EIEF WORKING PAPER SERIES

IEF

Einaudi Institute for Economics and Finance

EIEF Working Paper 28/13

November 2013

Price Dynamics with Customer Markets

by

Luigi Paciello

(EIEF and CEPR)

Andrea Pozzi

(EIEF and CEPR)

Nicholas Trachter

(Federal Reserve Bank of Richmond)

Price Dynamics with Customer Markets*

Luigi Paciello
Einaudi Institute for Economics and Finance and CEPR

Andrea Pozzi
Einaudi Institute for Economics and Finance and CEPR

Nicholas Trachter
Federal Reserve Bank of Richmond

December 29, 2015

Abstract

The customer base of a firm is an important and persistent determinant of its performance. We show that the interaction between firm pricing and customer dynamics affects the determination of price markups in response to idiosyncratic and aggregate shocks, and that such interaction varies with the cycle. We study an economy where the customer base of a firm is persistent because of search frictions preventing customers from freely relocating across suppliers of consumption goods. The key feature of our model is that the elasticity of the customer base to price - the *extensive margin* elasticity of demand - depends on the customers' endogenous opportunity cost of search. This results into a new channel affecting the cyclical relationship between consumer search, markups and price dispersion. In particular, an increase in the marginal utility of consumption increases customers' incentives to search resulting in higher demand elasticity and lower markups and price dispersion. To analyze the quantitative implications of customer markets, we calibrate the search friction using rich U.S. data on consumer shopping behavior and good prices. The estimated model predicts a procyclical response of search and a countercyclical response of markups to aggregate demand shocks amplifying their effect on output.

JEL classification: E30, E12, L16

Keywords: customer markets, price setting, product market frictions

---

*Corresponding Author: luigi.paciello@eief.it. Previous drafts of this paper circulated under the title "Price Setting with Customer Retention." We benefited from comments on earlier drafts at the Minnesota Workshop in Macroeconomic Theory, NBER Summer Institute 2014, 10th Philadelphia Search and Matching conference, Goods Markets Paris Conference 2014, 2nd Rome Junior Macroeconomics conference, 2nd Annual UTDT conference in Advances in Economics, ESSET 2013, MBF-Bicocca conference, MaCCI Mannheim, the Collegio Carlo Alberto Pricing Workshop, the Macroeconomy and Policy, EARIE 2014, IIOC 2015 and seminars at the Federal Reserve Bank of Minneapolis, Bank of France, Bank of Spain, Bocconi University, Columbia University, Federal Reserve Bank of Richmond, the Ohio State University, University of Pennsylvania, Paris School of Economics, Toulouse School of Economics, Macro Faculty Lunch at Stanford, Drexel University, Indiana University (Kelley) and University of Tor Vergata. We thank Fernando Alvarez, Lukasz Drozd, Huberto Ennis, Mike Golosov, Bob Hall, Christian Hellwig, Hugo Hopenhayn, Eric Hurst, Pat Kehoe, Francesco Lippi, Erzo Luttmer, Kiminori Matsuyama, Guido Menzio, Dale Mortensen, Jaromir Nosal, Ezra Oberfield, Nicola Pavoni, Facundo Piguillem, Valerie Ramey, Leena Rudanko and Katja Seim. Luigi Paciello thanks Stanford University for its hospitality. The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1 Introduction

The customer base of a firm -the set of customers buying from it at a given point in time- is an important determinant of firm performance. Its effects are long lasting, as customer-supplier relationships are subject to a certain degree of stickiness (Hall (2008)). Starting with Phelps and Winter (1970), a large literature has stressed that the price is an important instrument to attract and retain customers, and firms seek to maintain and grow their customer base through their pricing decisions. Analyses of the implications of such a mechanism for macroeconomic dynamics typically rely on reduced form specifications of the evolution of the customer base (Rotemberg and Woodford (1991)). However if shocks, either idiosyncratic or aggregate, affect customers' decisions to search for a new firm (Bai et al. (2012), Coibion et al. (2015), and Kaplan and Menzio (forthcoming)), leaving search behavior unmodeled may have important implications for the propagation of shocks to prices and demand. In this paper, we emphasize a novel channel by which aggregate shocks propagate to optimal price markups through their effect on consumer search behavior. In our model, shocks that increase the marginal utility of consumption not only make customers more eager to consume, but also push them to search harder for better sellers. This increases the mass of customers who are potentially up for grabs, motivating firms to lower their prices, and compress their markups, to retain them. As a result, both markups and price dispersion reduce and the effect of an aggregate demand shock is amplified.

We build a microfounded model of firm pricing with customer markets and focus on the interaction of pricing with customers' search intensity. The salient feature of our setting is that it delivers endogenous customer dynamics, arising as a consequence of customers search and exit decisions, which we explicitly model. Firms set prices taking into account their effects on the evolution of their customer base. Customers respond to price changes but their ability to reallocate across suppliers is impaired by the presence of search frictions.[1] Since the relevance of the mechanism we are exploring depends on the magnitude of the elasticity of the customer base to firm prices, we estimate the model exploiting novel data from a large US retailer documenting posted prices and tracking movements in its customer base.

We make two contributions to the literature. First, we obtain implications for the cyclical relationship between the average price markup, price dispersion and aggregate demand shocks which differ, even qualitatively, from those that would arise if customer search behavior were invariant to shocks. Empirically, we find a comovement between price dispersion and demand in U.S. data that is consistent with the predictions of the model. Second, we show that our

[1]Modeling the market friction as a search cost suits well our application since search costs have been found to importantly affect price dispersion in clustered retail markets (Sorensen (2000)) similar to those to which our empirical application refers to.

estimated model captures important features of the micro data. Namely, we find that it matches the existing evidence on the shape of the cross-sectional distribution of price levels (Kaplan and Menzio (2015)), in particular the dispersion and the kurtosis. Moreover, the model predicts high persistence of firm demand (revenues) induced by the stickiness of the customer base in line with the mechanism studied by Foster et al. (2008, 2016).

We endogenize customer dynamics by modeling the game between a firm and its customers. Customers start each period in the customer base of the firm from which they bought in the previous period. Every period, firms draw a new idiosyncratic productivity level, and post a price. Then, each customer can decide to pay an idiosyncratic search cost to observe the state of another randomly selected firm, compare it to that of her old supplier, and decide where to buy (*extensive* margin of demand). After these decisions have been made, each customer decides her purchased quantity of the good (*intensive* margin of demand). In this setting, firms face the common invest\harvest trade-off (Galenianos and Gavazza (2015)): charging a higher price and extracting more surplus from customers, versus posting a lower price to extract a lower surplus but from a larger mass of customers.

While being tractable, the model provides a rich laboratory to study how the relationship between customer and price dynamics is shaped, in equilibrium, by idiosyncratic production and search costs. The price posted by the firm and its current level of productivity determine the value that customers obtain if they remain in the firm's customer base. Customers of firms in the left tail of the distribution of such values are more likely to search and leave. As a result, firms have incentives to avoid being in the left tail of this distribution, which introduces an element of strategic complementarity in prices affecting the dispersion and the kurtosis of the price distribution.

The quantitative importance of customer markets for firm pricing depends on the magnitude of the elasticity of the customer base to firm prices. This calls for strengthening our evidence on the link between pricing and customer base dynamics, which currently rests mainly on anecdotes or surveys (Blinder et al. (1998), and Fabiani et al. (2007)). We exploit novel data from the retail industry to provide direct evidence that firm prices influence households' decision to stop buying from their current supplier. Our empirical analysis relies on scanner data documenting pricing and customer base evolution for a major U.S. retailer. The data contain information on purchases for a large sample of customers between 2004 and 2006. Household-level scanner data are particularly well suited to study customer base dynamics. First, we observe a wealth of details on all the shopping trips each household makes to the chain (list of goods purchased, prices, quantities, etc.). More importantly, we can infer when customers leave the retailer by looking at extended spells without purchases at the chain. Hence, these data allow us to study the relation between a customer's decision to abandon the firm and the price of the bundle of goods she consumes there. We show that an increase in the price significantly raises the probability that the household leaves the firm. This implies that the customer base is elastic to prices: for instance, if the price were 1% higher for a full year, our estimates imply that the customer base would decrease by 7%.

We use the estimated price elasticity of the customer base, jointly with additional moments from our data, to identify the key objects of the model: the distribution of search costs and the productivity process. We employ the estimated model to quantify the relevance of the customer market in explaining the patterns we observe in the data. Moreover, the estimated model offers a chance to investigate the relevance of customer dynamics in shaping the propagation of aggregate shocks. In our baseline model, a positive demand shock impacts (positively) customers' willingness to consume but they will also adjust their search intensity since it is now more valuable for them to be matched with good sellers. This implies that there are more consumers looking to switch, which forces firms to lower their markup to retain them. The result is that both markups and price dispersion shrink, magnifying the effect of the demand shock. In a counterfactual exercise, we show the effect of the same shock in an observationally equivalent economy where customers cannot adjust their search behavior in response to the aggregate shock. In this case, the extensive margin elasticity of demand is invariant to the aggregate demand shock and the predicted outcome is radically different. Markups and price dispersion increase in response to the positive demand shock attenuating the response to the shock. Using data from the Consumer Expenditure Survey, we document that price dispersion and aggregate demand appear to be negatively correlated, providing support to the quantitative relevance of the mechanism at the heart of our model.

Related Literature. Our paper relates to the seminal work by Phelps and Winter (1970) who study the pricing problem of a firm facing customer retention concerns. In their paper, the response of the firm's customer base to a change in the firm's price is modeled with an ad hoc function. We instead endogenize customer dynamics which arise as the outcome of customers' optimal search decisions in response to firms' pricing. Fishman and Rob (2003), Alessandria (2004), and Menzio (2007) also study the firm price-setting problem in models where search costs prevent customers from freely moving to the lowest price supplier. Fishman and Rob (2003) study the implications of customer markets for firm dynamics. Alessandria (2004) shows that such a model can generate large and persistent deviations from the law of one price, consistent with the empirical evidence on international prices. Menzio (2007) looks at the role of asymmetric information and commitment in the optimal pricing decision of the firm.

Unlike the literature cited above, we exploit micro data to discipline our model and provide a quantitative assessment of the relevance of customer markets for firm pricing and the propagation of aggregate shocks. This relates our findings to contributions that aim at documenting empirical stylized facts on the behavior of prices and markups. Our evidence on the shape of the price distribution ties in to the recent empirical work by Kaplan and Menzio (2015). While their focus is on customers and the price they pay for the same good (or bundle of goods), we are interested in the point of view of sellers and the price they charge.

Several other studies analyze the implications of product market frictions for business cycle fluctuations. Bai et al. (2012) analyze a demand-driven business cycle model where preference shocks affect consumers' search incentives and consumption by directly impacting production efficiency, so to show up as shocks to the Solow residual. We also emphasize that consumer search behavior amplify the transmission of aggregate demand shocks to output but through a different channel. In particular, we focus on the impact of cyclicality of consumer search on optimal markups, and show that markups fall when aggregate demand is high.

Petrosky-Nadeau and Wasmer (2015) and Kaplan and Menzio (forthcoming) study the interaction of labor and product market frictions, linking unemployment dynamics to consumer search effort. In these models, whether consumer search amplifies or dampens the recessionary implications of higher unemployment depends on how consumer search intensity comoves with unemployment. Empirical evidence on the sign of the relationship of search activity or shopping time with unemployment is mixed.[2] Coibion et al. (2015) document the relationship between the household expenditure allocation across retailers and unemployment and find that households pay on average a lower price when unemployment is higher. While we share with these papers the focus on the cyclicality of consumer search effort, we look at a different driver of the cyclicality of search. In particular, we show that when demand shocks affect the benefit of searching, and in particular the marginal utility of consumption, consumer search is procyclical leading to countercyclical markups and price dispersion. This channel amplifies the effect of demand shocks on output.

Because we study the impact of aggregate shocks on markups through their impact on the elasticity of demand, we also relate to the literature on deep habits (Ravn et al. (2006)). In this literature there is, typically, no *extensive* margin of demand as each consumer buys from all firms at any point in time, albeit with different habit and expenditure, and all the adjustment in demand takes place along the *intensive* margin. In our model, instead, the extensive margin plays a key role. Markups are tightly linked to customer dynamics and search behavior so that the quantitative predictions of the model depend on the magnitude of the extensive margin elasticity of demand which we measure using micro data.

[2] See Petrosky-Nadeau et al. (2014) for a discussion.

Another set of related contributions uses customer markets to address questions different from the ones we study here. Gourio and Rudanko (2014) explore the relationship between the firm's effort to capture customers and its performance. They show that customer markets have nontrivial implications for the relationship between investment and Tobin's q. Drozd and Nosal (2012) introduce in a standard international real business cycle model the notion that, when producers want to increase sales, they must exert effort to find new customers. This extension helps to rationalize a number of empirical findings on the dynamics of international prices and trade. Kleshchelski and Vincent (2009) examine the impact of customer markets on the pass-through of idiosyncratic cost shocks to prices, whereas we focus on how aggregate shocks affect the distribution of prices and aggregate output. Moreover, they focus on a symmetric equilibrium with no price dispersion, whereas price dispersion together with customer dynamics are important characteristics of both our model and data. Dinlersoz and Yorukoglu (2012) focus on the importance of customer markets for industry dynamics in a model where firms use advertising to disseminate information to uninformed customers. Shi (2011) studies a setting where firms cannot price discriminate across customers and use sales to attract new customers. Burdett and Coles (1997) study the role of firm size for pricing when firms use the price to attract new customers. Their work complements ours: price and customer dynamics in their setting are shaped by the heterogeneity in firm size (age). For us, the driving force is the heterogeneity in productivity. The industrial organization literature has also studied the implications of customer markets for a variety of subjects. For instance, Foster et al. (2016) stress their role in affecting firm survival and Einav and Somaini (2013) and Cabral (2014) focus on their effect on the competitive environment.

Finally, our paper relates to the recent work of Vavra (2014) documenting the cyclical properties of the distribution of price *changes*. He finds that the standard deviation of the distribution of price changes rises in a recession. We instead study the cyclical properties of the cross-sectional distribution of the price *levels*, and find that the standard deviation of this distribution (the price dispersion) is higher when consumption expenditure is higher.

The rest of the paper is organized as follows. In Section 2 we lay out the model and characterize the equilibrium. Section 3 presents the data and descriptive evidence of the relationship between customer dynamics and prices, and discusses identification and estimation of the model. In Section 4 we present some quantitative predictions of the model and compare them with empirical evidence from our data. In Section 5 we introduce an application of the model with the goal of studying the implications of customer markets for the dynamics of markups and price dispersion. Section 6 concludes.

2 The model

The economy is populated by a measure one of firms producing a homogeneous good and by a measure one of customers who consume it. The economy is in steady state and there are no foreseen aggregate shocks.

2.1 The problem of the firm

Firms produce the same homogeneous good. We assume a linear production technology $y = z\,\ell$ where ℓ is the production input, and z is the firm-specific productivity. Idiosyncratic productivity is distributed according to a conditional cumulative distribution function $F(z'|z)$ with bounded support $[\underline{z}, \bar{z}]$. We also assume that $F(\,z'|\,z_h)$ first order stochastically dominates $F(\,z'|\,z_l)$ for any $z_h > z_l$ to induce persistence in firm productivity. The profit per customer accrued to the firm is given by $\pi(p, z) \equiv d(p)(p - w/z)$, where p denotes the price, the constant $w > 0$ denotes the marginal cost of the input ℓ, and the function $d(\cdot)$ is a downward sloping demand function.[3] We assume that profits per customer are single-peaked in p.

Firms differ not only in their idiosyncratic productivity but also in the mass of customers buying from them. In particular, we denote by m the firm's *customer base* which is defined as the mass of customers who bought from that firm in the previous period, adjusted for an exogenous attrition rate δ.Starting from a given customer base m, the mass of customers actually buying from the firm in the current period is determined in equilibrium and we conjecture, and later verify, that it is given by the function $\mathcal{M}(m, p, z)$ depending on the price and productivity of the firm in the current period, as well as on the customer base.

We assume a constant probability κ of a firm exiting the market. Once a firm exits the market it loses all customers and its value is zero. An exiting firm is replaced by a new firm which starts with a customer base m_0, and draws a productivity z_0 from the invariant productivity distribution $\bar{F}(z)$ associated to the conditional distribution $F(z'|z)$.[4]

We study a stationary Markov Perfect equilibrium where pricing strategies are a function of the current state. Firms set prices every period without commitment and without discriminating across customers.[5] As there are no aggregate shocks, the aggregate state is constant and the relevant state for the firm problem in period t is the pair $\{z, m\}$. The firm

[3]In Appendix E, we extend this framework adding a model of the labor market to endogenize the wage w.

[4]The invariant distribution is obtained by solving $\bar{F}(z) = \int_{\underline{z}}^{\bar{z}} F(z|x)d\bar{F}(x)$ for all $z \in [\underline{z}, \bar{z}]$.

[5]See Nakamura and Steinsson (2011) for a model of pricing with customer markets where a commitment to a price path can be sustained in equilibrium.

pricing problem in its recursive form solves

$$\tilde{W}(z,m) \;=\; \max_{p}\; \mathcal{M}(m,p,z)\,\pi(p,z) \;+\; \beta\,(1-\kappa)\int_{\underline{z}}^{\bar{z}} \tilde{W}(z',m')\;dF(z'|\,z) \tag{1}$$
$$\text{s.t.} \qquad m' = (1-\delta)\,\mathcal{M}(m,p,z),$$

where $\tilde{W}(z,m)$ denotes the firm value at the optimal price. The price impacts firm value through two channels. First, it affects the level of profits per customer as in standard models of firm pricing. Given our assumption of single-peakedness of the profit function $\pi(p,z)$, there is a unique level of p that maximizes the profits per customer. Second, the price p affects the dynamics of the customer base. In fact, it influences the mass of customers buying from the firm in the current period, and, if there is persistence in the evolution of the customer base, the mass of customers buying from the firm in future periods. As a result, the pricing problem of the firm is dynamic in nature.

We study an environment where there is persistence in the customer base, as in Phelps and Winter (1970) and Rotemberg and Woodford (1999). These models assume a functional form for the evolution of the customer base where the mass of customers served by a firm is given by the product of its original customer base and a growth rate, which depends on its (relative) price. Our conjectured law of motion for customers preserves this standard structure and is given by:

$$\mathcal{M}(m,p,z) \equiv m\,\Delta(p,z)\;. \tag{2}$$

This similarity notwithstanding, there are two important innovations that we introduce. First, while the customer evolution is typically characterized with ad-hoc functional form assumptions, our $\Delta(\cdot)$ function is endogenous and results from the solution to a game between the firm and its customers. It depends on the equilibrium distribution of prices as well as on the distributions of productivity and search costs. Accounting for this dependence matters for the estimation where we will match micro moments obtained from customers' decisions. Moreover, it has important implications when using the model for policy experiments, as we will illustrate with the application in Section 5.

