

Rocheteau, G., P.-O. Weill, and T.-N. Wong (2021). An heterogeneous-agent new-monetarist model with an application to unemployment. *Journal of Monetary Economics 117*, 64–90.

Sangani, K. (2023). Markups across the income distribution: Measurement and implications. Available at SSRN.

Sara-Zaror, F. (2025). Inflation, price dispersion, and welfare: The role of consumer search. Technical report, SSRN. Available at SSRN: https://ssrn.com/abstract=4127502.

Stantcheva, S. (2024). Why do we dislike inflation? Technical Report 32300, National Bureau of Economic Research. NBER Working Paper.

Stroebel, J. and J. Vavra (2019). House prices, local demand, and retail prices. *Journal of Political Economy 127*(3), 1391–1436.

Varian, H. R. (1980). A model of sales. *American Economic Review 70*(4), 651–659.

A Proofs and derivations

A.1 Partial-equilibrium comparative statics with respect to inflation

This appendix derives the results of Section 4. Throughout, the posted price distribution $F(p)$ is held fixed, the asset is money with $r = -\pi$, and $\tau = 0$. We work under the simplification $\lambda \to \infty$, so that the idle and searching states merge into a single value function $W(a, e)$ satisfying

$$\rho W(a,e) = \max_{c,s} \left[u(c) - \delta(s) + \alpha V(a,e,s)\right] + W_a(a,e)\left(-\pi a + e - c\right) + \mathcal{P}_{e,e'} W, \tag{39}$$

where $\mathcal{P}_{e,e'} W \equiv P_{e,e'}[W(a,e') - W(a,e)]$, and the net purchase surplus is $\hat{V}(p; a, e) = U(\hat{y}) + W(a - p\hat{y}, e) - W(a, e)$ with $\hat{y}$ satisfying $U'(\hat{y}) = p\, W_a(a - p\hat{y}, e)$ at an interior solution and $\hat{y} = (a - \underline{a})/p$ when the borrowing constraint binds. At the borrowing constraint, the results below are unchanged with finite λ. We maintain the following regularity conditions: W is twice continuously differentiable in a on $(\underline{a}, \infty)$ with one-sided derivatives at $\underline{a}$, and differentiable in π, with $\Psi \equiv \partial W/\partial \pi$, $\Psi_a \equiv \partial \Psi/\partial a$, and Ψ_{aa} continuous.

Lemma 3. *In partial equilibrium with $F(p)$ fixed,*

$$\frac{\partial \hat{V}(p; a, e)}{\partial \pi} = \Psi(a - p\hat{y}, e) - \Psi(a, e). \tag{40}$$

Proof. Differentiating $\hat{V}$ with respect to π,

$$\frac{\partial \hat{V}}{\partial \pi} = \underbrace{\left[U'(\hat{y}) - p\, W_a(a - p\hat{y}, e)\right]}_{=0 \text{ by the FOC for } \hat{y}} \frac{\partial \hat{y}}{\partial \pi} + \Psi(a - p\hat{y}, e) - \Psi(a, e).$$

The bracketed term vanishes at an interior solution. When the borrowing constraint binds, $\hat{y} = (a - \underline{a})/p$ is independent of π, so $\partial \hat{y}/\partial \pi = 0$ and (40) holds directly. □

Since $\hat{V}$ enters the purchase probability only through the logit formula (12) and enters the effort condition (14) only through EV, with $\partial EV/\partial \pi = \varrho \cdot \partial \hat{V}/\partial \pi$, Lemma 3 delivers the two derivative formulas in (33).

The equation for Ψ and its slope. Differentiating (39) with respect to π, using the envelope theorem to eliminate the indirect effects through the optimal c and s, and using $\partial EV/\partial\pi = \varrho \cdot \partial\hat{V}/\partial\pi$ together with Lemma 3 inside V:

$$\rho\,\Psi(a,e) = \underbrace{-a\,W_a(a,e)}_{\text{direct erosion}} + \underbrace{\alpha\int_p \varrho(p;a,e)\,[\Psi(a-p\hat{y},e)-\Psi(a,e)]\,dG(p;s)}_{\text{search option value}} + \underbrace{\Psi_a(a,e)\,\dot{a}(a,e)}_{\text{wealth drift}} + \mathcal{P}_{e,e'}\Psi, \quad (41)$$

where $\dot{a}(a,e) = -\pi a + e - c(a,e)$. Differentiating (41) with respect to a,

$$\rho\,\Psi_a(a,e) = \underbrace{-W_a - a\,W_{aa}}_{\text{erosion slope}} + \underbrace{\alpha\frac{\partial}{\partial a}\int_p \varrho\,[\Psi(a-p\hat{y},e)-\Psi(a,e)]\,dG(p;s)}_{\text{option value slope}} + \underbrace{\Psi_{aa}\,\dot{a} + \Psi_a\,\dot{a}_a}_{\text{drift slope}} + \mathcal{P}_{e,e'}\Psi_a. \quad (42)$$