Second, we generalize the law of motion so that it can depend not only on the price the firm sets but also on its productivity. This extension allows us to to study the mapping from the distribution of productivities to the distribution of prices. It also proves useful when we bring the model to the data, since having heterogeneity in productivity helps us to match the cross-sectional variation in prices. Our formulation does, however, share an important feature with classic customer market models: the growth rate of the customer base does not depend on the initial mass of customers. This property allows for a substantial simplification of the firm's problem. In particular, it can be obtained that the value function of a firm is homogeneous of degree one in m, i.e. $\tilde{W}(z,m) = m\ \tilde{W}(z,1) \equiv m\ W(z)$, where using equation (1) and $\mathcal{M}(m,p,z) = m\,\Delta(p,z)$, it is immediate to show that $W(z)$ solves[6]

$$W(z) \;=\; \max_{p}\ \Delta(p,z)\,\pi(p,z) + \Delta(p,z)\,\beta\,(1-q)\int_{\underline{z}}^{\bar{z}} W(z')dF(z'|\,z)\ , \tag{3}$$

where $q \equiv \kappa + \delta - \kappa\,\delta$ is the probability of exogenous dissolution of the firm-customer match due to either firm or customer random exit. The relevant state to the firm pricing problem is its productivity, as the level of the customer base affects the firm value multaplicatively. The solution to the firm problem in equation (3) gives an optimal pricing strategy that depends on productivity and we denote by $\hat{p}(z)$.

We emphasize that, while the initial level of the customer base does not affect the optimal price, its evolution does. This follows as a change in the price affects the growth rate of the customer base, i.e., the value of $\Delta(p,z)$, and given the persistence of the customer base, it affects the firm value in the current period as well as in future periods. Our framework is well suited to capture the relationship between firm prices and customer dynamics when this is driven by variation in idiosyncratic productivity; extending it to encompass how firm size affects this relationship is an interesting direction for future research.[7]

The objective of the firm maximization problem can be expressed as the product of two terms, $W(z) \equiv \Delta(\hat{p}(z),z)\,\Pi(\hat{p}(z),z)$, where $\Pi(p,z)$ denotes the expected present discounted value of each customer to the firm. Under the assumption that the functions $\Delta(p,z)$ and $\pi(p,z)$ are differentiable in p, the first order condition to the firm problem is given by

$$\frac{\partial \Pi(p,z)}{\partial p}\ \frac{p}{\Pi(p,z)} = -\frac{p}{\Delta(p,z)}\frac{\partial \Delta(p,z)}{\partial p}\ , \tag{4}$$

where we define $\varepsilon_m(p,z) \equiv \partial \log(\Delta(p,z))/\partial \log(p)$ as the *extensive margin* elasticity of demand. We will discuss conditions under which equation (4) is necessary and sufficient in Section 2.3. The function $\Pi(p,z)$ is maximized at the static profit maximizing price,

$$p^*(z) \equiv \frac{\varepsilon_d(p)}{\varepsilon_d(p)-1}\ \frac{w}{z}\ , \tag{5}$$

where we define $\varepsilon_d(p) \equiv \partial \log(d(p))/\partial \log(p)$ as the *intensive margin* elasticity of demand.

[6]Under the assumption that the discount rate β is low enough so that the maximization operator in equation (3) is a contraction, by the contraction mapping theorem we can conclude that our conjecture about homogeneity of $\tilde{W}(z,m)$ is verified.

[7]Incidentally, our setup lends itself well to the data we use for the estimation where we observe prices from several firms but the customer base of only one chain. See section 3 for details.

Equation (4) implies however that, due to concerns about customer dynamics, the optimal price is in general different from the one that maximizes static profits.

If the growth in the customer base is non-increasing in the price, equation (4) implies that setting a price above the static profit maximizing price is never optimal. Hence, $\hat{p}(z) \leq p^*(z)$ for all z. Moreover, if the growth in the customer base is strictly decreasing in the price in a neighborhood of the static profit-maximizing price $p^*(z)$, the optimal price is pushed downwards with respect to it, i.e. $\hat{p}(z) < p^*(z)$. The first order condition in equation (4) illustrates the trade-off the firm faces when setting the price in a region where customer retention is a concern. The optimal price balances the marginal benefit of an increase in price (more profit per customer) with the cost (decrease in the customer base). The requirement that the solution to the firm problem must satisfy the first order condition implies that we study equilibria where the firm objective, and in particular $\Delta(p, z)$, is differentiable in p. In Section 2.3, we will derive the necessary equilibrium properties that guarantee that these properties are satisfied.

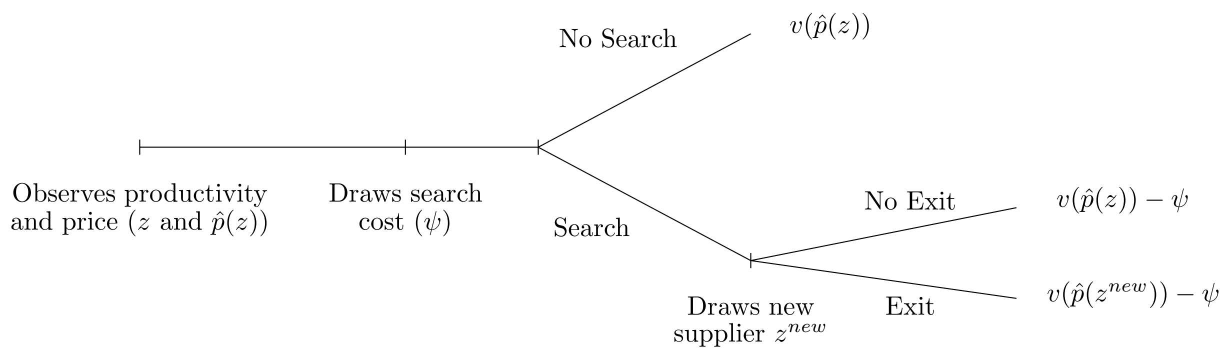
2.2 The problem of the customer

Customers value the good sold by the firms described in the previous section according to the function $v(p)$, denoting the customer surplus associated to the demand function $d(p)$. We assume that $v(p)$ is continuously differentiable with $v'(p) < 0$, and bounded above with $\lim_{p \to 0^+} v(p) < \infty$. These properties are satisfied in standard models of consumer demand.

Each customer starts the period in the customer base of the firm she bought from in previous period. At the beginning of every period, a customer can be randomly reallocated to a new entrant because either the firm she was matched with exited (with probability κ) or with probability δ the customer herself leaves for random reasons (for instance she moved to a different city). We allow for random exit to acknowledge that price dynamics, the object we study in detail in this paper, are unlikely to account for all the exits observed in the data. Conditioning on a firm surviving, random exit is i.i.d. across customers of that firm.

After random relocation has taken place, the customer observes perfectly the state of the firm she is matched to; in particular she observes its productivity. Given the equilibrium pricing function of the firm, this allows her to assess the probability distribution of the path of prices of that firm. After observing the state of her current match, the customer decides whether she wants to pay a search cost to draw another firm. The search cost $\psi \geq 0$ is measured in units of customer surplus, it is idiosyncratic to each customer and it is drawn each period from a cumulative distribution $G(\psi)$, with an associated density $g(\psi)$. For tractability, we restrict our attention to density functions that are continuous on all the support. Heterogeneity, albeit transitory, in search costs makes the customer base a

Figure 1: The problem of a customer matched to a firm with productivity z



continuous function of the price and allows us to study firms' pricing decisions that are not necessarily knife-edge in the trade-off between maximizing demand and markups.

The customer can search at most once per period. Search is random, with the probability of drawing a particular firm being proportional to its customer base m. As in Fishman and Rob (2005), this assumption captures the idea that consumers search for new suppliers not by randomly sampling firms but by randomly sampling other consumers. On the technical side, this is the key assumption that will allow us to solve for an equilibrium where the value of a firm scales up multiplicatively with its customer base. Conditional on searching, the customer observes the state of the new match and then makes another decision concerning whether to exit the customer base of her initial firm and match to the new firm. In particular, the customer compares the distribution of the path of current and future prices at the two firms and buys from the firm offering higher expected value. Finally, we assume that a customer cannot recall a particular firm once she exits its customer base. Figure 1 summarizes timing and payoffs of the problem of the customer.

We next characterize the customer problem. Let $V(p, z, \psi)$ denote the value function of a customer i who has drawn a search cost ψ and is matched to firm j, which has current productivity z and posted price p. This value function solves the following problem,

$$V(p, z, \psi) = \max \left\{ \bar{V}(p, z) \, , \, \hat{V}(p, z) - \psi \right\} \, , \tag{6}$$

where $\bar{V}(p, z)$ is the customer's value if she does not search, and $\hat{V}(p, z) - \psi$ is the value if she does search. The value in the case of not searching is

$$\bar{V}(p,z) = \xi\, v(p) + \beta\,(1-q)\,\mathbb{E}_G\left[\int_{\underline{z}}^{\bar{z}} V(\hat{p}(x), x, \psi')\, dF(x|z)\right] + \beta\, q\,\mathbb{E}_G\left[\int_{\underline{z}}^{\bar{z}} V(\hat{p}(x), x, \psi')\, d\bar{F}(x)\right] , \tag{7}$$

where ξ is an utility shifter set equal to 1 in our steady state analysis.[8] We notice that the state of the firm problem depends on the productivity z because the pricing function $\hat{p}(\cdot)$ mapping future productivity into prices in the Markov equilibrium makes productivity z a sufficient statistic for the distribution of future prices at the firm. We also notice that the state of the firm problem includes the current price p, despite the fact that in equilibrium productivity is enough to determine the current price, as this notation is needed to study the game between the firm and its customers where the firm could, in principle, deviate from the equilibrium price. Finally, the expectation operator $\mathbb{E}_G[\cdot]$ refers to the realization of future search costs which are drawn from the i.i.d. distribution G. The value function $\bar{V}(p,z)$ is strictly decreasing in p and increasing in z.

Given the specifics of the search technology, the value to the customer if searching is given by

$$\hat{V}(p,z) = \int_{-\infty}^{+\infty} \max\left\{\bar{V}(p,z)\,,\, x\right\} dH(x)\ , \tag{8}$$

where the customer takes expectations over all possible draws of potential new firms, and where $H(\cdot)$ is the equilibrium cumulative distribution of continuation values from which the firm draws a new potential match when searching.

We are now ready to describe the customer's optimal search and exit policy rules. Such policies are characterized by simple cut-off rules. The customer matched to a firm with productivity z charging price p searches if she draws a search cost $\psi \leq \hat{\psi}(p,z)$, where

$$\hat{\psi}(p,z) \equiv \int_{\bar{V}(p,z)}^{\infty} \big(x - \bar{V}(p,z)\big)\, dH(x) \geq 0$$

is the threshold to search. Conditional on searching, the customer exits if she draws a new firm promising a continuation value $\bar{V}^{new}$ larger than the current match, i.e. $\bar{V}^{new} \geq \bar{V}(p,z)$. Notice that the threshold $\hat{\psi}(p,z)$ is strictly increasing in p. The dependence on the price is straightforward, following from its effect on the surplus $v(p)$ that the customer can attain in the current period. The intuition behind the dependence on the firm's productivity is that, as searching is a costly activity, the decision of which firm to patronize is a dynamic

[8] We will study the response of the economy to a shock to ξ in Section 5.

one, and involves comparing the value of remaining in the customer base of the current firm with the value of searching. Because of the Markovian structure of prices, the customer's expectation about future prices is completely determined by the firm's current productivity. We notice that if the continuation value is increasing in z (a sufficient condition is that $\hat{p}(z)$ is decreasing) then the threshold $\hat{\psi}(p, z)$ is decreasing in z.

2.3 Equilibrium

In this section we define an equilibrium, discuss its existence, and characterize its general properties. First we derive the equilibrium dynamics of the customer base as a function of price and productivity, given the optimal search and exit strategy of the customers. Given customers' optimal decision rule, the mass of customers buying from a firm with productivity z and charging price p is given by $\mathcal{M}(m, p, z) = m\, \Delta(p, z)$, with

$$\Delta(p, z) \equiv 1 - \underbrace{G\left(\hat{\psi}(p, z)\right)\left(1 - H(\bar{V}(p, z))\right)}_{\text{customer outflow}} + \underbrace{Q\left(\bar{V}(p, z)\right)}_{\text{customer inflow}}, \tag{9}$$

where $G(\hat{\psi}(p, z))$ is the fraction of customers searching from the firm customer base, a fraction $1 - H(\bar{V}(p, z))$ of which actually finds a better match and exits the customer base of the firm. The mass m is the probability that searching customers in the whole economy draw the firm as a potential match. The function $Q(\bar{V}(p, z))$ denotes the equilibrium mass of searching customers currently matched to a firm with continuation value smaller than $\bar{V}(p, z)$. Therefore, the product $m\, Q(\bar{V}(p, z))$ amounts to the mass of new customers entering the customer base. Equation (9) verifies the conjecture about the equilibrium customer dynamics made in Section 2.1.

We are now ready to define and discuss the equilibrium. We study equilibria where the continuation values to customers is non-decreasing in productivity, implying that customers' rank of firms coincides with their productivity. This is a natural outcome as more productive firms are better positioned to offer lower prices and therefore higher values to customers.

Definition 1 *Let $\mathcal{V}(z) \equiv \bar{V}(\hat{p}(z), z)$ and $p^*(z)$ be given by equation (5). We study stationary Markovian equilibria where $\mathcal{V}(z)$ is non-decreasing in z and $\hat{p}(z) \geq p^*(\bar{z})$ for all $z \in [\underline{z}, \bar{z}]$. A stationary equilibrium is then*

(i) search and exit strategies that solve the customer problem in equations (6)-(8);

(ii) a firm pricing strategy $\hat{p}(z)$ that solves equation (4) for each z;

(iii) a customer base for new entrant firms $m_0 = q/\kappa$, with $q = \kappa + \delta - \kappa\, \delta$;

(iv) a dynamic of the customer base at a surviving firm with productivity z given by $m' = (1-\delta)\,\Delta(\hat{p}(z), z)\, m$, where $\Delta(\cdot)$ is given by equation (9);

(v) an invariant distribution of customers $K(\cdot)$ over productivities, that for each z solves

$$K(z) = (1-q) \int_{\underline{z}}^{z} \int_{\underline{z}}^{\bar{z}} \Delta(\hat{p}(x), x)\, dF(s|x)\, dK(x) + q \int_{\underline{z}}^{z} d\bar{F}(x) \ ;$$

(vi) two invariant distributions, $H(\cdot)$ and $Q(\cdot)$, that solve

$$H(x) = K(\hat{z}(x)) \quad \text{and} \quad Q(x) = \int_{\underline{z}}^{\hat{z}(x)} G(\hat{\psi}(\hat{p}(z), z))\, dK(z) \ ,$$

for each $x \in [\mathcal{V}(\underline{z}), \mathcal{V}(\bar{z})]$, where $\hat{z}(x) = \max\{z \in [\underline{z}, \bar{z}] : \mathcal{V}(z) \le x\}$.

The next proposition states conditions under which the equilibrium that we evaluate exists and characterizes some of its properties.

Proposition 1 *Let productivity be i.i.d. with $F(z'|z_1) = F(z'|z_2)$ continuous and differentiable for any z' and any pair $(z_1, z_2) \in [\underline{z}, \bar{z}]$, and let $G(\psi)$ be differentiable for all $\psi \in [0, \infty)$, with $G(\cdot)$ differentiable and not degenerate at $\psi = 0$. There exists an equilibrium as in Definition 1 where $\hat{p}(z)$ satisfies equation (4), and*

(i) $\hat{p}(z)$ is strictly decreasing in z, with $\hat{p}(\bar{z}) = p^*(\bar{z})$ and $p^*(\bar{z}) < \hat{p}(z) < p^*(z)$ for $z < \bar{z}$, implying that $\mathcal{V}(z)$ is strictly increasing. Moreover, the optimal markups are given by

$$\mu(p, z) \equiv \frac{p}{w/z} = \frac{\varepsilon_d(p)}{\varepsilon_d(p) - 1 + \varepsilon_m(p, z)\, \Pi(p, z)/(d(p)\, p)} \ , \tag{10}$$

where $p = \hat{p}(z)$ for each z.

(ii) $\hat{\psi}(\hat{p}(z), z)$ is strictly decreasing in z, with $\hat{\psi}(\hat{p}(\bar{z}), \bar{z}) = 0$ and $\hat{\psi}(\hat{p}(z), z) > 0$ for $z < \bar{z}$, implying that $\Delta(\hat{p}(z), z)$ is strictly increasing, with $\Delta(\hat{p}(\bar{z}), \bar{z}) > 1$ and $\Delta(\hat{p}(\underline{z}), \underline{z}) < 1$.

The proof of the proposition can be found in Appendix A. Here we just point out that, while the results of Proposition 1 refer to the case of i.i.d. productivity shocks, numerical results in Section 4 show they hold even in the case of persistent productivity processes.

We now comment on the properties of the equilibrium highlighted in the proposition. The equilibrium is characterized by price dispersion: this is important, as price dispersion is what motivates customers to search. Price dispersion is driven by heterogeneity in firm productivity, as in Reinganum (1979), and by the level and dispersion of search frictions.[9] More productive firms charge lower prices and, therefore, offer higher continuation value to customers. If all the firms had the same productivity, Proposition 1 would imply a unique equilibrium where the price is that maximizing static profits, $p^*(z)$, and as a result the customer base of every firm would be constant.[10] The equilibrium is also characterized by dispersion in customer base growth: more productive firms grow faster, and there is a positive mass of lower productivity firms that have a shrinking customer base and a positive mass of higher productivity firms that are expanding their customer base.

Optimal markups in equation (10) depend on three distinct terms: $\varepsilon_d(p)$, $\varepsilon_m(p,z)$, and $\bar{\pi}(p,z) \equiv \Pi(p,z)/(d(p)\,p)$. The terms $\varepsilon_d(p)$ and $\varepsilon_m(p,z)$ represent the price elasticities of quantity purchased (per-customer) and of customer growth, respectively. We notice that the elasticity of total firm demand to the price, i.e. $m\,\Delta(p,z)\,d(p)$, is given by $\varepsilon_d(p)+\varepsilon_m(p,z)$. An increase in price reduces total current demand both because it reduces quantity per customer (*intensive margin effect*) and because it reduces the number of customers (*extensive margin effect*). Moreover, the optimal markup solves a dynamic problem as a loss in customers has persistent consequences for future demand due to the inertia in the customer base. This dynamic effect is captured by the term $\bar{\pi}(p,z)$, which measures the firm present discounted value of a customer scaled by the current revenues. It follows that active customer markets are associated with a strictly lower markup than the one that maximizes static profit; the lower the markup, the larger the product $\varepsilon_m(p,z)\,\bar{\pi}(p,z)$.

Finally, it is useful to discuss two interesting limiting cases of our model reported in the following corollary (see Appendix B for a proof).

Corollary 1 *Let search costs be scaled as $\psi \equiv n\,\tilde{\psi}$, where $n > 0$. That is, let the value function in equation (6) be* $\max\left\{\bar{V}(p,z)\,,\,\hat{V}(p,z) - n\tilde{\psi}\right\}$. *Let $\pi(p^*(\bar{z}),\underline{z}) > 0$ and the assumptions of Proposition 1 be satisfied.*

(1) Let $n \to \infty$. Then: (i) the optimal price maximizes static profits, i.e. $\hat{p}(z) = p^*(z)$ for all $z \in [\underline{z},\bar{z}]$, and (ii) there is no search in equilibrium.

(2) Let $n \to 0$. Then, (i) there is no price dispersion, i.e. $\hat{p}(z) = p^*(\bar{z})$ for all $z \in [\underline{z},\bar{z}]$, and (ii) there is no search in equilibrium.

These two limiting cases highlight the tight relationship between size of the search cost, competition for customers and price dispersion. The first limiting case concerns the equilibrium when we let search costs diverge to infinity. The model then reduces to one where

[9] For tractability we will abstract from (possible) equilibria where symmetric firms charge different prices.

[10] This special case is useful to understand our relation to Diamond's (1971) results. Our model delivers equilibrium price dispersion as a result of heterogeneity in productivity. If productivity was homogeneous as in Diamond (1971) the monopoly price would be the only equilibrium price.

customer base concerns are not present. Because the customer base is unresponsive to prices, the firm problem reverts to a standard price-setting problem commonly studied in the macroeconomics literature: the firm sets the price p, taking into account only its impact on static demand $d(p)$. In equilibrium, optimal prices maximize static profits, i.e. $\hat{p}(z) = p^*(z)$ for all $z \in [\underline{z}, \bar{z}]$. There is price dispersion, and there is no search in equilibrium. The second limiting case concerns the equilibrium when search costs become arbitrarily small. In this case, as the scale of search costs becomes arbitrarily small, equilibrium prices approach the static profit maximizing price of the most productive firm, $\hat{p}(\bar{z})$. As a result, there is no price dispersion and customers do not search.

2.4 Parametrization of the model

To quantify our model, we need to make parametric assumptions and pick parameter values. In Section 3, we present data which we use to guide our parameter choice. Here, we present our functional form assumptions and discuss the calibration of parameters on which our data provide no information.

To mirror the frequency of our data, we assume that a period in the model corresponds to a week. We fix the firm discount rate to $\beta = 0.9992$ corresponding to an annual discount factor of about 4%. We set the firm exit rate $\kappa = 0.0009$, corresponding to a yearly exit rate of 4.5% and consistent with estimates reported by Evans (1987) for U.S. firms. We assume that customers have logarithmic utility in consumption. Consumption is defined as a composite of two types of goods $c \equiv \left[d^{\frac{\theta-1}{\theta}} + n^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}$, with $\theta > 1$ governing the demand elasticity. The first good (that we label d) is supplied by firms facing product market frictions as described in this section; the other good (n) acts as a numeraire and it is sold in a frictionless centralized market. The customer budget constraint is given by $p\,d + n = I$, where I is the agent's nominal income, which we normalize to one.[11] We set the nominal wage equal to the price of the numeraire good, so that $w = 1$.[12] As a result, we obtain a standard downward sloping demand function $d(p) = I/P\,(p/P)^{-\theta}$ where $P = (p^{1-\theta} + 1)^{\frac{1}{1-\theta}}$ is the price of the consumption basket. We set θ so that the average elasticity of demand (including both extensive and intensive margins) is equal to 4, in the range of values standard in the macro literature (see Burstein and Hellwig (2007) for a discussion) and similar to the average across the product categories reported in Chevalier et al. (2003) who, like us, analyze grocery products.[13]

[11] In Appendix E we show that I can be derived based on a model of the labor market.

[12] This is equivalent to assume that the numeraire good n is produced by a competitive representative firm with linear production function and unitary labor productivity. See Appendix E for details.

[13] The average elasticity of demand is obtained by summing over the intensive and extensive margins at

We assume that the logarithm of idiosyncratic firm productivity evolves according to an AR(1) process, $\log(z') = \rho \log(z) + \varepsilon$, where ε is i.i.d. normally distributed, $\varepsilon \sim N(0, \sigma)$.[14] Finally, we assume that customers draw their search cost from a Gamma distribution with shape parameter ζ, and scale parameter λ. The Gamma is a flexible distribution and fits the assumptions we made over the G function in the specification of the model. In particular, for $\zeta > 1$, we obtain that the distribution of search costs is differentiable at $\psi = 0$.[15] The parameters governing the productivity process, ρ and σ, the Gamma distribution, ζ and λ, are picked to match statistics from micro data on customer and price dynamics described in the next section. Finally, we set the exogenous customer attrition rate δ so that the overall customer attrition rate in the model (summing random exits and those driven by prices) matches the fraction of the households in our data that experience an exit from the customer base of the supermarket chain discussed in the next section (22% on a yearly basis).

3 Data and estimation

We complement the theoretical analysis with an empirical investigation that relies on scanner data from a large U.S. supermarket chain. In this section we document that the price posted by a firm influences customers' decisions to leave the firm, therefore supporting the mechanism at the heart of our model. In Section 2.4, we will also use a measure of the size of this effect to estimate the model and quantify the importance of customer markets in shaping firm price setting.

3.1 Data sources and variable construction

The empirical analysis exploits two main pieces of data. The first dataset (henceforth, "retailer price data") has been previously used and documented by Eichenbaum et al. (2011). It consists of store level weekly[16] revenues and quantities for the full set of UPCs purchasable at the stores of a large US supermarket chain between the years 2004 and 2006. The chain is a major grocery retailer operating over one thousand outlets across 10 states. The data only cover a subset of the stores but we observe at least one store for each of the price areas

the firm level, and then aggregating over firms: $\int_{\underline{z}}^{\bar{z}} [\varepsilon_m(\hat{p}(z), z) + \varepsilon_d(\hat{p}(z))] \, dK(z)$.

[14] Operationally, we approximate the AR(1) through a discrete Markov chain with a methodology proposed by Tauchen (1986).

[15] In the estimation procedure we will discuss later we do not impose any constraints on the values the parameter ζ can take. Our unconstrained point estimate lies in the desired region.

[16] The retailer changes the price of the UPCs at most once per week, hence the frequency of the retailer price data captures the entire time variation.

setup by the retailer.[17] The products sold by the chain are mostly food (packaged and non packaged) and household supplies (detergents, personal hygiene products, etc.). The data also contain a measure of cost provided by the retailer for each UPC-store-week.[18] This represents the replacement cost for the chain, i.e. the cost for the retailer of restocking the product. It includes the wholesale price but also other costs associated with logistics (delivery to the store, etc.). Eichenbaum et al. (2011) treat this measure as a good approximation of the retailer's marginal cost.

The second and most crucial source of information we exploit (henceforth, "retailer consumers panel") is a companion dataset consisting of cashier register records on purchases by a panel of households carrying a loyalty card of the same chain providing the price data.[19] The empirical study of the interaction of both consumer shopping behavior and retail prices is one of the novel aspects of our analysis relative to the exiting literature. For every trip made to the chain between June 2004 and June 2006 by customers in the sample, we have information on the date of the trip, store visited, and list of all goods purchased (as identified by their UPC) at any store of the chain, as well as quantity and price paid. The customers in our sample make an average of 150 shopping trips to the chain over the two years; if those trips were uniformly distributed, that would imply visiting a store of the chain six times per month. The average expenditure per trip is $69 for the average household. There is a great deal of variation (the 10th percentile is $29; the 90th is $118) explained, among other things, by income and family size of the different households.[20]

We use these two databases to construct the variables needed to measure the comovement between the customer's decision to exit the customer base and the price of her typical basket of goods posted at the chain: (i) an indicator signaling when the household is exiting the chain's customer base, and (ii) the price of the household basket. Below we briefly describe the procedure followed to obtain them; the details are left to Appendix C.

We consider every customer shopping at the retailer in a given week as belonging to the chain's customer base in that week. We assume that a household has exited the customer base when she has not shopped at the chain for eight or more consecutive weeks. The decision

[17] Price areas are geographic areas where stores of the same chain post the same prices.

[18] The acronymous UPC stands for Universal Product Code.

[19] The chain is able to associate the loyalty cards belonging to different members of a same family to a single household identifying number, which is the unit of observation in our data. Therefore, in the analysis we use the terms "customer" and "household" interchangeably. The household identifier also allows us to track members of a same household when they lose (and replace) their individual loyalty card.

[20] Our data do not include information on purchases by customers not carrying a loyalty card of the chain. However, the chain pushes for high penetration of the loyalty card (only card holders can enjoy the price promotions) and most households use it regularly (Einav et al. (2010)). Moreover, the focus of this study, is on "regular" customers who can be meaningfully said to be part of the customer base of a firm. This seems to be more the case for individuals who sign up for a loyalty card than for occasional shoppers who do not.

to exit is imputed as the last time the customer visited the chain. Although brief spells without purchases can be justified with alternative explanations (e.g. consuming inventory or going on vacation), the typical customer is unlikely to experience an eight-week spell without shopping for reasons other than having switched to a different chain. In fact, for the average household in our sample, four days elapse between consecutive grocery trips and the 99th percentile of this statistic is 28 days, half the length of the absence we require before inferring that a household is buying its groceries at a competing chain. This suggests that the eight-week window is a conservative choice.

We construct the price of the basket of grocery goods usually purchased by the households in the following fashion. We identify the collection of UPCs belonging to a household's basket (K_i) using the retailer consumers panel and build the index as a weighted average of the price of such goods, with price information taken from the retailer price data. The price of the basket of customer i, shopping at store j in week t is then

$$p_{it}^{j} = \sum_{u \in K_i} \omega_{iu}\, p_{ut}^{j} \ , \qquad \omega_{iu} = \frac{\sum_t E_{iut}}{\sum_{u \in K_i} \sum_t E_{iut}} \ , \tag{11}$$

where p_{ut}^{j} is the price of UPC u in week t at the store j where customer i shops, and E_{iut} is the expenditure (in dollars) by customer i in UPC u in week t. Note that the price of the basket is household specific because households differ in their choice of grocery products (K_i) and in the weight such goods have in their budget (ω_{iu}).

This brief description highlights how our data are well suited to study customer base dynamics. Not only are they rich in detail about households' grocery consumption (products purchased, expenditure, prices, etc.) but they also include a panel dimension which is crucial to observe the evolution of the customer base of the supermarket chain. Furthermore, given the significant market share held by the chain, its wide geographical spread and the fact that it offers all the mainstream grocery packaged good products, the behavior of its customers provides insights that are likely to generalize to the population of retail shoppers at large. At the same time, exploiting data from a single retail chain has limitations preventing us from investigating other relevant questions related to the topic of this study. The most obvious one is that we no longer track customers once they leave the chain. Hence, we cannot provide any empirical insights on their post-exit behavior or on the characteristics of the firms they join.

Finally, it is also worth addressing that the settings analyzed in the model and our application differ in some respects. In fact, in the theoretical model we studied the behavior of customers buying from firms producing a single homogeneous good; our application documents the exit decisions of customers from supermarket stores where they buy bundles of goods.[21] However, under the assumption that customers' behavior depends on the price of the whole basket of goods they typically buy at the supermarket, we can focus on the resulting price index of the customer basket in order to retrieve the response of the customer base to variation in prices. Since customers baskets are largely composed of package goods, which are standardized products, the assumption that the basket is a homogenous good is not unwarranted.

3.2 Evidence on customer base dynamics

In Figure 2 we plot the survival function for our sample of customers, that is the probability of remaining in the customer base of the firm as a function of the length of the household spell as a customer. In the plot, we explore the sensitivity of customer base evolution to our definition of exit by displaying two survival functions. The solid line refers to our baseline definition; the dashed line represents the survival probability if we extend to three months the absence spell required to determine that the customer has exited. The first thing to notice is that, regardless of the definition of exit we adopt, the odds of exiting the customer base evolve smoothly. The second noteworthy fact emerging from the plot is that the customer base is quite sticky: the probability of never exiting the customer base in 90 weeks is between 70% and 80%. These values lend support to industry estimates on the customer attrition rate surveyed in Gourio and Rudanko (2014).

To sharpen our understanding of the determinants of a customer's decision to exit the customer base of the firm she is currently shopping from, we estimate a linear probability model where the dependent variable is an indicator for whether the customer has left the customer base of the chain in a particular week. Our aim is to capture the effect of the price posted by the chain for the basket of goods purchased by the customer on her decision to exit. In Table 1, we report results of regressions of the following form,

$$Exit_{it} = b_0 + b_1 \log(p_{it}) + b_2 \log(\bar{p}_{it}) + b_3 tenure_{it} + X_i' c + \varepsilon_{it} \ . \tag{12}$$

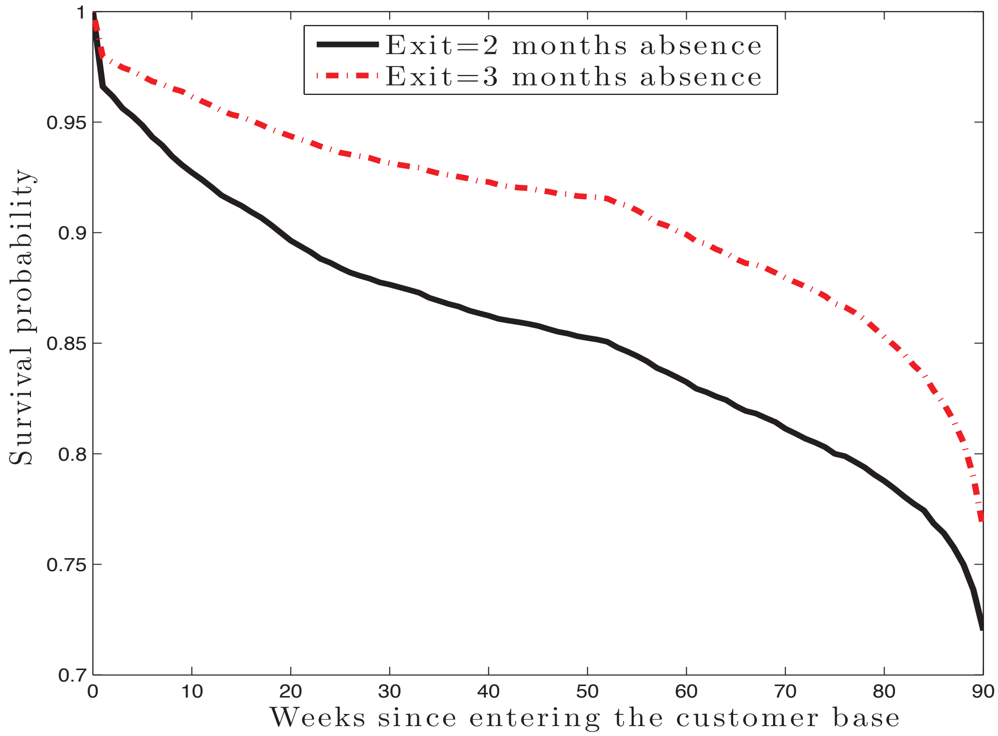
Our main interest is on the coefficient of the retailer price of the basket, b_1, which is a direct measure of the average extensive margin elasticity in our model.[22,23] Existing theories on the

[21] The choice of focusing on the customer base of the store rather than that of one of the branded product it sells is data driven. With data from a single chain we cannot track the evolution of the customer base of a single brand. In fact, if we observed customers no longer buying a particular brand we could only infer that they are not buying it at the chain we analyze, but we could not rule out that they are buying it elsewhere.

[22] To ease notation, we have dropped the j superscript: it is implicit that p_{it} is the price of the basket purchased by consumer i in week t at the store j where she usually shops.

[23] The actual extensive margin elasticity in our model also includes the elasticity of the customer inflow to prices. However, we can claim that b_1 maps into our measure of elasticity because in the model the price

Figure 2: Survival in the customer base



Notes: The figure plots the survival function for our sample of households, where failure is defined as exit from the customer base. We report two survival functions for different criteria to determine whether the customer has exited: two consecutive months without shopping at the chain (our baseline definition, continuous black line), and three consecutive months without shopping at the chain (dashed red line). To ensure that all individuals have similar potential length for their spells, we only consider the first spell as customer for those having multiple ones and we only retain households whose first trip at the chain occurs within the first 40 weeks in our sample.

link between prices and customer dynamics (Phelps and Winter (1970)) stress that a firm's ability to retain its customers should be influenced by its idiosyncratic price variations but not from aggregate shocks that move the competitors' prices as well. To isolate idiosyncratic price variations, we control for the prices posted by the competitors in the same market of the chain using information from the IRI Marketing data set.[24] This data allows us to compute for every customer the average (cross retailers) price of her basket in the market where she lives ($\bar{p}_{it}$). To further control for sources of aggregate variation, we include in the regression year-week fixed effects that account for time-varying drivers of the decision of exiting the customer base common across households (e.g., disappearances due to travel during holiday

affects the probability that the customer switches, but not the arrival rate.

[24] A detailed description of the data can be found in Bronnenberg et al. (2008). All estimates and analyses in this paper based on Information Resources Inc. data are by the authors and not by Information Resources Inc. We provide additional details on the IRI data and on the construction of the price index for the competitors of the chain in Appendix C.

season).

The coefficient b_1 is then identified by *UPC-chain* specific shocks as those triggered, for example, by the expiration of a contract between the chain and the manufacturer of a UPC. We also observe the price of a same good moving differently in different stores within the chain, for instance due to variation in the cost of supplying the store linked to logistics (e.g. fluctuations in the price of gas affect differently stores at different distance from the warehouse). Since these shocks can hit differently goods with different intensity in delivery cost (e.g. refrigerated vs. non-refrigerated goods), *UPC-store* specific shocks also contribute to our identification. We do not need to assume that such shocks will make a supermarket uniformly more expensive than the competition. Shocks that affect the convenience of a chain with respect to a subset of goods suffice to induce the customers who particularly care about those goods to leave. Kaplan and Menzio (2015) use a different scanner data to provide ample evidence for this type of variation. They report that the bulk of price dispersion arises not from the difference between high-price and low-price stores but from dispersion in the price of a particular good (or product category) even among stores with similar overall price level. Since the retailer price in equations (12) can be endogenous if the chain conditions to variables unobserved to the econometrician that also influence the customer's decision to leave, we instrument it with the cost of the basket. Similarly to what we did to obtain the price of the basket, we calculate the cost of a basket obtained as the weighted average of the replacement cost of the UPCs included in it.