Evaluation at the borrowing constraint. Suppose $\underline{a} \leq 0$ and evaluate at the lowest income state $\underline{e}$, writing $\dot{a} \equiv \dot{a}(\underline{a},\underline{e})$. The state constraint delivers $\dot{a} \geq 0$. Since $c(\underline{a},\underline{e}) = \min\left\{(u')^{-1}(W_a(\underline{a},\underline{e})),\, \underline{e}-\pi\underline{a}\right\}$ and $c(\underline{a},\underline{e}) > 0$ by the Inada condition on u, the drift is bounded above by the inflow, $\dot{a} \leq \underline{e} - \pi\underline{a}$, and the binding case $\dot{a} = 0$ is equivalent to $W_a(\underline{a},\underline{e}) \leq u'(\underline{e}-\pi\underline{a})$ — an equilibrium property, not a primitive one. Evaluate each term of (42) at $(\underline{a},\underline{e})$ (derivatives are one-sided):

i) *Erosion slope.* $-W_a - a\,W_{aa}|_{(\underline{a},\underline{e})} = -W_a(\underline{a},\underline{e})\,[1-R(\underline{a})]$, where $R(a) \equiv -aW_{aa}(a,\underline{e})/W_a(a,\underline{e})$. Since $\underline{a} \leq 0$ and $W_{aa} < 0$, we have $R(\underline{a}) \leq 0 < 1$, so the erosion slope is at most $-W_a(\underline{a},\underline{e}) < 0$. No approximation is required.[15]

ii) *Option value slope.* At $\underline{a}$, lumpy expenditure is constrained at every p, so $p\hat{y} = a - \underline{a} = 0$ and $\partial(p\hat{y})/\partial a = 1$. Differentiating the search option value term with respect to a and evaluating at $(\underline{a},\underline{e})$: every term containing the difference $\Psi(a-p\hat{y},\underline{e}) - \Psi(a,\underline{e})$ vanishes, and the remaining term is

$$\alpha\int_p \varrho\left[\Psi_a(\underline{a},\underline{e})\left(1-\frac{\partial(p\hat{y})}{\partial a}\right) - \Psi_a(\underline{a},\underline{e})\right]dG = -\alpha\,\bar{\varrho}(\underline{a},\underline{e})\,\Psi_a(\underline{a},\underline{e}),$$

where $\bar{\varrho}(\underline{a},e) \equiv \int_p \varrho(p;\underline{a},e)\,dG(p;s) \in (0,1)$ is the expected purchase probability.

iii) *Drift slope.* The term $\Psi_a\,\dot{a}_a$ has $\dot{a}_a(\underline{a},\underline{e}) = -\pi - c_a \leq 0$ and is absorbed into the effective

[15]When $\underline{a} > 0$ (a case we do not use), the condition $R(\underline{a}) < 1$ must be imposed; under a steady-state approximation for W and CRRA preferences it is equivalent to $\underline{a} < e/[\pi(1+\sigma)]$, which holds by a wide margin under standard calibrations.

discount below. The term $\Psi_{aa}\,\dot{a}$ is bounded in magnitude by $|\Psi_{aa}(\underline{a},\underline{e})|\,\dot{a}$, and vanishes when the constraint binds ($\dot{a}=0$).

Collecting terms,

$$\left[\rho+\alpha\bar{\varrho}(\underline{a},\underline{e})-\dot{a}_a(\underline{a},\underline{e})+\textstyle\sum_{e'\neq\underline{e}}P_{\underline{e},e'}\right]\Psi_a(\underline{a},\underline{e})=-W_a(\underline{a},\underline{e})\,[1-R(\underline{a})]+\Psi_{aa}(\underline{a},\underline{e})\,\dot{a}+\textstyle\sum_{e'\neq\underline{e}}P_{\underline{e},e'}\,\Psi_a(\underline{a},e'). \tag{43}$$

Proposition 1's hypothesis that the lowest income state is low enough and sufficiently persistent is, precisely,

$$|\Psi_{aa}(\underline{a},\underline{e})|\;\dot{a}\;+\;\left(\textstyle\sum_{e'\neq\underline{e}}P_{\underline{e},e'}\right)\max_{e'\neq\underline{e}}\left|\Psi_a(\underline{a},e')\right|\;<\;W_a(\underline{a},\underline{e})\,[1-R(\underline{a})]\,. \tag{44}$$

The left side collects the two forces that could overturn the erosion slope. The first is the wealth drift: since $\dot{a}\leq\underline{e}-\pi\underline{a}$, it is controlled by an upper bound on $\underline{e}$, and it vanishes altogether when the constraint binds ($\dot{a}=0$). The second is income switching, controlled by the persistence of the lowest income state. All objects in (44) are computable in the calibrated model.

Proof of Proposition 1. The bracket on the left side of (43) is strictly positive: $\rho>0$, $\alpha\bar{\varrho}>0$, $-\dot{a}_a\geq 0$, and the transition rates are non-negative. The right side is bounded above by

$$-W_a(\underline{a},\underline{e})\,[1-R(\underline{a})]+|\Psi_{aa}(\underline{a},\underline{e})|\,\dot{a}+\left(\textstyle\sum_{e'\neq\underline{e}}P_{\underline{e},e'}\right)\max_{e'\neq\underline{e}}\left|\Psi_a(\underline{a},e')\right|,$$

where the maximum is finite because e takes finitely many values and Ψ_a is continuous. Under condition (44) this bound is strictly negative, so $\Psi_a(\underline{a},\underline{e})<0$. By continuity of Ψ_a in a, there is $\bar{\varepsilon}>0$ with $\Psi_a(a,\underline{e})<0$ for all $a\in(\underline{a},\underline{a}+\bar{\varepsilon})$. Fix $e=\underline{e}$ for parts i) and ii).

Part i). For $a\in(\underline{a},\underline{a}+\bar{\varepsilon})$ the constraint binds at every p in the support of F, so $p\hat{y}=a-\underline{a}\equiv\varepsilon>0$ and, by Lemma 3,

$$\frac{\partial\hat{V}(p;a,\underline{e})}{\partial\pi}=-\int_0^{\varepsilon}\Psi_a(a-t,\underline{e})\,dt\equiv\Delta_\varepsilon>0,$$

independent of p. Then $\partial\varrho/\partial\pi=\eta\,\varrho(1-\varrho)\,\Delta_\varepsilon>0$ for every p.

Part ii). From (33), $\delta''(s)\,\partial s/\partial\pi=\alpha\,\Delta_\varepsilon\int_p\varrho(p;a,\underline{e})\,\omega(p)\,dF(p)$. As $a\downarrow\underline{a}$, $\hat{V}(p;a,\underline{e})=U(\varepsilon/p)+$

$W(\underline{a}, \underline{e}) - W(a, \underline{e}) \to 0$ uniformly in p, using $U(0) = 0$ and continuity of U and W; hence $\varrho(p; a, \underline{e}) \to 1/2$ uniformly in p. Since $\int_p \omega(p) dF(p) = -e^{-\mu}$ and $\int_p |\omega(p)| dF(p)$ is bounded, for a close enough to $\underline{a}$,

$$\int_p \varrho(p; a, \underline{e})\, \omega(p)\, dF(p) \;\leq\; -\frac{e^{-\mu}}{2} + \sup_p \left| \varrho(p; a, \underline{e}) - \frac{1}{2} \right| \int_p |\omega(p)|\, dF(p) \;<\; 0,$$