In our empirical specification we acknowledge that, unlike posited in the model, customers are heterogenous in more dimensions than their cost to search. The limited number of exits occurring in our sample implies that the within unit variation in the dependent variable is low. Therefore, we cannot control nonparametrically for cross-household heterogeneity using household or store fixed effects. Instead, we include in our specification a rich set of covariates that control for the main characteristics affecting store choice: demographics, location, market characteristics, and tenure. The demographic variables (age, income, and education) are matched from Census 2000. We calculate, using data on grocery shop location by Reference US, both the distance (in miles) between a household's residence and the closest store of the chain and that to the closest supermarket of a competing brand. We account for market structure by controlling for the total number of supermarket stores in the zip code of residence of the customer. To pick up the heterogeneity in the type of goods different customers include in their basket, we control for the price volatility of the customer-specific basket and for its price in the first week in the sample, as a scaling factor. Finally, we calculate customer tenure, defined as the number of consecutive weeks the customer has spent in the customer base of the chain, and include it in the regression to account for the fact that long-term customers of the chain may be less willing to leave it *ceteris paribus*.

Table 1: Effect of the price of the basket on the probability of exiting the customer base

Exiting: Missing at least 8 consecutive weeks	**(1)**	**(2)**	**(3)**	**(4)**
$\log(p_{it})$	0.14**	0.16**	0.15**	
	(0.066)	(0.080)	(0.064)	
$\log(p_{it})$*Walmart entry		0.018**		
		(0.009)		
$\log(\bar{p}_{it})$	0.001		0.001	0.000
	(0.001)		(0.001)	(0.001)
$\log(p_t^{j(i)})$				0.01
				(0.001)
Tenure	-0.002***	-0.003***	-0.004***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)
Observations	52,670	66,182	52,101	52,670

Notes: An observation is a household-week pair. The results reported are calculated through two-stage least squares where we use the logarithm of the cost of the basket (constructed based on the replacement cost provided for each UPC by the retailer) as instrument for the logarithm of the price of the basket. In column (3), the exit of the customer is assumed to have occurred in the first week of absence in the eight (or more) weeks spell without purchase at the chain rather than the week of the last shopping trip before the hiatus. In column (4), the price of the household basket is substituted with a price index for the store where the customer shops (identical for all the customers shopping at the same store). We trim from the sample households in the top and bottom 1% in the distribution of the number of trips over the two years. Coefficients on a series of variables are not reported for brevity: demographic controls matched from Census 2000 (ethnicity, family status, age, income, education, and time spent commuting) as well as distance from the closest outlet of the supermarket chain and distance from the closest competing supermarket (provided by the retailer). The logarithm of the price of the household basket in the first week in the sample and the standard deviation of changes in the log-price of the household basket over the sample period are included as a controls in all specifications. Week-year fixed effects are also always included. Standard errors are in parenthesis and account for within-household correlation through a two-steps feasible-GLS estimator. ***: Significant at 1% **: Significant at 5% *: Significant at 10%.

The main specification in column (1) shows that the basket price posted by the retailer significantly impacts the probability of leaving. The effect is also quantitatively important. A weekly price elasticity of the customer base equal to 0.14 implies that if the retailer's prices were 1% higher for a full year, the customer base would decrease 7%. The coefficient on the competitors' price, which we would expect to enter with a negative sign, is not significant. This may be due to the fact that the IRI data only allow us to imperfectly capture competitors' behavior. In fact, the IRI dataset contains price information only on a subset of the goods included in a customer's basket, although it arguably covers all the major product categories. Furthermore, the IRI data do not contain detailed information on the location of the outlets. This introduces measurement error in our construction of the set of stores a customer considers as options for her shopping. The negative coefficient on tenure confirms the intuition that the longer the relationship between a firm and a customer, the less likely they are to be interrupted. Among the several individual characteristics we control for, it is worth mentioning that distance from stores of the chain and distance from the closest competing store enter with the expected sign. Customers living in proximity of a store of the chain are less likely to leave it, and those living closer to competitors' stores are more inclined to do so.

Documenting that changes in the price of the goods included in their regular consumption basket affect customers' decision to abandon the chain not only provides a compelling motivation to our decision to model the link between customer base and pricing policy but also lends support to the central tenet of the growing literature on customer markets. Preexisting evidence of this relationship is based on survey data where firms report concerns about customer retention as the main reason for their reluctance to adjust prices (see Blinder et al. (1998), and Fabiani et al. (2007)). To the best of our knowledge, we are the first to document this fact using micro data based on actual customers' decisions.

Additional columns in Table 1 present robustness checks of our main result. In column (2), we experiment with an alternative way to control for the effect of competition: we exploit episodes of entry by Walmart, a major retailer with which our chain is in direct competition. We use data from Holmes (2011) to identify the date of entry by a Walmart supercenter, i.e. a store selling groceries on top of general discount goods, in a zip code where the retailer we study also operates a supermarket. The resulting event study allows us to measure the effect of our retailer price on the probability of exit controlling for the most relevant change in the competitive environment. The coefficient obtained falls in the same ballpark as the estimate in the main specification, which is reassuring on the effectiveness of the IRI price in measuring the competitors' behavior. In column (3), we modify the assumption on the imputation of the date of exit. Rather than assuming that the customer left on the occasion of her last trip to the store, we posit that the exit occurred in the first week of her absence. Even in this case, the main result is unaffected. In column (4), we replace the price of the individual basket with a price index for the store where she buys ($p_t^{j(i)}$). The store price is a price index of a composite bundles of goods for each store so to accommodate the multiproduct nature of grocery retailing (Smith (2004)) and its construction resembles that of the price index for the customer individual baskets. It is calculated as the average of the prices of the goods sold by the store, weighted by the amount of revenue they generate.[25] Formally the price

[25]In principle, we would want to include the prices of all the UPCs carried by the store. In practice, this

index for store j in week t is:

$$p_t^j = \sum_{u \in A_j} \omega_u^j \, p_{ut}^j \;, \qquad \omega_u^j = \frac{\sum_t R_{ut}^j}{\sum_{u \in A_j} \sum_t R_{ut}^j} \;, \tag{13}$$

where p_{ut}^j is the price of UPC u in week t at store j, A_j is the set of goods in assortment at store j and R_{ut}^j are revenues from UPC k to store j in week t. By construction, $p_t^{j(i)}$ is identical for all the customers shopping in the same store.

The coefficient on $p_t^{j(i)}$ is negative and not significant, confirming the importance of being able to construct individual specific baskets in order to make inference on customers' behavior. Finally, we performed a placebo test to investigate whether it is possible to obtain results with the same level of significance of our main specification out of pure chance. We estimated our main specification 1,000 times each time with a different dependent variable where exits from the customer base, while kept constant in number, are randomly assigned to customers. We find that only in 2.8% of the cases the simulation yields a price coefficient that is positive and significant at 5%.

3.3 Estimation

Our data allow us to estimate the key parameters of the model -those characterizing the idiosyncratic productivity process and the search cost distribution- using a minimum-distance estimator. The productivity process influences the variability of prices, which is necessary for customers to obtain any benefit from search. The parameters of the search cost distribution, on the other hand, directly determine how costly it is to search. Below we discuss the moments in the data that allow us to pin down those parameters. The discussion is provided only for the sake of intuition; given the nonlinearity of the model, all the moments contribute to the identification of all the parameters.

Our theoretical model describes how persistence and volatility of productivity (ρ and σ, respectively) determine autocorrelation and volatility of the resulting firm prices. We therefore estimate ρ and σ by matching the autocorrelation and the volatility of the logarithm of firm prices observed in the data. In practice, we estimate an AR(1) process for the store price index defined in equation (13), controlling for time and store fixed effects. We find that the autocorrelation of log-prices in the data is equal to 0.61. For the volatility of the process, we target the unconditional weekly standard deviation of the residuals, which gives us an estimate of the volatility of the store price index (0.027).

is not possible because the information on price is missing for some UPCs in certain store-weeks. Therefore, the price index for the store is computed using a constant set of UPCs for which we have a complete time series of prices at the store during our sample.

To identify the parameters of the search cost distribution we exploit the estimates of the relationship between the price and the probability of exiting the customer base discussed in the previous section. Recall that we have assumed that the search cost ψ is i.i.d. distributed with a cumulative distribution function $G(\psi) \sim Gamma(\zeta, \lambda)$. The parameter λ governs the scale of the search cost distribution. A higher λ implies higher search cost on average and lower propensity to search of customers (see the discussion in Section 2.3). Thus, the parameter λ speaks directly to the size of the extensive margin elasticity of our model, and we estimate it by matching the average effect of log-prices on the exit probability predicted by the model to its counterpart in the data, measured by the estimate of the parameter b_1 reported in column (1) of Table 1.[26]

The parameter ζ measures the inverse of the coefficient of variation of the search cost distribution: the larger ζ, the flatter the density. A price increase raises the search cost threshold determining the customer indifferent between searching and not searching ($\hat{\psi}(z)$), and triggers an increase in the mass of customers searching ($G(\hat{\psi}(z))$). However, such variation in G is smaller the flatter the search cost density. Therefore, larger values of ζ are associated with smaller variability in the extensive margin elasticity of demand to price. Since in our estimated model the elasticity of demand is negatively correlated with firm productivity, a smaller variability of the demand elasticity also induces higher average pass-through of productivity shocks. For instance, in the extreme case where the demand elasticity is constant, the optimal price is a constant markup over marginal cost and the pass-through is complete. Thus, we can pin down ζ using information on the price pass-through of cost shocks: high pass-through signals a high value for the shape parameter ζ. To this end, we estimate pass-through using the retailer price data by regressing the log-price index of each store in a given week on its log-cost index. We obtain a pass-through of 17% which we use as target for the estimation (see Appendix H for details on the estimation procedure).[27]

Finally, we set the exogenous customer attrition rate δ so that the overall customer attrition rate in the model (summing random exits and those driven by prices) matches the fraction of the households in our sample that experience an exit from the customer base of the supermarket chain (0.0044 on a weekly basis).

We define $\Omega \equiv [\zeta \ \lambda \ \rho \ \sigma]'$ as the vector of parameters of interest and estimate it with a

[26] It is important to recall that this parameter was estimated controlling for a number of household-specific characteristics (demographics, distance from the store, etc.) which affect the decision of the store where to shop.

[27] Our finding is not inconsistent with evidence of complete pass-through presented by Eichenbaum et al. (2011) using the same data. First, they measure pass-through conditional on price adjustment; whereas we look at the unconditional correlation between prices and costs. Second, they deal with UPC-level pass-through while we measure pass-through of a basket of goods. If retailers play strategically with the pricing of different products, for example lowering margins on some UPC to compensate the cost increase they experienced on others, we can obtain both high UPC-level pass-through and low basket-level pass-through.

Table 2: Parameter estimates

	Estimates
Persistence of productivity process, ρ	0.80 [0.70, 0.90]
Volatility of productivity innovations, σ	0.15 [0.11, 0.20]
Scale parameter of search cost distribution, λ	0.375 [0.16, 0.47]
Shape parameter of search cost distribution, ζ	2.00 [1.5, 2.5]

Notes: 99% confidence intervals reported in parenthesis are computed by block bootstrap.

minimum-distance estimator. We denote by $v(\Omega)$ the vector of the moments predicted by the model as a function of parameters in Ω, and by v_d the vector of their empirical counterparts. The n^{th} iteration of the estimation procedure unfolds according to the following steps:

1. Pick values for the parameters ρ_n, σ_n, λ_n and ζ_n from a given grid,
2. Solve the model and obtain the vector $v(\Omega_n)$,
3. Evaluate the objective function $(v_d - v(\Omega_n))'\Sigma\ (v_d - v(\Omega_n))$. Where Σ is a weighting matrix that we assume to be the identity matrix.

We select as estimates the parameter values from the proposed grid that minimize the objective function. Implementing step 2 requires solving a fixed point problem in equilibrium prices. In particular, given our definition of equilibrium and the results of Proposition 1, we look for equilibria where prices are in the interval $[p^*(\bar{z}), p^*(\underline{z})]$. In principle, our model could have multiple equilibria; however, numerically we always converge to the same equilibrium despite starting from different initial conditions. More details are provided in Appendix D.

The estimation results are summarized in Table 2. Persistence and volatility of the productivity process are easy to interpret. To provide context for the parameters of the search cost distribution, we perform a back-of-envelope calculation of the dollar equivalent value of the search costs the estimated distribution implies. The average search cost for agents who search is equivalent to a 1.8% permanent reduction in weekly income.[28]

[28] Assuming a working week of 40 hours, this is equivalent to a time search cost of about 40 minutes.

4 Quantitative implications: the distribution of prices

In this section we use the parameter estimates reported in Table 2 to characterize the quantitative properties of our model (henceforth "baseline economy"). We show that the model is able to produce predictions in line with recent empirical evidence on the distribution of price levels. Moreover, we perform counterfactual exercises to highlight the role of customer dynamics for equilibrium prices. In particular, we vary the size of the search costs directly affecting the extensive margin elasticity of demand.

Panel (a) of Figure 3 reports the distribution of (log-) price levels implied by the model. The distribution of prices shows a high concentration of prices around the mean, and relatively fat tails. As a result excess kurtosis is 5.3. These characteristics are similar to those documented by empirical studies looking at the distribution of prices. In particular, Kaplan and Menzio (2015) report a shape for the price distribution of homogeneous goods in the grocery sector similar to the one delivered by our model. The empirical distribution is characterized by a high concentration of prices around the mean, much higher than what implied, for instance, by a normal distribution.

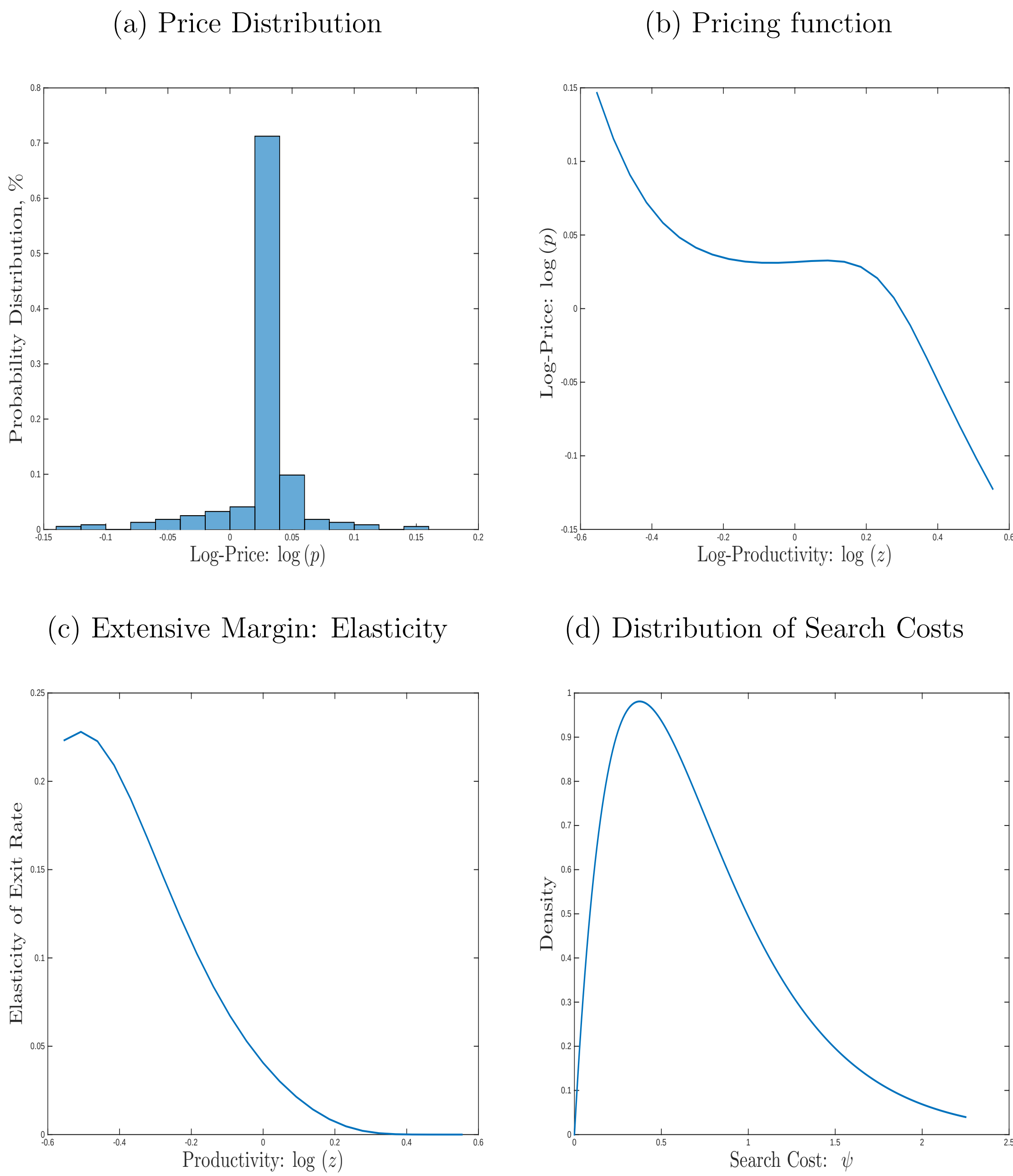
To understand the properties of the price distribution, we need to look at the relationship between the pricing function and the extensive margin of demand. The price distribution depends on the equilibrium pricing function, whose shape can be explained by the behavior of extensive margin of demand.[29] We report all of these objects in additional panels of Figure 3. The pricing function (panel (b)) is almost flat at intermediate levels of productivity thus explaining the high concentration of prices. At low and high levels of productivity, instead, the pricing function is steeper implying a higher pass-through of productivity shocks to prices. This accounts for the fat tails of the price distribution.

The pricing function is in turn shaped by the extensive margin of demand. Panel (c) of Figure 3 shows that the extensive margin elasticity steeply decreases with productivity at intermediate levels of productivity. Therefore the optimal markup increases with productivity so that the pass-through of productivity to price is incomplete. As can be seen in equation (10), an increase in production cost is then offset by a reduction in markups explaining the flatness of the pricing function in that region. This effect is more marked, the stronger is the relationship between productivity and demand elasticity.

The most productive firms face low risk that customers will leave since they offer high expected value to their customers relative to the average firm. As a result, they have low exit rates: only customers drawing both tiny search costs and an even better match will end up

[29] Equation (10) implies that optimal prices also depend on the intensive margin elasticity (ε_d) and the value of a customer ($\bar{\pi}$). Their role is however quantitatively small. Therefore, here we concentrate only on the role of the extensive margin of demand.