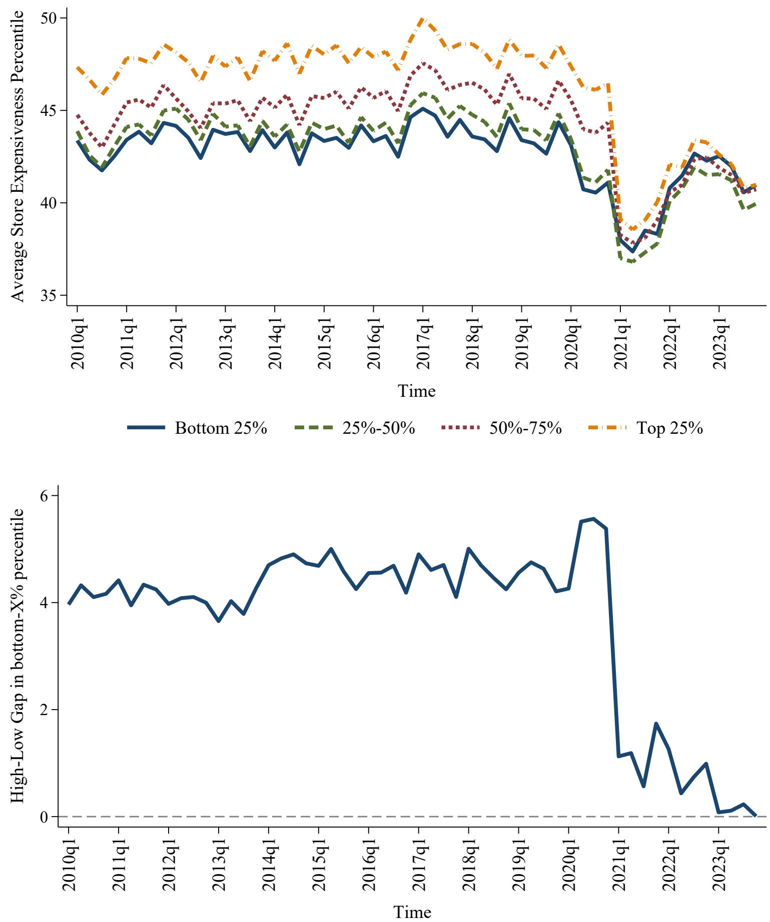
so $\partial s / \partial \pi < 0$. (Shrink $\bar{\varepsilon}$ if necessary.) $\square$

Intermediate and high wealth. Away from the constraint, $\hat{y} > 0$ and all terms of (42) are active; the option value slope opposes the erosion slope, the drift slope depends on unsigned higher-order objects, and the sign of Ψ_a is ambiguous. As $a \to \infty$ under CRRA flow utility with $\sigma > 1$, $W_a \to 0$ so the erosion slope vanishes; and since $\hat{y}$ remains bounded once the constraint is slack, $\Psi(a - p\hat{y}, e) - \Psi(a, e) \approx -p\hat{y}\,\Psi_a \to 0$, so the option value channel vanishes as well. The partial-equilibrium effects of inflation on ϱ and s therefore vanish in magnitude at high wealth, as summarized in Table 1.

B Data Appendix and Additional Figures

Figure 16 repeats the analysis in Figures 1 using expenditure weights rather than trip weights. The qualitative pattern is nearly identical. During the low-inflation years, higher-income households shop at systematically more expensive stores, with the top-bottom gap fluctuating around four to five percentile points. Following the inflation surge, the gap narrows markedly, driven primarily by a downward shift in the store positioning of higher-income households. The persistence of this compression under spending weights indicates that the result does not arise mechanically from differences in basket sizes or expenditure intensity, but reflects a genuine reallocation across price segments.

Figure 17 examines the sensitivity of the threshold-based measure to different definitions of the “low-price” segment. The figure plots the difference between top and bottom income quartiles in the share of trips to stores below the 20th, 25th, 30th, and 35th percentiles of the contemporaneous expensiveness distribution. Across all cutoffs, the qualitative pattern is the same. During the low-inflation period, the gap is negative and stable, indicating that lower-