Figure 3: Equilibrium price dynamics



Notes: The histogram in panel (a) portrays the distribution of prices grouped into bins of size equal to 0.02 log-points. In panel (b), we plot the optimal log-prices as a function of productivity. In panel (c), we plot the extensive margin elasticity. In panel (d) we plot the density of search costs. All the objects refer to the baseline model whose parameters are reported in Table 2.

exiting. This combination is a low probability event, roughly insensitive to small variations in productivity because the estimated density of search costs has little mass close to the origin (panel (d)) meaning that variations in the threshold to search are associated to small variations in the mass of customers searching.[30]

As productivity decreases, the threshold to search as well as the probability of drawing a better match increase. Moreover, the decision to search and exit becomes more sensitive to changes in productivity because the estimated density of search costs increases, implying that variations in the threshold to search are associated with significant variations in the mass of customers searching. Finally, as productivity approaches the left tail of the distribution, the extensive margin elasticity flattens again. This happens because customers paying higher prices at lower productivity firms substitute towards the numeraire good (good n). Therefore, everything else being equal, variations in the price of the good with product market friction (good d) have less of an impact on the utility of these customers.[31]

To sum up, the incomplete pass-through of idiosyncratic cost shocks to prices and high clustering of prices around the mean are tightly linked in our model to variation in the elasticity of demand that different firms face. In particular, variation in the elasticity of the extensive margin of demand is what truly drives those results. This is a distinctive characteristic of our model relatively to alternative models of endogenous markups where demand elasticity varies because of the interaction of variation in market share with the intensive margin of demand (see for instance Atkeson and Burstein (2008)). As we discussed above, in our model the variation in the elasticity of demand may not be necessarily associated with the variation in market share. In fact, firms at intermediate levels of productivity, and therefore with intermediate market shares on average, face steeper variation in elasticity than the least productive firms. This feature seems to find indirect confirmation in the data, since it is key to obtain excess kurtosis of the distribution of prices.

To further illustrate how customer markets play a key role in generating the distribution of prices we obtain, we perform a counterfactual exercise varying their strength. In Figure 4 we report the price distributions (left panel) and the pricing functions (right panel) from three economies identical but for the level of the search cost. The first is our baseline economy, which we compare to an alternative where we increase the search costs by a factor of two and another one where we increase the search costs by a factor of five.

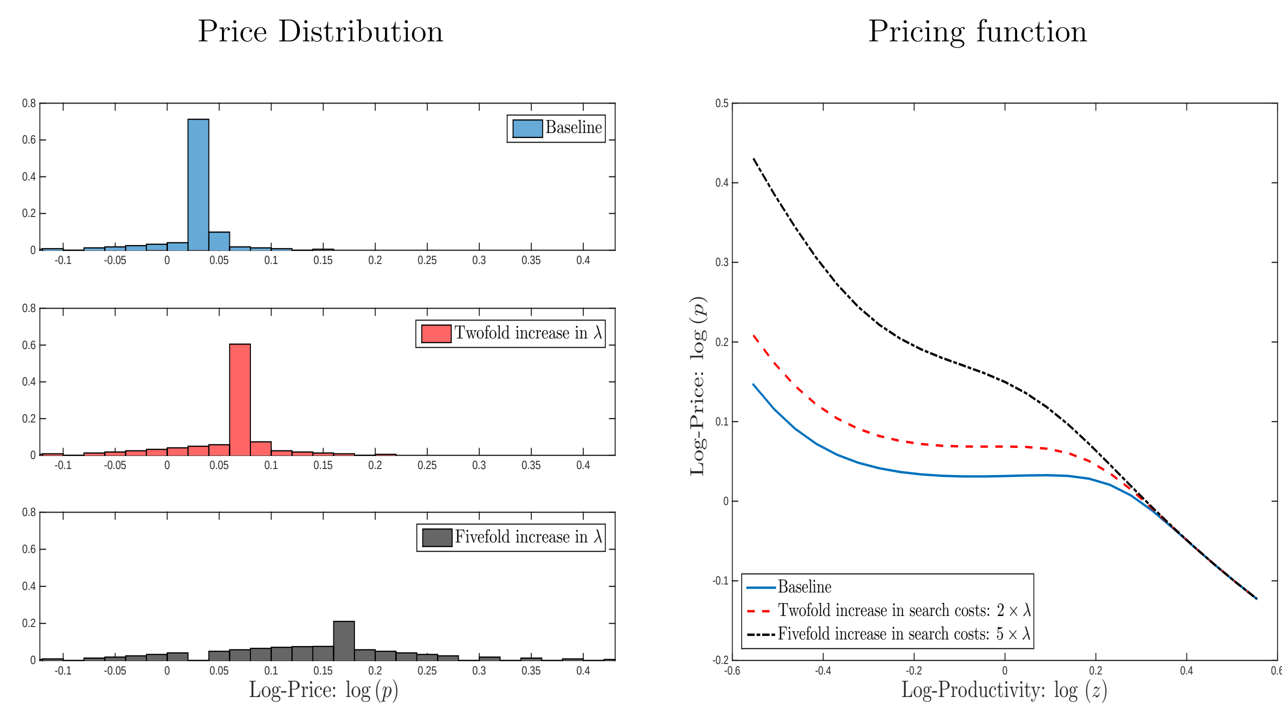
As we increase the scale of the search costs (increasing λ), it is harder for customers to exit, therefore competition for customers declines. This results in a price distribution characterized by a higher mean, a higher standard deviation and a lower kurtosis. All of these effects are due to the impact of higher search costs on the optimal pricing function.

[30] In Appendix H, we present evidence consistent with the fact that higher productivity firms display higher cost pass-through.

[31] Notice that the threshold to search at low productivity firms is in the increasing part of the density of search cost, so the flattening of the extensive margin elasticity at low productivity is not due to a smaller mass of customers exiting at the margin.

When the scale of the search cost grows, firms charge higher price for each level of productivity and pass-through productivity shocks more. This reflects that the extensive margin elasticity becomes less sensitive to productivity changes when customers search costs are higher. In fact, as we showed in Corollary 1, the model then converges to monopolistic competition where the extensive margin elasticity is constant.

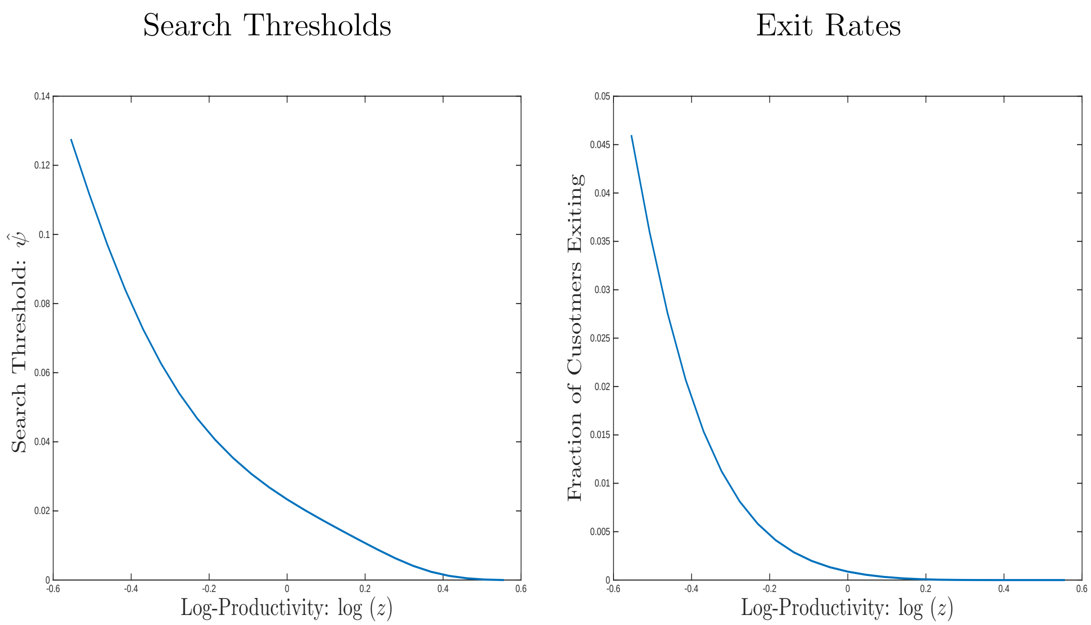
Figure 4: Comparative statics: the role of search costs



Notes: In the left panel, we plot the distribution of log-prices and in the right panel the optimal pricing function as a function of productivity.The blue solid line and histogram refer to our baseline economy with parameters reported in Table 2. The black and red objects refer to two counterfactuals where we increase the size of search costs ($\lambda = 1.65$ & $\lambda = 0.7$ respectively).

Finally, in Figure 5 we illustrate the predictions of our model regarding customer dynamics. We plot the search cost thresholds (left panel) and the fraction of customers exiting the customer base (right panel) as a function of the productivity z of the firm the customer is matched to. There is substantial variation in exit rates across firms. Firms above the median in terms of productivity lose less than 0.25% of their customers every week but low productivity firms can see their customer base erode significantly. Recent literature has emphasized the role of the customer base as an important and persistent determinant of the level of firm demand (Foster et al. (2008), Foster et al. (2016)). Our findings are fully consistent with this view. In fact, Table 3 reports in the top row the probability that a firm in the top 25% of the distribution of demand is still in the same quartile after 1 month, 1 quarter and 1 year, re-

Figure 5: Equilibrium customer dynamics



Notes: In panel (a), we plot the threshold for the search cost, below which customers search as a function of productivity. In panel (b) we plot the fraction of customers exiting the customer base of a firm. All the objects refer to the baseline model whose parameters are reported in Table 2.

Table 3: Persistence of demand

	Total sales: $d \times m$			Productivity: z		
	1 month	1 quarter	1 year	1 month	1 quarter	1 year
Top 25%	0.84	0.81	0.75	0.32	0.26	0.25
Bottom 25%	0.90	0.88	0.84	0.32	0.26	0.25

Notes: The top row reports the probability that a firm in the top quartile of the sales distribution stays there after 1 month, 1 quarter and 1 year. The bottom row displays the same statistics for the bottom quartile of the distribution. We report separately the results for the distribution of total sales (left panel) and for the distribution of productivity (right panel). The statistics are obtained by simulating our baseline economy with parameters reported in Table 2.

spectively. The bottom row shows the same statistics for the bottom 25% of the distribution. We look at two different outcomes: total sales (left panel) and productivity (right panel). The latter can be linked to the intensive margin of demand $d(p)$: each customer expands or contracts her demand for the customer market good depending on the price she faces, the persistence of which only depends on productivity. The total demand for the firm, however, is the sum of the demands of all of its customers. Even firms with declining productivity, which will post higher prices and, therefore, see their demand per customer shrink, can have strong total demand if they have a large base of customers. Our statistics do show that firm level demand is very persistent, much more so than the underlying exogenous process of firm productivity. For instance, a firm in the top 25% of the demand distribution is more that 80% likely to stay in the same group after a quarter; whereas a firm in the top 25% of the distribution of productivity only has a 26% chance of staying in the same group after one quarter. We also find asymmetry in persistence between the top and bottom quartiles of the distribution of demand: persistence at the bottom is stronger than persistence at the top. This happens because low ranked sellers have on average low productivity and lose customers at a faster rate than that at which high productivity firms, i.e. top sellers, gain them.

5 Aggregate demand, competition and price dispersion

Starting with Phelps and Winter (1970), the role of customer markets in shaping aggregate markup dynamics, in particular regarding the propagation of demand shocks (Rotemberg and Woodford (1991, 1999)), has been actively debated. Our model offers new insights to the analysis of how aggregate shocks affect markups and output. First, we show that by endogenizing the extensive margin of demand we obtain implications on the cyclical relationship between average markups, customers' search intensity and demand which substantially differ from that of traditional customer market models: markups are reduced in response to positive demand shocks. Second, our framework delivers a relationship between the extent of competition and the degree of price dispersion in a market. Given the wide and increasing availability of detailed data on prices, this provides a handy tool to measure the cyclicality of competition and, therefore, markups.

5.1 Demand shocks in a customer market economy

We consider our baseline economy in steady state at $t = t_0$, calibrated as described in Section 2.4 and Table 2, and we augment the model to endogenize household income I and to accommodate shocks to the aggregate state. In particular, we consider the dynamics following an unforeseen aggregate shock that takes the economy temporarily away from the steady state and assume that after the shock has hit, there is perfect foresight in the path of the aggregate state. This experiment requires the model to allow for the possibility that the aggregate state varies over time. Therefore, the key equations of the model will now be indexed by a time subscript t, capturing the dynamics in the aggregate state. As we want to study the effects of aggregate shocks in general equilibrium, we also need to endogenize household income. We do so by adding a simple model of a perfectly competitive labor market, where the household trades off (linear) utility from leisure with labor income. The household takes the wage as given, and the wage is determined in a centralized market to clear labor demand.[32]

We assume that the representative household is divided into a mass one of shoppers/customers and a representative worker. The worker takes care of supplying labor in perfectly competitive labor market, and then shares labor income equally across the shoppers who instead take care of buying goods according to the model described in Section 2.[33] The expected discounted utility of the household is given by

$$\int_0^1 V_t(p(i), z(i), \psi(i))\, di - \sum_{T=t}^{\infty} \beta^{T-t} \frac{\ell_T^{1+\phi}}{1+\phi} \quad , \tag{14}$$

where $V_t(\cdot)$ is the equivalent to equation (6) augmented with a subscript t to account for variation in the aggregate state, and represents the expected discounted utility from consumption. The index i denotes a customer, and the household utility from consumption is obtained by integrating over all agents i. For simplicity, we set $\phi = 0$ so to have linear disutility in labor. The worker chooses the path of labor supply ℓ_t that maximizes household preferences in equation (14). In particular the worker trades off higher disutility of labor ℓ_t with higher labor income $w_t\,\ell_t$ to be distributed equally across all customers. Thus, the worker internalizes the impact that higher labor income will have on the shopper decisions both in terms of search activity and consumption allocation, but cannot discriminate across shoppers.

The production technology of the good sold in the perfectly competitively market (good n) is linear in labor, with unitary productivity. Perfect competition in the market for variety n and in the labor market implies that workers are paid a wage equal to the marginal productivity of labor so that $w_t = 1$ for all t.

We model the aggregate shock as a once and for all unforeseen preference shock to the consumption/leisure margin that shifts the utility of consumption, i.e. the function $v(\cdot)$ in equation (7), by a multiplicative factor $\xi_t = \hat{\xi} > 1$ at $t = t_0$. It is important to emphasize that the qualitative results of this experiment extend to other types of demand shocks. For instance, in Appendix F we consider government spending shocks as well as preference shocks to the consumption/saving margin, and obtain similar results. The relevant characteristics

[32] Additional details on how we augment the model to study aggregate shocks and on the specifics of our stylized labour market model are provided in Appendix E.

[33] We are implicitly assuming that the worker cannot discriminate across the different shoppers. This assumption reduces the dimensionality of the problem, removing heterogeneity in income across consumers. This is a common shortcut in the literature (see Shi (1997)).

of these shocks for our analysis is that they all increase the marginal utility of consumption. After the shock, the preference parameter ξ_t mean reverts to its steady state value of 1 according to an AR(1) process, i.e. $\log(\xi_t) = \rho_\xi \log(\xi_{t-1})$ with $\rho_\xi < 1$ for $t > t_0$. Operationally, we guess the path of the distribution of customers over productivity $K_t(z)$, and then solve backward for the optimal pricing and customer decisions at each time $t \geq t_0$ as described in Appendix E starting from the final steady state (which is identical to the initial one) at $t = t_0 + T$ for some large T. Then we update our guess about $K_t(z)$ and iterate until convergence.

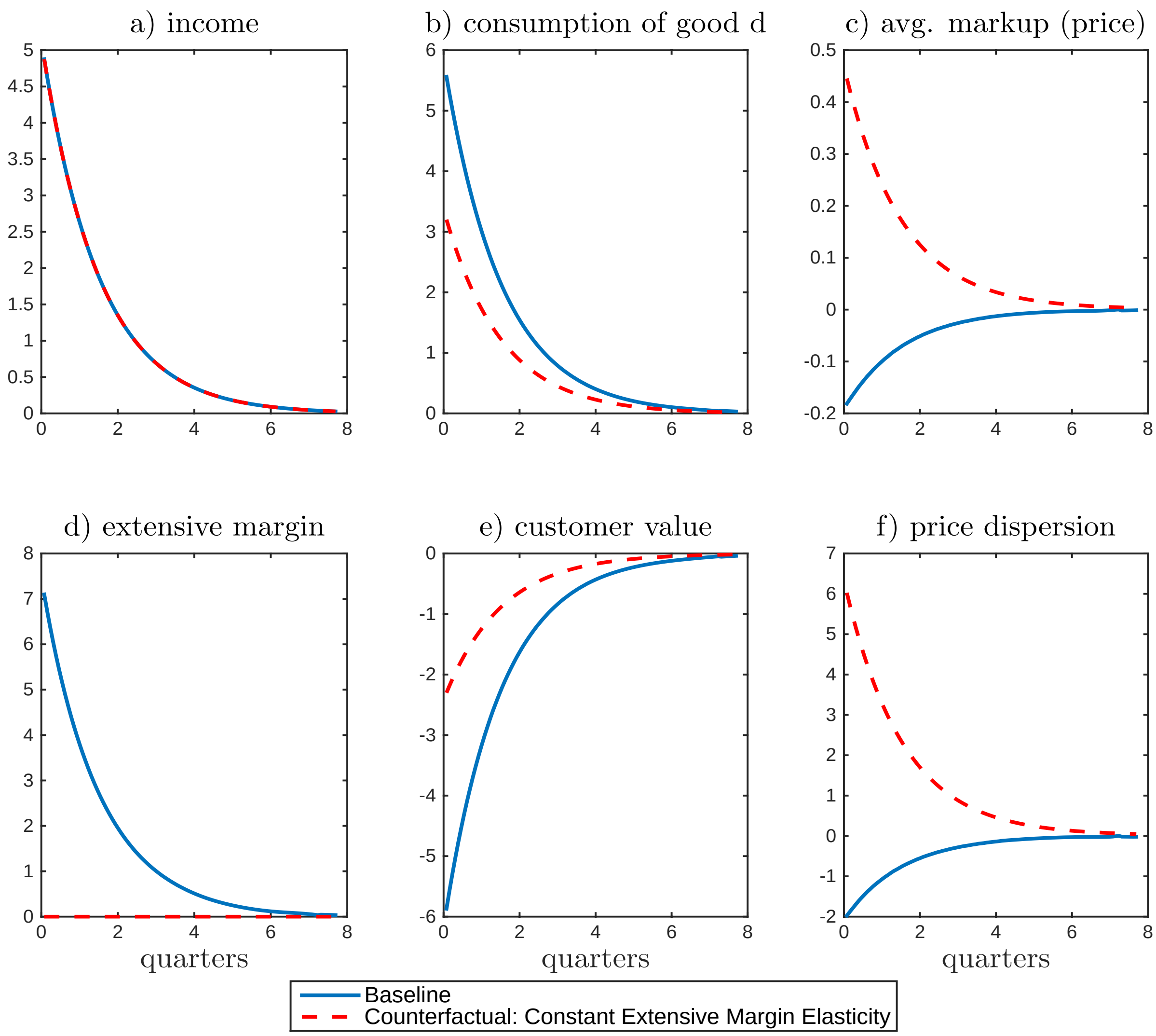
Figure 6 plots the impulse responses of several variables of interest to a 5% preference shock setting $\xi_{t_0} = 1.05$ with autocorrelation $\rho_\xi = 0.95$, corresponding to a half-life of approximately one quarter. We contrast results from our model with those implied by an economy where the extensive margin elasticity is invariant to the aggregate shock. This is obtained by taking our baseline economy and fixing the extensive margin elasticity faced by each firm to the value observed in the original steady state.[34] This is akin to the typical model of customer market (Rotemberg and Woodford (1991, 1999)) where the extensive margin is introduced with an ad-hoc function invariant to the aggregate state. We refer to this benchmark as the "counterfactual economy".