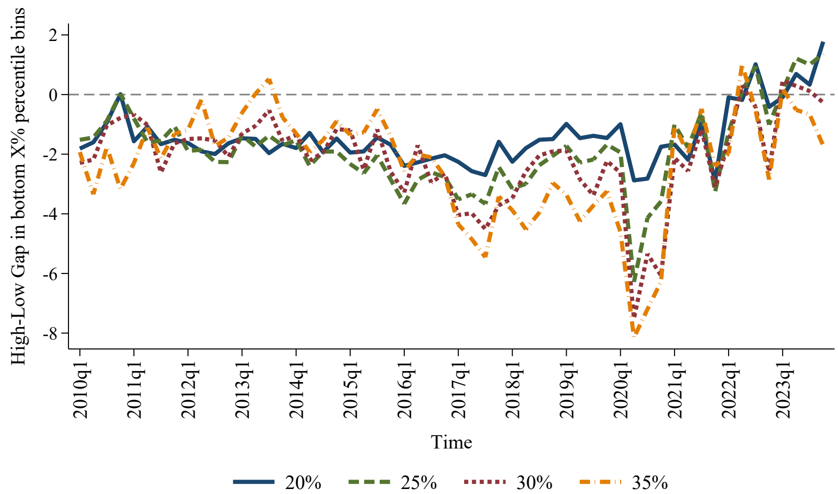


Notes: Panel A plots the mean expensiveness percentile of visited stores separately for each income quartile. Panel B plots the difference between the top and bottom income quartiles. Percentiles are computed within each DMA and quarter. All measures are spending-weighted.

Figure 16: Income Differences in Store Choice (Spending-weighted)

income households devote a larger share of trips to cheaper stores. When inflation accelerates, the gap rises toward zero, reflecting a relative increase in low-price shopping among higher-income households. While the magnitude varies mechanically with the threshold, the compression of the income gradient during the inflation episode is robust to these alternative definitions.

As a robustness exercise, we construct a time-invariant measure of store expensiveness. The goal is to separate changes in shopping behavior from potential shifts in the relative price ranking of stores over time. To do so, we estimate persistent store-level price effects using data from 2016-2019, a period of relatively low and stable inflation. Specifically, we estimate a specification



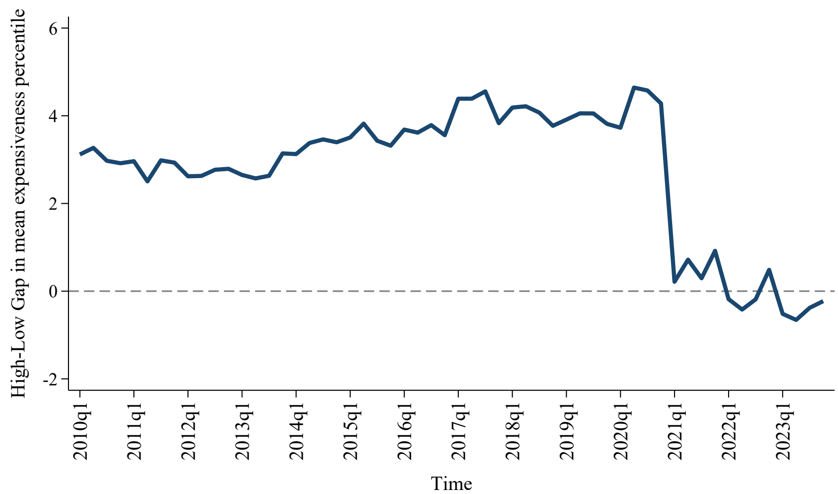
Notes: The figure plots the difference between top and bottom income quartiles in the share of trips to stores below alternative percentile thresholds of the contemporaneous expensiveness distribution (20%, 25%, 30%, and 35%). Percentiles are computed within each DMA and quarter. All measures are trip-weighted.

Figure 17: High-Low Income Gaps in Bottom X% Group Over Time

with store-DMA fixed effects, product-by-quarter effects, and DMA-by-quarter effects, so that the estimated store component captures persistent differences in price levels net of product composition and local shocks. We then assign each store a fixed percentile in the pre-period distribution and use this ranking throughout the sample.

Figure 18 plots the top-bottom income gap under this fixed ranking. The solid line shows the difference in mean expensiveness percentiles, and the dashed line shows the corresponding gap in the share of trips to stores in the bottom 30 percent of the pre-period distribution. During the low-inflation years, the gap is stable and economically meaningful, with higher-income households consistently shopping at more expensive stores.

When inflation rises, the high-low separation narrows substantially under the fixed ranking as well. The mean-percentile gap declines by roughly one percentile point, and the bottom-segment gap moves toward zero. The persistence of gap compression under a fixed store ordering indicates that the result is not driven by changes in the relative price ranking of stores, but reflects genuine reallocation of shopping across price segments during the inflation episode.



Notes: The figure plots the difference between high- and low-income households (Top 25% minus Bottom 25%) in two measures of store choice: the mean expensiveness percentile (left axis) and the share of trips to the bottom 30% of stores (right axis). Percentiles are computed using a fixed pre-period ranking of store expensiveness. All measures are trip-weighted. The vertical line marks the onset of the inflation episode.

Figure 18: High-Low Income Gaps in Store Choice Over Time