The equilibrium income, in units of the numeraire good, is given by $I_t = \xi_t$.[35] In both economies, the positive demand shock leads to a rise in consumption (panel b)). However, the price response (panel c)) goes in the opposite direction. In our baseline economy, average price (which, given that the cost process is invariant to the shock, behaves in the same way as the markup) falls; whereas it rises in the counterfactual scenario. This explains why we observe a bigger effect on consumption in our model. These differences highlight the key role of an endogenous extensive margin of demand, which is present in the baseline setting but shut down in the counterfactual. In the counterfactual economy, a positive, mean reverting, demand shock implies that customers are more keen to buy in the current period relative to future periods, reducing the value of retaining a customer to the firm ($\bar{\pi}$, panel e)) and leading to a pro-cyclical response of markups. The decline in customer value also occurs in our baseline setting. However, in our baseline model, the fact that customers value consumption more not only increases their demand but it also motivates them to search harder for firms with lower prices. The spike in the mass of customers searching increases the extensive margin elasticity (ε_m, panel d)) and toughens competition for customers. In

[34]In practice, optimal firm pricing is obtained from equation (10), where $\hat{\varepsilon}_m(\hat{p}(z), z)$ is kept constant at its steady state value for each z, whereas the other objects are recomputed in the equilibrium of the counterfactual economy after the shock.

[35]This is due to the assumptions of log-utility of consumption and linear disutility of labor. Notice also that in steady state $I_t = 1$, as in the parametrization of the model studied in Section 2.4, because ξ_t converges to 1.

Figure 6: Impulse responses to an aggregate demand shock, in % deviations from s.s.



Notes: All plots report the impulse response to the same 5% preference shock for different outcomes of the model; the figures are averages where firms are weighted by the size their customer base. The blue solid line documents the response in our baseline model; the dashed red line shows the results in the counterfactual economy, where the extensive margin elasticity is invariant to the aggregate shock.

other words, the fact that more customers may be looking for a new supplier provides an opportunity for firms to snatch customers away from competitors, offering favorable prices by lowering their markups. Finally, the outcomes of the two models also diverge when it comes to dispersion in prices (panel f)): price dispersion declines in our setting and sharply rises in the counterfactual economy. This divergence can be easily understood by recalling the results of Corollary 1. In the limiting case where competition for customers is perfect, the economy converges to an equilibrium with a degenerate price distribution where all firms charge one price (no price dispersion). As illustrated in the comparative statics analysis in Figure 4, an increase in competition for customers (obtained through a variation in the scale of search cost λ) is associated with both lower average prices and lower price dispersion. Thus, as the positive demand shock increases the benefit of search in our baseline model, it triggers to more competition and lower price dispersion. In contrast, in the counterfactual model customers do not adjust their search intensity and a demand shock implies less competition. Thus, we observe more price dispersion.

5.2 Search cost as shopping time

In our model search costs take the form of utility costs and their distribution is invariant to the aggregate state. This assumption is not shared by a recent literature that focuses on the cyclicality of shopping time (Kaplan and Menzio (forthcoming)), and in particular on the idea that the opportunity cost of shopping time decreases in a recession.[36] In this section we assess whether the conclusions drawn based upon our model are robust to defining search costs in units of (shopping) time and allowing search to be more costly in expansions.

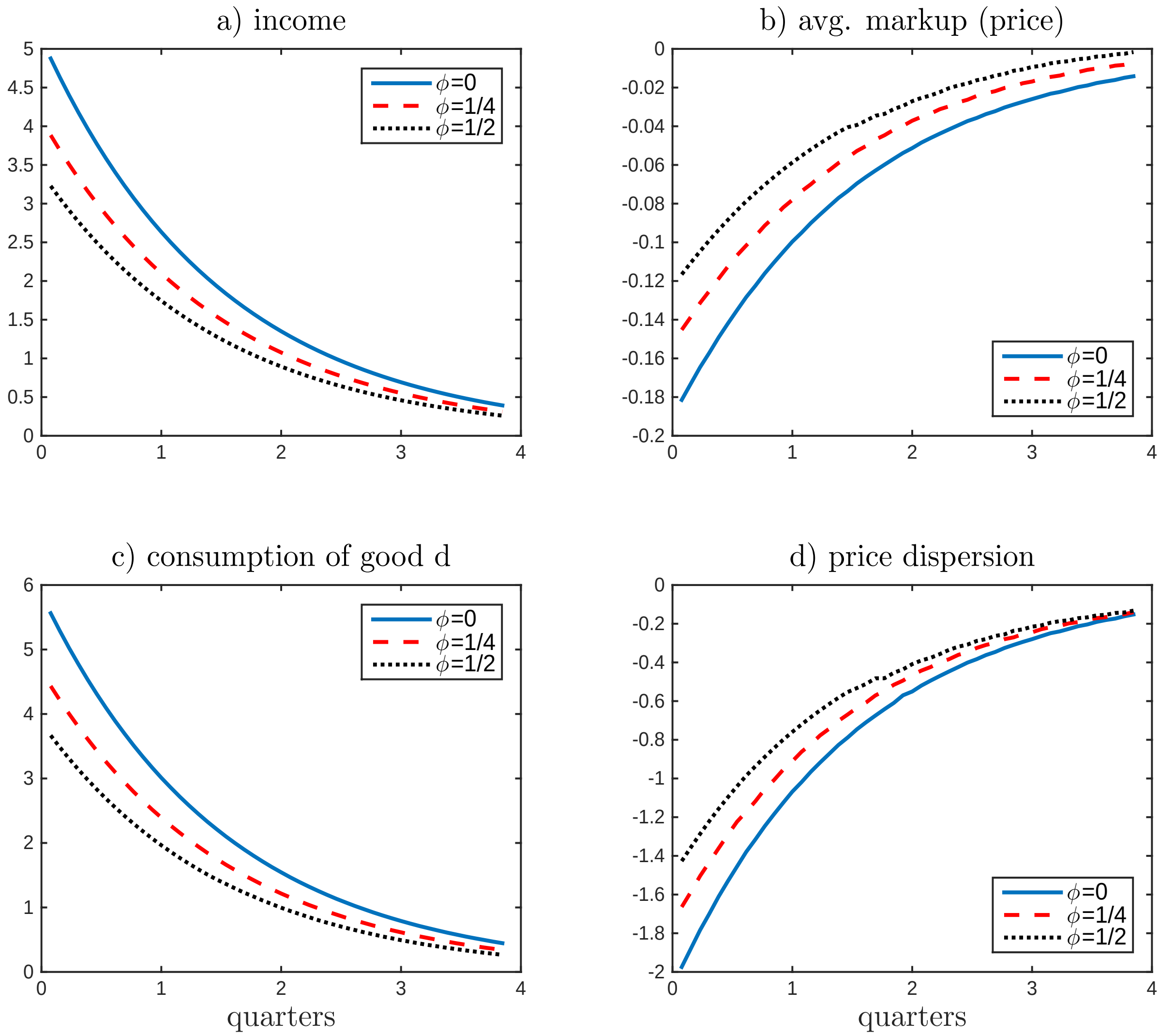
We consider the same framework and aggregate shock described in Section 5, but with two differences: i) the household is composed of a mass one of shoppers/workers, so that each shopper/worker internalizes the disutility from supplying working/shopping time; ii) when searching the shopper has to spend ψ units of time, where ψ is distributed i.i.d. according to G as in the baseline model. As in the previous case, we assume for simplicity that all workers supply the same amount of working time ℓ_t, which they take as given and is set at the household level. They also all receive the same amount of labor proceedings.[37]

As a result of our assumptions, a shopper/customer who draws a search cost ψ has a period flow utility $v_t(p)$ - $\frac{\ell_t^{1+\phi}}{1+\phi}$ in period t if she does not search, and $v_t(p)$ - $\frac{(\ell_t+\psi)^{1+\phi}}{1+\phi}$ if she does. If $\phi > 0$, an increase in aggregate labor ℓ_t is associated with a higher marginal disutility of searching, thus introducing an element of cyclicality in the opportunity cost of shopping

[36]For a summary on the empirical evidence on the cyclicality of shopping time, see Petrosky-Nadeau et al. (2014).

[37]Further details on this extension are provided in Appendix G.

Figure 7: Impulse responses to a demand shock as a function of ϕ, in % deviations from s.s.



Notes: All plots report the impulse response to the same 5% preference shock for different outcomes of the model; the figures are averages where firms are weighted by the size their customer base. The blue solid line documents the response in our baseline model, corresponding to $\psi = 0$; the dashed red line and dotted black lines shows the results when $\psi = 1/4$ and 1/2 respectively.

time. Our baseline model is nested in this setting as for $\phi = 0$ search costs are invariant to the aggregate state. Hence, we calibrate the extended model so that for $\phi = 0$ it is isomorphic to the model considered in Section 5, and we consider how such a model responds to an aggregate demand shock as we vary ϕ. The parameter ϕ is equal to the inverse of the Frisch labor supply elasticity. Values of the Frisch elasticity used by macroeconomists to calibrate general equilibrium models typically range between 2 and 4, corresponding to a range of ϕ between 1/4 and 1/2 (see Peterman (2016) for a discussion).

Figure 7 plots impulse responses of income, markup, demand and price dispersion in our economy for three values of $\phi \in \{0, 1/4, 1/2\}$ to the same aggregate shock to the consumption/leisure margin ξ_t considered in Section 5. As discussed above, the case $\phi = 0$ is identical to the model we considered in the previous section and is reported as benchmark. For $\phi > 0$, the increase in the benefit of searching originating directly from the demand shock is partially mitigated by the increase in the marginal disutility of searching associated with the increase in employment. However, at our parameter estimates, a positive shock to preferences still results in an increase in the mass of searching customers. Therefore, our model maintains its distinctive prediction even in a framework where we allow the cost of search to scale in response to aggregate shocks.

5.3 The cyclicality of price dispersion in the data

In our model a toughening of the competition among firms is signalled not only by a reduction in markups but also by a reduction in price dispersion. This result is noteworthy because, unlike markups which are traditionally difficult to measure reliably, there is abundance of data on prices which can be collected as easily as scrapping shopping sites on the Internet. In this section, we take advantage of the convenient availability of price data by exploiting the IRI database to construct market level measures of price dispersion and correlate them with a proxy for the state of local demand. By doing so, we are providing evidence that relates to a novel prediction of our model we illustrated in Section 5.1: competition and demand should be positively related.[38]

Since in our model prices are flexible, variation in price dispersion only arises from variation in optimal markups. Theories of nominal price rigidities suggest that time varying price dispersion can also arise from time varying inflation due to staggered and non coordinated price adjustments. Although quantitative analyses of this relationship find that staggered adjustment has a small impact on price dispersion for moderate levels of inflation (Alvarez et al. (2015)), we need to purge the impact of nominal rigidities if we want to look at price dispersion through the lens of our model. We do so by calculating dispersion for a given UPC at time t only across stores that have adjusted the price of that good at that time. This procedure is not necessary for the store basket, whose price changes every period.

We take a Metropolitan Statistical Area (MSA) as our market definition and use household average yearly dollar expenditure in food consumption reported in the Consumer Expenditure

[38] See Appendix I for details.

Table 4: Price dispersion and demand

	UPCs	*Basket*
Expenditure in food	-0.02*** (0.007)	-0.06* (0.033)
Market f.e.	No	Yes
Market×category f.e.	Yes	No
Obs.	264,517	141

Notes: The table reports the result of regressions of market-level price dispersion on expenditure in food in the market (MSA). Price dispersion for each UPC is computed using data from IRI as the yearly standard deviation of the price of the UPC normalized by the cross-stores average price for the UPC in the week in a market. The construction of price dispersion for the store basket is analogous. Expenditure in food is the yearly average for the market as reported in the CEX.

Survey (CEX) for eighteen MSAs as a proxy of the state of demand. Given that the frequency of the CEX data is yearly, we compute price dispersion for each MSA at the same frequency. In Table 4 we regress price dispersion on the logarithm of food expenditure. The results for price dispersion at the UPC level rely on a cross section of products from each of the 31 categories in the IRI data in each of the 18 markets. We include market-category fixed effect so that the relationship between UPC price dispersion and market demand is identified exploiting time variation within a market for products in the same category. When we analyze the price dispersion at the basket level, we have a single good for every market and therefore we only have to include market fixed effects.

The outcome of this exercise is hard to predict ex-ante because we cannot tell whether the demand variation registered in the CEX data arises from variation in marginal utility of consumption (relative to leisure) or from supply shocks affecting income. In the former case, the qualitative result would be the one of the experiment described in Figure 6, where higher consumption is associated to increased search activity. If instead the variation is due to labor supply shocks affecting income, then the cost of search might comove positively with consumption (Section 5.2), potentially leading to the opposite conclusion in terms of the correlation between price dispersion and aggregate demand.

While our regression cannot discard the presence of either channel, the fact that we find that price dispersion both measured at the basket and UPC level is negatively correlated with consumption expenditure signals, through the lens of our model, that the mechanism explored in Section 5.1 is at least as important as the income channel to explain fluctuations in price dispersion and search intensity.

6 Concluding remarks

This paper provides an assessment of the importance of customer dynamics in shaping firm pricing strategies. We develop a rich yet tractable model to assess the role that customer markets play in determining price setting. Since the model predictions crucially depend on the price elasticity of the customer base, we exploit novel data documenting prices and customer base evolution of a large US retail firm to inform the parametrization of the model. We find that incorporating customer dynamics in a model of pricing helps explaining salient features of prices emerging both from our data and independently documented relying on other sources. This new evidence on the forces involved in price formation should be of great interest for the increasing number of both micro and macroeconomists analyzing price data for an assortment of purposes. Finally, we use our framework to study the effect of accounting for customer dynamics on the propagation of aggregate shocks. Our microfoundation linking customer dynamics to search and exit decisions by individual customers results in a countercyclical response of markup to demand shocks because consumers search more intensively in periods when they value consumption more. As competition for customers intensifies in periods of high demand, our model also predicts a negative relationship between aggregate demand and price dispersion of which we find evidence in the data.

Our study relies on a number of simplifying assumptions, whose relaxation seems of interest for future research. First, for tractability we refrain from explicitly modeling persistent heterogeneity in customers search/opportunity costs (although we control for these factors in the empirical analysis) and we do not allow for price discrimination. The presence of customers heterogeneity in shopping behavior is well documented (Aguiar and Hurst (2007)), which makes studying its implications for optimal pricing and customer dynamics an important topic. Due to lack of data, we do not consider the role of advertising in generating demand dynamics (Hall (2014)). While our conjecture is that the analysis of the pricing incentives presented in this paper would still apply, we think that extending the analysis to advertising, as well as to other strategies to attract and retain customers, and confront the results with direct firm level evidence, could provide new insights about firms' behavior.

References

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

Alessandria, G. (2004). International Deviations From The Law Of One Price: The Role Of Search Frictions And Market Share, *International Economic Review* **45**(4): 1263–1291.

Alvarez, F., Beraja, M., Gonzalez-Rozada, M. and Neumeyer, A. (2015). From hyperinflation to stable prices: Argentina's evidence on menu cost models, *working paper*.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices, *American Economic Review* **98**(5): 1998–2031.

Bai, Y., Rull, J.-V. R. and Storesletten, K. (2012). Demand Shocks that Look Like Productivity Shocks, *working paper*.

Blinder, A. S., Canetti, E. R. D., Lebow, D. E. and Rudd, J. B. (1998). *Asking About Prices: A new Approach to Understanding Price Stickiness*, New York: Russell Sage Foundation.

Bronnenberg, B. J., Kruger, M. W. and Mela, C. F. (2008). The IRI Marketing data set, *Marketing Science* **27**(4): 745–748.

Burdett, K. and Coles, M. G. (1997). Steady state price distributions in a noisy search equilibrium, *Journal of Economic Theory* **72**(1): 1 – 32.

Burstein, A. and Hellwig, C. (2007). Prices and market shares in a menu cost model, *NBER Working Papers 13455*, National Bureau of Economic Research, Inc.

Cabral, L. (2014). Dynamic pricing in customer markets with switching costs, *Working paper*.

Chevalier, J. A., Kashyap, A. K. and Rossi, P. E. (2003). Why don't prices rise during periods of peak demand? Evidence from scanner data, *American Economic Review* **93**(1): 15–37.

Coibion, O., Gorodnichenko, Y. and Hong, G. H. (2015). The cyclicality of sales, regular and effective prices: Business cycle and policy implications, *American Economic Review* **105**(3): 993–1029.

Diamond, P. A. (1971). A model of price adjustment, *Journal of Economic Theory* **3**(2): 156–168.

Dinlersoz, E. M. and Yorukoglu, M. (2012). Information and Industry Dynamics, *American Economic Review* **102**(2): 884–913.

Drozd, L. A. and Nosal, J. B. (2012). Understanding international prices: Customers as capital, *American Economic Review* **102**(1): 364–95.

Eichenbaum, M., Jaimovich, N. and Rebelo, S. (2011). Reference prices and nominal rigidities, *American Economic Review* **101**(1): 234–262.

Einav, L., Leibtag, E. and Nevo, A. (2010). Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?, *Quantitative Marketing and Economics* **8**(2): 207–239.

Einav, L. and Somaini, P. (2013). A model of market power in customer markets, *Journal of Industrial Economics* **61**(4): 938–986.

Evans, D. S. (1987). Tests of Alternative Theories of Firm Growth, *Journal of Political Economy* **95**(4): 657–74.

Fabiani, S., Loupias, C., Martins, F. and Sabbatini, R. (2007). *Pricing decisions in the euro area: how firms set prices and why*, Oxford University Press, USA.

Fishman, A. and Rob, R. (2003). Consumer inertia, firm growth and industry dynamics, *Journal of Economic Theory* **109**(1): 24 – 38.

Fishman, A. and Rob, R. (2005). Is bigger better? Customer base expansion through word-of-mouth reputation, *Journal of Political Economy* **113**(5): 1146–1175.

Foster, L., Haltiwanger, J. and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?, *American Economic Review* **98**(1): 394–425.

Foster, L., Haltiwanger, J. and Syverson, C. (2016). The slow growth of new plants: Learning about demand?, *Economica* **83**(329): 91–129.

Galenianos, M. and Gavazza, A. (2015). A quantitative analysis of the retail market for illicit drugs, *mimeo*, LSE.

Gourio, F. and Rudanko, L. (2014). Customer capital, *The Review of Economic Studies* **81**: 1102–1136.

Hall, R. E. (2008). General equilibrium with customer relationships: A dynamic analysis of rent-seeking, *working paper*.

Hall, R. E. (2014). What the cyclical response of advertising reveals about markups and other macroeconomic wedges, *working paper*.

Holmes, T. (2011). The diffusion of Wal-Mart and economies of density, *Econometrica* **79**: 253–302.

Kaplan, G. and Menzio, G. (2015). The morphology of price dispersion, *International Economic Review* **56**(4): 1165–1206.

Kaplan, G. and Menzio, G. (forthcoming). Shopping externalities and self-fulfilling unemployment fluctuations, *Journal of Political Economy* .

Kleshchelski, I. and Vincent, N. (2009). Market share and price rigidity, *Journal of Monetary Economics* **56**(3): 344–352.

Menzio, G. (2007). A search theory of rigid prices, *working paper*.

Nakamura, E. and Steinsson, J. (2011). Price setting in forward-looking customer markets, *Journal of Monetary Economics* **58**(3): 220–233.

Peterman, W. (2016). Reconciling micro and macro estimates of the frisch labor supply elasticity, *Economic Inquiry* **54**(1): 100–120.

Petrosky-Nadeau, N. and Wasmer, E. (2015). Macroeconomic dynamics in a model of goods, labor, and credit market frictions, *Journal of Monetary Economics* **72**: 97–113.

Petrosky-Nadeau, N., Wasmer, E. and Zeng, S. (2014). Shopping time, *Available at SSRN 2499871* .

Phelps, E. S. and Winter, S. G. (1970). Optimal price policy under atomistic competition, *in* G. C. Archibald, A. A. Alchian and E. S. Phelps (eds), *Microeconomic Foundations of Employment and Inflation Theory*, New York: Norton.

Ravn, M., Schmitt-Grohe, S. and Uribe, M. (2006). Deep habits, *Review of Economic Studies* **73**(1): 195–218.

Reinganum, J. F. (1979). A simple model of equilibrium price dispersion, *Journal of Political Economy* **87**(4): 851–58.

Rotemberg, J. J. and Woodford, M. (1991). Markups and the business cycle, *NBER Macroeconomics Annual 1991, Volume 6*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 63–140.

Rotemberg, J. J. and Woodford, M. (1999). The cyclical behavior of prices and costs, *working paper*.

Shi, S. (1997). A divisible search model of fiat money, *Econometrica* **65**(1): 75–102.

Shi, S. (2011). Customer relationship and sales, *working paper*.

Smith, H. T. (2004). Supermarket choice and supermarket competition, *Review of Economic Studies* **71**: 235–263.

Sorensen, A. T. (2000). Equilibrium price dispersion in retail markets for prescription drugs, *Journal of Political Economy* **108**(4): 833–850.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions, *Economics Letters* **20**(2): 177–181.

Vavra, J. (2014). Inflation dynamics and time-varying volatility: New evidence and an ss interpretation, *Quarterly Journal of Economics* **129**(1).

Appendix - Not for publication

A Proof of Proposition 1

The following lemma discusses some key properties of the optimal price useful to prove Proposition 1.

Lemma 1 *Let $\Delta(p,z)$ be continuous in p, and let $\varepsilon_m(p,z) \equiv \partial \log(\Delta(p,z))/\partial \log(p)$. If a price $\bar{p}(z)$ exists such that $\varepsilon_m(p,z) > 0$ for all $p > \bar{p}(z)$, and $\varepsilon_m(p,z) = 0$ for all $p \leq \bar{p}(z)$, then we have $\hat{p}(z) \in [\bar{p}(z), p^*(z))$ if $\bar{p}(z) < p^*(z)$, and $\hat{p}(z) = p^*(z)$ otherwise.*

The proof of the lemma is an immediate implication of equation (4). We next prove the results of Proposition 1.

Monotonicity of prices. Monotonicity of optimal prices follows from an application of Topkis' theorem. In order to apply the theorem to the firm problem in equation (3) we need to establish increasing differences of the firm objective $\Delta(p,z)\,\Pi(p,z)$ in $(p,-z)$. Under the standard assumptions we stated on $\pi(p,z)$, it is easy to show that $\Pi(p,z)$ satisfies this property. The customer base growth function does not in general verify the increasing difference property. However, let $\bar{p}(z)$ denote the price p that solves $\bar{V}(p,z) = \mathcal{V}(\bar{z})$. We have that $\Delta(p,z)$ is continuous, strictly decreasing in p for all $p > \bar{p}(z)$, and constant for all $p \leq \bar{p}(z)$. Under the assumption of i.i.d. productivity, $\Delta(p,z)$ is independent of z, which is sufficient to obtain the result. We first show that optimal prices $\hat{p}(z)$ are non-increasing in z. Given, that productivity is i.i.d. and that we look for equilibria where $\hat{p}(z) \geq p^*(\bar{z})$, we have that $\bar{p}(z) = p^*(\bar{z})$ for each z. From Lemma 1 we know that, for a given z, the optimal price $\hat{p}(z)$ belongs to the set $[p^*(\bar{z}),\ p^*(z)]$. Over this set, the objective function of the firm,

$$W(p,z) = \Delta(p,z)\left(\pi(p,z) + \beta\,\text{constant}\right)\,, \tag{15}$$

is supermodular in $(p,-z)$. Notice the i.i.d. assumption implies that future profits of the firm do not depend on current productivity as future productivity, and therefore profits, are independent from it. Similarly, $\Delta(p,z)$ does not depend on z, as the expected future value to the customer does not depend on the productivity of the current match as future productivity is independent from it. Abusing notation, we replace $\Delta(p,z)$ by $\Delta(p)$. To show that $W(p,z)$ is supermodular in $(p,-z)$ consider two generic prices p_1, p_2 with $p_2 > p_1 > 0$ and productivities $z_1, z_2 \in [\underline{z}, \bar{z}]$ with $-z_2 > -z_1$. We have that $W(p_2,z_2) - W(p_1,z_2) \leq W(p_2,z_1) - W(p_1,z_1)$

if and only if

$$\Delta(p_2)d(p_2)(p_2-w/z_2)-\Delta(p_1)d(p_1)(p_1-w/z_2) \leq \Delta(p_2)d(p_2)(p_2-w/z_1)-\Delta(p_1)d(p_1)(p_1-w/z_1),$$

which, since $\Delta(p_2)d(p_2) < \Delta(p_1)d(p_1)$ as $d(p)$ is strictly decreasing and $\Delta(p)$ is non-increasing, is indeed satisfied if and only if $z_2 < z_1$. Thus, $W(p, z)$ is supermodular in $(p, -z)$. By application of the Topkis Theorem we readily obtain that $\hat{p}(z)$ is non-increasing in z.

Existence of equilibrium. Next we prove existence of an equilibrium. The fixed point problem is a mapping from candidate functions of equilibrium prices, $\hat{p}(z)$, to the firm's optimal pricing strategy, $\hat{p}(z)$. Notice that $W(p, z)$ in equation (15) is continuous in (p, z). By the theorem of maximum, $\hat{p}(z)$ is upper hemi-continuous and $W(\hat{p}(z), z)$ is continuous in z. Given that $\hat{p}(z)$ is non-increasing in z it follows that $\hat{p}(z)$ has a countably many discontinuity points. We thus proceed as follows. Let $\hat{\mathcal{P}}(z)$ be the set of prices that maximize the firm problem. Whenever a discontinuity arises at some $\tilde{z}$ (so that $\hat{\mathcal{P}}(\tilde{z})$ is not a singleton), we modify the optimal pricing rule of the firm and consider the convex hull of the $\hat{\mathcal{P}}(\tilde{z})$ as the set of possible prices chosen by the firm with productivity $\tilde{z}$. The constructed mapping from $\mathcal{P}(z)$ to $\hat{\mathcal{P}}(z)$ is then upper-hemicontinous, compact and convex valued. We then apply Kakutani's fixed point theorem to this operator and obtain a fixed point. Finally, notice that since the convexification procedure described above has to be applied only a countable number of times, the set of convexified prices has measure zero with respect to the density of z. Hence, they do not affect the fixed point.

It is important to point out that differentiability of the distribution of productivity F is not needed for the existence of an equilibrium. We assume it to ensure that $H(\cdot)$ and $Q(\cdot)$ are almost everywhere differentiable so that equation (4) is a necessary condition for optimal prices (see below). However, even when F is not differentiable and the first order condition cannot be used to characterize the equilibrium, an equilibrium with the properties of Proposition 1 exists where $\hat{p}(z)$ and $\hat{\psi}(\hat{p}(z), z)$ are monotonic in z but not necessarily strictly monotonic for all z.

Necessity of the first order condition. We show that Q and H are almost everywhere differentiable, so that Lemma 1 implies that equation (4) is necessary for an optimum. We guess that $\hat{p}(z)$ is strictly decreasing and almost everywhere differentiable. It immediately follows that $\mathcal{V}(z)$ is strictly increasing in z and almost everywhere differentiable. Then, given the assumption that F is differentiable, we have that K is differentiable. From $H(x) = K(\mathcal{V}^{-1}(x))$ it follows that H is also almost everywhere differentiable. Given that G and H are differentiable, so is Q. Then the first order condition in equation (4) is necessary for an optimum, which indeed implies that $\hat{p}(z)$ is strictly decreasing and differentiable in z in any neighborhood of the first order condition. Finally, given that $\hat{p}(z)$ has a countably many discontinuity points, it has countably many points where it is not differentiable, and the first order condition does not apply at those points, but applies everywhere else. These points have measure zero with respect to the density of z and therefore $\hat{p}(z)$ is almost everywhere differentiable.

Proof of Point (i). We already proved that $\hat{p}(z)$ is non-increasing in z. The proof that $\hat{p}(z)$ is strictly decreasing follows by contradiction. Consider that $\hat{p}(z_1) = \hat{p}(z_2) = \tilde{p}$ for some $z_1, z_2 \in [\underline{z}, \bar{z}]$. Also, without loss of generality, assume that $z_1 < z_2$. Given that we already established the necessity of the first order condition presented in equation (4) when prices are monotonic, suppose that it is satisfied at the pair $\{z_2, \tilde{p}\}$. Notice that, because of the assumed i.i.d. structure of productivity shocks together with $\pi_z(p, z) < 0$, it is not possible that the first order condition is also satisfied at the pair $\{z_1, \tilde{p}\}$. Moreover, because the first order condition is necessary and we already established that $\hat{p}(z)$ cannot be increasing at any z, we conclude that the optimal price at z_1 is strictly larger than at z_2. That is, $\hat{p}(z_1) > \hat{p}(z_2)$. Notice that this verifies the conjecture used to prove the necessity of the first order condition, which in turn validates the use of equation (4) here.[39]

Notice that, because $\hat{p}(z)$ is strictly decreasing in z, the fact that $v'(p) < 0$ together with i.i.d. productivity, implies, through an application of the contraction mapping theorem, that $\mathcal{V}(z) = \bar{V}(\hat{p}(z), z)$ is increasing in z.

Proof of Point (ii). $\hat{\psi}(p, z) \geq 0$ immediately follows its definition. The fact that $\mathcal{V}(z)$ is strictly increasing in z implies that $\hat{\psi}(\hat{p}(\bar{z}), \bar{z}) = 0$ and that $\hat{\psi}(\hat{p}(z), z)$ and $\Delta(\hat{p}(z), z)$ are strictly increasing in z. Because of price dispersion, some customers are searching, which guarantees that $\Delta(\hat{p}(\bar{z}), \bar{z}) > 1$. Likewise, $\Delta(\hat{p}(\underline{z}), \underline{z}) < 1$.

[39] If prices are not strictly decreasing, this argument cannot be used as the first order condition is not necessary. However, it is possible to prove that $\hat{p}(z)$ is strictly decreasing in z for some region of z. The argument follows by contradiction. Suppose that $\hat{p}(z)$ is everywhere constant in z at some $\tilde{p}$. Then $\bar{p}(z) = \tilde{p}$ for all z. If $\tilde{p} > p^*(\bar{z})$, then $\tilde{p}$ would not be optimal for firm with productivity $\bar{z}$, which would choose a lower price. If $\tilde{p} = p^*(\bar{z})$, then continuous differentiability of G together with $H = G = Q = 0$ at the conjectured constant equilibrium price imply that the first order condition is locally necessary for an optimum, and a firm with productivity $z < \bar{z}$ would have an incentive to deviate according to equation (4), and set a strictly higher price than $\tilde{p}$. Finally, the result that $\hat{p}(z) < p^*(z)$ for all $z < \bar{z}$ and that $\hat{p}(\bar{z}) = p^*(\bar{z})$ follows from applying Lemma 1, and using that $\hat{p}(z) \geq \hat{p}(\bar{z})$ and $\bar{p}(z) = \hat{p}(\bar{z})$ for all z.

B Proof of Corollary 1

Part (1): Start by noticing that, because the mean of $G(\psi)$ is positive, the expected value of searching diverges to $-\infty$ as n diverges to infinity. Because prices are finite for all $z \in [\underline{z}, \bar{z}]$, the value of not searching is bounded. As a result, customers do not search so that firms do not face customer base concerns. Formally, $\bar{p}(z) \to \infty$ for all $z \in [\underline{z}, \bar{z}]$. Because $p^*(z)$ is finite for all $z \in [\underline{z}, \bar{z}]$, it follows immediately that $p^*(z) < \bar{p}(z)$ for all $z \in [\underline{z}, \bar{z}]$. Then, using Lemma 1 we obtain that $\hat{p}(z) = p^*(z)$ for all $z \in [\underline{z}, \bar{z}]$.

Part (2): From Proposition 1 we have that, in equilibrium, the highest price is $\hat{p}(\underline{z})$. Moreover, under the assumptions of Proposition 1, the first order condition is a necessary condition for optimality of prices. We use this to show that, as n approaches zero, $\hat{p}(\underline{z})$ has to approach $\hat{p}(\bar{z}) = p^*(\bar{z})$. In equilibrium, it is possible to rewrite equation (4), evaluated at $\{\hat{p}(\underline{z}), \underline{z}\}$, as $LHS(\hat{p}(\underline{z}), n) = RHS(\hat{p}(\underline{z}), n)$, where

$$\begin{aligned} LHS(\hat{p}(\underline{z}), n) \quad &\equiv \quad G'(\hat{\psi}(\hat{p}(\underline{z}), \underline{z})/n)\hat{\psi}_p(\hat{p}(\underline{z}), \underline{z})/n + \\ &\quad + \left(G(\hat{\psi}(\hat{p}(\underline{z}), \underline{z})/n)H'(\bar{V}(\hat{p}(\underline{z}), \underline{z})) + Q'(\bar{V}(\hat{p}(\underline{z}), \underline{z}))\right)\bar{V}_p(\hat{p}(\underline{z}), \underline{z}) \ , \\ RHS(\hat{p}(\underline{z}), n) \quad &\equiv \quad -\frac{\pi_p(\hat{p}(\underline{z}), \underline{z})}{\Pi(\hat{p}(\underline{z}), \underline{z})}\left(1 - G(\hat{\psi}(\hat{p}(\underline{z}), \underline{z})/n)\right) \ , \end{aligned}$$

given that $H(\bar{V}(\hat{p}(\underline{z}), \underline{z})) = Q(\bar{V}(\hat{p}(\underline{z}), \underline{z})) = 0$. Suppose that as $n \downarrow 0$, $\hat{\psi}(\hat{p}(\underline{z}), \underline{z})$ does not converge to zero. Then, $G\left(\frac{\hat{\psi}(\hat{p}(\underline{z}), \underline{z})}{n}\right) \uparrow 1$ as $n \downarrow 0$. This implies that $\lim_{n \downarrow 0} RHS(\hat{p}(\underline{z}), n) > 0$. Consider now the function $LHS(\hat{p}(\underline{z}), n)$. Again, suppose that as $n \downarrow 0$, $\hat{\psi}(\hat{p}(\underline{z}), \underline{z})$ does not converge to zero. Notice that the second term of the function approaches a finite number as $\bar{V}_p(\hat{p}(\underline{z}), \underline{z})$ is bounded by assumptions on $v(p)$ and $H'(\bar{V}(\hat{p}(\underline{z}), \underline{z}))$ and $Q'(\bar{V}(\hat{p}(\underline{z}), \underline{z}))$ being bounded as a result of Proposition 1. Moreover, as long as $\hat{p}(\underline{z}) > \bar{p}(z) = p^*(\bar{z})$, we have that $\hat{\psi}_p(\hat{p}(\underline{z}), \underline{z}) > 0$ so that $\hat{\psi}_p(\hat{p}(\underline{z}), \underline{z})/n$ diverges as n approaches zero. This means that $G'\left(\frac{\hat{\psi}(\hat{p}(\underline{z}), \underline{z})}{n}\right)\frac{\hat{\psi}_p(\hat{p}(\underline{z}), \underline{z})}{n}$ is divergent, and therefore the first order condition cannot be satisfied.

This analysis concluded that, if $\hat{\psi}(\hat{p}(\underline{z}), \underline{z})$ does not converge to zero as n becomes arbitrarily small, the first order condition, i.e. equation (4), cannot be satisfied. This occurs because $LHS(\hat{p}(\underline{z}), n)$ would diverge to infinity, while $RHS(\hat{p}(\underline{z}), n)$ would remain finite. It then follows that, as n approaches zero, a necessary condition is that $\hat{\psi}(\hat{p}(\underline{z}), \underline{z})$ also approaches zero. This condition can be restated as requiring that $\hat{p}(\underline{z})$ approaches $\bar{p}(z)$ as n approaches zero. Moreover, given the assumptions of Proposition 1, $\bar{p}(z) = \hat{p}(\bar{z}) = p^*(\bar{z})$.

In the end, if $\hat{p}(\underline{z})$ approaches $p^*(\bar{z})$ as n becomes arbitrarily small (so that $\hat{\psi}(\hat{p}(\underline{z}), \underline{z}) \to 0$ and $\hat{\psi}_p(\hat{p}(\underline{z}), \underline{z}) \to 0$), we have that $\lim_{n \downarrow 0} LHS(\hat{p}(\underline{z}), n) < \infty$ and $\lim_{n \downarrow 0} RHS(\hat{p}(\underline{z}), n) < \infty$ as $\pi_p(p^*(\bar{z}), \underline{z})$ is bounded as $\pi(p^*(\bar{z}), \underline{z}) > 0$. However, if $\hat{p}(\underline{z})$ does not approach $p^*(\bar{z})$ as n

becomes arbitrarily small, we have that $LHS(\hat{p}(\underline{z}), n)$ diverges as n approaches zero, while $LHS(\hat{p}(\underline{z}), n)$ remains finite. As the first order condition has to be satisfied in equilibrium, a necessary condition is that, as n approaches zero, the highest price in the economy, i.e. $\hat{p}(\underline{z})$, has to approach the lowest price in the economy, i.e. $p^*(\bar{z})$.

C Data sources and variables construction

This appendix provides additional information on the data sources presented in Section 3. We also document more in depth the procedure used to construct the main variable used to empirically assess the relevance of the extensive margin of demand.

C.1 Data and selection of the sample

The retailer that provided both the price data and the consumer panel is a large supermarket chain that operates over 1,000 stores across the United States. It is a high/low supermarket chain selling grocery goods as well as household supplies; it could be compared to Kroeger or Tesco.

Sampling and representativeness

The Consumer Panel data include complete purchase data for over 11,000 customers of the chain sampled for the major markets for the retailer, excluding those where it operates under acquired brands. Households are tracked through usage of the supermarket loyalty card; purchases made without using the card are not recorded. However, the chain ensures that the loyalty card has a high penetration, for instance by keeping to a minimum the effort needed to register. Furthermore, nearly all promotional discounts are tied to ownership of a loyalty card, which provides a strong incentive to sign up and use it. Therefore, we can consider the customers in our sample as representative of the population of non casual shoppers at the chain.

The Price Data cover 270 stores. This is about a fifth of the stores operated by the retailer; however, the chain sets different prices for the same UPC in different geographic areas, called “price areas”. The set of stores for which the retailer provided information was designed so that at least one store for each price area would be included.

C.2 Variables construction

Exit from the customer base

The dependent variable in the regression presented in equation (12) is an indicator for whether a customer is exiting the customer base of the chain. With data on grocery purchases at a single retail chain it is hard to definitively assess whether a customer has abandoned the retailer to shop elsewhere or she is simply not purchasing groceries in a particular week, for instance because she is just consuming her inventory. In fact, we observe households when they buy groceries at the chain but do not have any information on their shopping at competing grocers. Our choice is to assume that a customer is shopping at some other store when she has not visited any supermarket store of the chain for at least eight consecutive weeks. The *Exit* dummy is then constructed so that it takes value of one in correspondence to the last visit at the chain before a spell of eight or more weeks without shopping there. Table 5 summarizes shopping behavior for households in our sample. It is immediate to notice that an eight-week spell without purchase is unusual, as customers tend to show up frequently at the stores. This strengthens our confidence that customers missing for an eight-week period have indeed switched to a different retailer.

Table 5: Descriptive statistics on customer shopping behavior

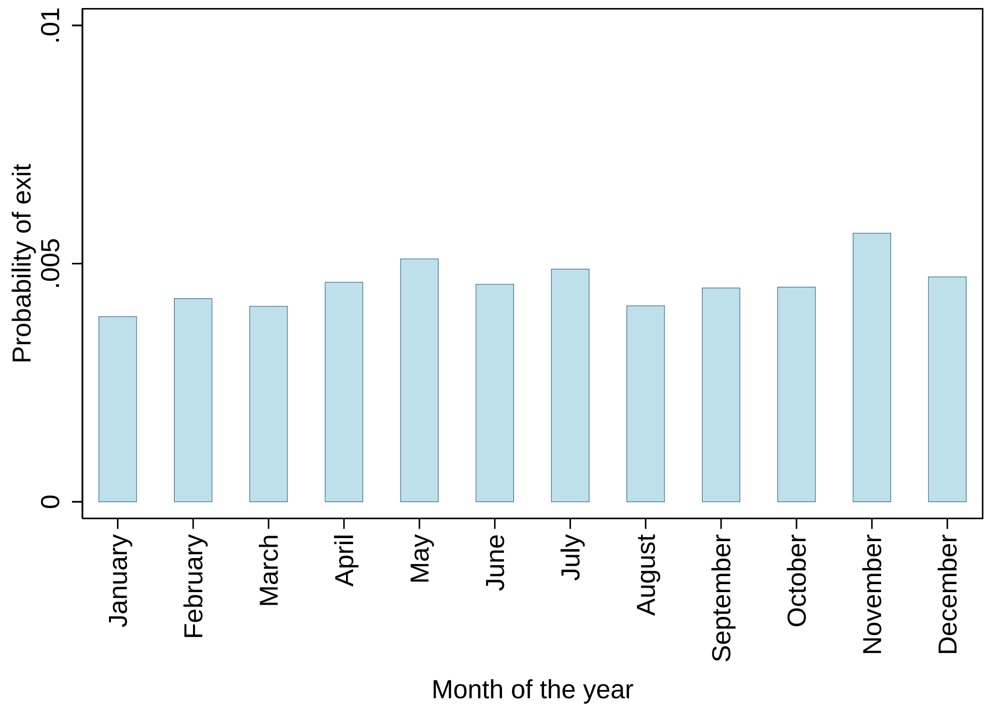
	Mean	*Std.dev.*	*25th pctile*	*75th pctile*
Number of trips	150	127	66	200
Days elapsed between consecutive trips	4.2	7.5	1	5
Expenditure per trip ($)	69	40	40	87
Frequency of exits	0.003	0.065		

In Figure 8 we document the seasonality in exit rates. We find that the probability of exit is roughly stable across months.

Weekly UPC prices

The Consumer Panel reports information on the price paid conditional on a certain item having been bought by the customer. Therefore, if we do not observe at least one household in our sample buying a given item in a store in a week, we cannot infer the price of the item in that store-week. However, our definition of basket requires us to be able to attach a price to each of the items composing it in every week, even when the customer does not shop. The issue can be solved using the Price data which report information on weekly store revenues and quantities, regardless of the shopping decisions of the households in our Consumer Panel. We use data on store level revenues and quantities sold in the Price data to compute Unit

Figure 8: Survival in the customer base



Notes: The figure plots the unconditional probability of exit, computed as the ratio of the number of exits and the number of shopping trips by customers of the chain, by month. The definition of exit adopted for the plot is our baseline one: lack of shopping trips at the chain for 8 consecutive weeks or more.

value prices as

$$UVP_{tu}^{j} = \frac{TR_{tu}^{j}}{Q_{tu}^{j}} \; ,$$

where TR represent total revenues and Q the total number of units sold of good u in week t in store j.

As explained in Eichenbaum et al. (2011), this only allows us to recover an average price for goods that were on promotion. In fact the same good will be sold to loyalty card carrying customers at the promotional price and at full price to customers who do not have or use a loyalty card. Without information on the fraction of these two types of customers it is not possible to recover the two prices separately. Furthermore, since prices are constructed based on information on revenues, missing values can originate even in this case if no unit of a specific item is sold in a given store in a week. This is, however, a rare occurrence and involves only infrequently purchased UPCs, which are unlikely to represent important shares of the basket for any of the households in the sample. For the analysis, we only retain UPCs with at most two nonconsecutive missing price observations and impute price for the missing observation interpolating the prices of the contiguous weeks.

In order to use unit value prices calculated from store-level data to compute the price of the basket of a specific household, we need to determine to which price area the store(s) at which she regularly shops belong. This information is not supplied by the retailer that kept the exact definition of the price areas confidential. A possible solution is to infer in which price areas the store(s) visited by a household are located by comparing the prices contained in the household panel with those in the store data. In principle the household data should give information on enough UPC prices in a given week to identify the price area representative store whose pricing they are matching. However, even though two stores belonging in the same price area should have the same prices, they may not have the same unit value prices if the share of shoppers using the loyalty card differs in the two stores. Therefore, we choose to restrict our analysis to the set of customers shopping predominantly (over 80% of their grocery expenditure at the chain) in one of the 270 stores for which the chain provided complete store-level data. This choice is costly in terms of sample size: only 1,336 households (or 12% of the original sample) shop at one of the 270 stores for which we have store-level price data. However, since the 270 representative stores were randomly chosen, the resulting subsample of households should not be subject to any selection bias.

Composition of the household basket and basket price

The household scanner data deliver information on all the UPCs a household has bought through the sample span. We assume that all of them are part of the household basket and, therefore, the household should care about all of those prices. Some of the items in the household's basket are bought regularly, whereas others are purchased less frequently. We take this into account when constructing the price of the basket by weighting the price of each item by its expenditure share in the household budget. The price of household i's basket purchased at store j in week t is computed as:

$$p_{it}^{j} = \sum_{u \in K_i} \omega_{iu}\, p_{ut}^{j} \ , \qquad \omega_{iu} = \frac{\sum_t E_{iut}}{\sum_u \sum_t E_{iut}} \ ,$$

where K^i is the set of all the UPCs (u) purchased by household i during the sample period, p_{ut}^{j} is the price of a given UPC u in week t at the store $costj$ where the customer shops. E_{iut} represents expenditure by customer i in UPC u in week t and the ω_{iu}'s are a set of household-UPC specific weights. There is the practical problem that the composition of the consumer basket cannot vary through time; otherwise basket prices for the same customer in different weeks would not be comparable. This requires that we drop from the basket all UPCs for which we do not have price information for every week in the sample. However, the price information is missing only in instances where the UPC registered no sales in a particular week. It follows that only low market-share UPCs will have missing values and, therefore, the UPCs entering the basket computation will represent the bulk of each customer's grocery expenditure. The construction of the cost of the basket follows the same procedure where we substitute the unit value price with the measure of replacement cost provided by the retailer.

We choose to calculate the weights using the total expenditure in the UPC by the household over the two years in the sample. This can lead to some inaccuracy in identifying the goods the customer cares for at a given point in time. For example, if a customer bought only Coke during the first year and only Pepsi during the second year of data, our procedure would have us give equal weight to the price of Coke and Pepsi throughout the sample period. If we used a shorter time interval, for example using the expenditure share in the month, we would correctly recognize that she only cares about Coke in the first twelve months and only about Pepsi in the final 12 months. However, weights computed on short time intervals are more prone to bias induced by pricing. For example, a two-weeks promotion of a particular UPC may induce the customer to buy it just because of the temporary convenience; this would give the UPC a high weight in the month. The effect of promotion is instead smoothed when we compute weights using expenditure over the entire sample period.

The construction of the price of the competitors occurs in two steps. First, we use the IRI data and the same procedure described above to obtain a price for the basket of each consumer at every store located in her same Metropolitan Statistical Area. Next, we average those prices across stores to obtain the average market price of the consumer basket. In particular, the price is computed as:

$$\bar{p}_{it} = \sum_{j \in m(i)} s^j \sum_{u \in K_i} \omega_{iu} p^j_{ut}, \quad \omega_{iu} = \frac{\sum_t E_{iut}}{\sum_u \sum_t E_{iut}}\ , \quad s^j = \frac{\sum_t R^j_t}{\sum_{j' \in M} \sum_t R^{j'}_t}$$

where $m(i)$ is the market of residence for customer i and R^j_t represents revenues of store j in week t. In other words, in the construction of the competitors' price index stores with higher (revenue-based) market shares weight more.

D Numerical solution of the model

In order to solve the model, we start by setting the parameters. The parameters β, κ and I are constant throughout the numerical exercises. For the set of estimated parameters $\Omega_n = [\lambda_n,\ \zeta_n,\ \rho_n,\ \sigma_n]'$, we set a search grid. The grid is different for each parameter, as they differ both in their levels and in the sensitivity of the statistics of interest to their variation. We consider a grid with an interval of 0.01 for σ, 0.05 for ρ, 0.5 for ζ, and 0.01 for λ. Each Ω_n corresponds to a particular combination of parameters among these grids. For each Ω_n we set θ to obtain $E[\varepsilon_d(z) + [\varepsilon_m(z)] = 4$ and the exogenous customer attrition rate δ so to match the fraction of customers that exit the customer base of the supermarket chain in a week, i.e. 0.0044.

We next describe how we solve for the equilibrium of the model for a given combination of parameters. We start by discretizing the AR(1) process for productivity to a Markov chain featuring $N = 25$ different productivity values. We then conjecture an equilibrium function $\hat{p}(z)$. Given our definition of equilibrium and the results of Proposition 1, we look for equilibria where $\hat{p}(z) \in [p^*(\bar{z}), p^*(\underline{z})]$ for each z, and $\hat{p}(z)$ is decreasing in z. Our initial guess for $\hat{p}(z)$ is given by $p^*(z)$ for all z. We experiment with different initial guesses and found that the algorithm always converges to the same equilibrium.

Given the guess for $\hat{p}(z)$, we can compute the continuation value of each customer as a function of the current price and productivity, i.e. $\bar{V}(p, z)$, and solve for the optimal search and exit thresholds. Given $\hat{p}(z)$ and the customers' search and exit thresholds we can solve for the distributions of customers $Q(\cdot)$ and $H(\cdot)$ as defined in Definition 1. Notice that the latter also amounts to solve for a fixed point in the space of functions. Here, standard arguments for the existence of a solution to invariant distribution for Markov chains apply. Therefore, the assumption that $F(z'|z) > 0$ and $\Delta(\hat{p}(z), z) > 0$ ensure the existence of a unique $K(z)$. Finally, given $Q(\cdot)$, $H(\cdot)$, $\hat{p}(z)$ and $\bar{V}(p, z)$, we solve the firm problem and the obtain optimal firm prices given by the function $\hat{p}(z)$. We use $\hat{p}(z)$ to update our conjecture about equilibrium prices $\hat{p}(z)$, and iterate this procedure until convergence to a fixed point where $\hat{p}(z) = \hat{p}(z)$ for all $z \in [\underline{z}, \bar{z}]$.

Once we have solved for the equilibrium of the model at given parameter values. We then evaluate the objective function $(v_d - v(\Omega_n))'\Sigma\ (v_d - v(\Omega_n))$ at each iteration. We assume the weighting matrix Σ to be the identity matrix. We select as estimates the parameter values from the proposed grid that minimize the objective function and check that the optimum in the interior of the assumed grid.

E Extension: unforeseen aggregate shocks

In this appendix we provide details on how the model can be extended to use it to evaluate the role of aggregate shocks. In particular we consider the dynamics following an unforeseen aggregate shock that takes the economy temporarily away from the steady state, and study its convergence back to the initial steady state. To do so we need to allow for the possibility that the aggregate state varies over time. This means that the key equations of the model (listed below) will now be indexed by a time subscript t, capturing the dynamics in the aggregate state. As we want to study the effects of aggregate shocks in general equilibrium we also add a stylized model of the labor market so that household income is endogenously determined. The worker chooses the path of ℓ_t that maximizes household preferences in equation (14).

The production technology of the perfectly competitively sold good (good n) is linear in labor, so that its supply is given by $y_t^n = Z_t \, \ell_t^n$, where Z_t is aggregate productivity, and ℓ_t^n is labor demand by this firm. The production technology of the other good (good d) is also linear in labor, so that its supply is given by $y_t^j = Z_t \, z_t^j \, \ell_t^j$, where Z_t is aggregate productivity, and ℓ_t^n is labor demand by this firm, where j indexes one particular producer. Perfect competition in the market for variety n and in the labor market implies that workers are paid a wage equal to the marginal productivity of labor so that $w_t = q_t \, Z_t$. Equilibrium in the labor markets requires $\ell_t = \ell_t^n + \int_0^1 \ell_t^j \, dj$.

The value function of each shopper is given by

$$V_t(p, z, \psi) = \max \left\{ \bar{V}_t(p, z) \, , \, \hat{V}_t(p, z) - \psi \right\} \ , \tag{16}$$

where

$$\hat{V}_t(p, z) = \int_{-\infty}^{+\infty} \max \left\{ \bar{V}_t(p, z) \, , \, x \right\} dH_t(x) \ , \tag{17}$$

and

$$\begin{aligned} \bar{V}_t(p, z) = \xi_t \, v_t(p) + \beta \, (1 - q) \, \mathbb{E}_G \left[\int_{\underline{z}}^{\bar{z}} V_{t+1}(\hat{p}_{t+1}(x), x, \psi') \, dF(x|z) \right] + \\ + \beta \, q \, \mathbb{E}_G \left[\int_{\underline{z}}^{\bar{z}} V_{t+1}(\hat{p}_{t+1}(x), x, \psi') \, d\bar{F}(x) \right] \ . \end{aligned} \tag{18}$$

where ξ_t is the exogenous utility shifter, with

$$v_t(p) = \max_{d, \ n} \ \frac{\left(d^{\frac{\theta-1}{\theta}} + n^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}(1-\gamma)}}{1 - \gamma} \tag{19}$$

$$\text{s.t.} \quad p \, d + \, n \leq I_t \ , \tag{20}$$

The first order condition to the problem in equations (19)-(20) delivers the following standard downward sloping demand function for variety d

$$d_t(p) = \frac{I_t}{P} \left(\frac{p}{P} \right)^{-\theta} \ . \tag{21}$$

where $P = ((p)^{1-\theta}+1)^{\frac{1}{1-\theta}}$ is the price of the consumption basket. The solution to the shopper search problem gives a threshold

$$\hat{\psi}_t(p,z) \equiv \int_{\bar{V}_t(p,z)}^{\infty} \left(x - \bar{V}_t(p,z)\right) dH_t(x) \geq 0 \ .$$

The equilibrium pricing function $\hat{p}_t(z)$ is given by the solution to the firm pricing problem

$$W_t(z) \ = \ \max_p \ \Delta_t(p,z)\,\pi(p,z) + \Delta_t(p,z)\,\tilde{\beta}_t\,(1-q) \int_{\underline{z}}^{\bar{z}} W_{t+1}(z')dF(z'|\,z) \ , \tag{22}$$

where

$$\Delta_t(p,z) \ \ \equiv 1 - \underbrace{G\left(\hat{\psi}_t(p,z)\right)\left(1 - \ H_t(\bar{V}_t(p,z))\right)}_{\text{customers outflow}} + \underbrace{Q_t\left(\bar{V}_t(p,z)\right)}_{\text{customers inflow}} \ , \tag{23}$$

and

$$\tilde{\beta}_t \equiv \beta\,\frac{\int_0^1 \left(c_{t+1}(i)\right)^{-\gamma}/P_{t+1}(i)\,di}{\int_0^1 \left(c_t(i)\right)^{-\gamma}/P_t(i)\,di} \ ,$$

where $\int_0^1 \left(c_t(i)\right)^{-\gamma}/P_t(i)\,di$ is the household marginal increase in utility with respect to nominal income; $c_t(i)$ denotes customer i's consumption basket in period t, and $P_t(i)$ is the associated price index.

The equilibrium distributions $H_t(\cdot)$ and $Q_t(\cdot)$ are given

$$H_t(x) = K_t(\hat{z}(x)) \quad \text{and} \quad Q_t(x) = \int_{\underline{z}}^{\hat{z}(x)} G(\hat{\psi}_t(\hat{p}_t(z),z))\,dK_t(z) \ ,$$

for each $x \in [\mathcal{V}_t(\underline{z}), \mathcal{V}_t(\bar{z})]$, where $\hat{z}(x) = \max\{z \in [\underline{z},\bar{z}] : \mathcal{V}_t(z) \leq x\}$, $\mathcal{V}_t(z) = \bar{V}_t(\hat{p}_t(z),z)$, and

$$K_t(z) = (1-q)\int_{\underline{z}}^{z}\int_{\underline{z}}^{\bar{z}} \Delta_{t-1}(\hat{p}_{t-1}(x),x)\,dF(s|x)\,dK_{t-1}(x) + q\int_{\underline{z}}^{z} d\bar{F}(x) \ .$$

F Preference versus government spending shocks

In this section we compare the response of our baseline economy to two other types of aggregate demand shocks, and show that it predicts similar patterns for markups and price dispersion irrespectively of the type of demand shock. The first type of demand shock is a preference shock to the consumption/leisure margin increasing the utility flow of consumption relatively to utility from leisure. In this case the utility flow of consumption is given by $\xi_t\, v_t(p)$

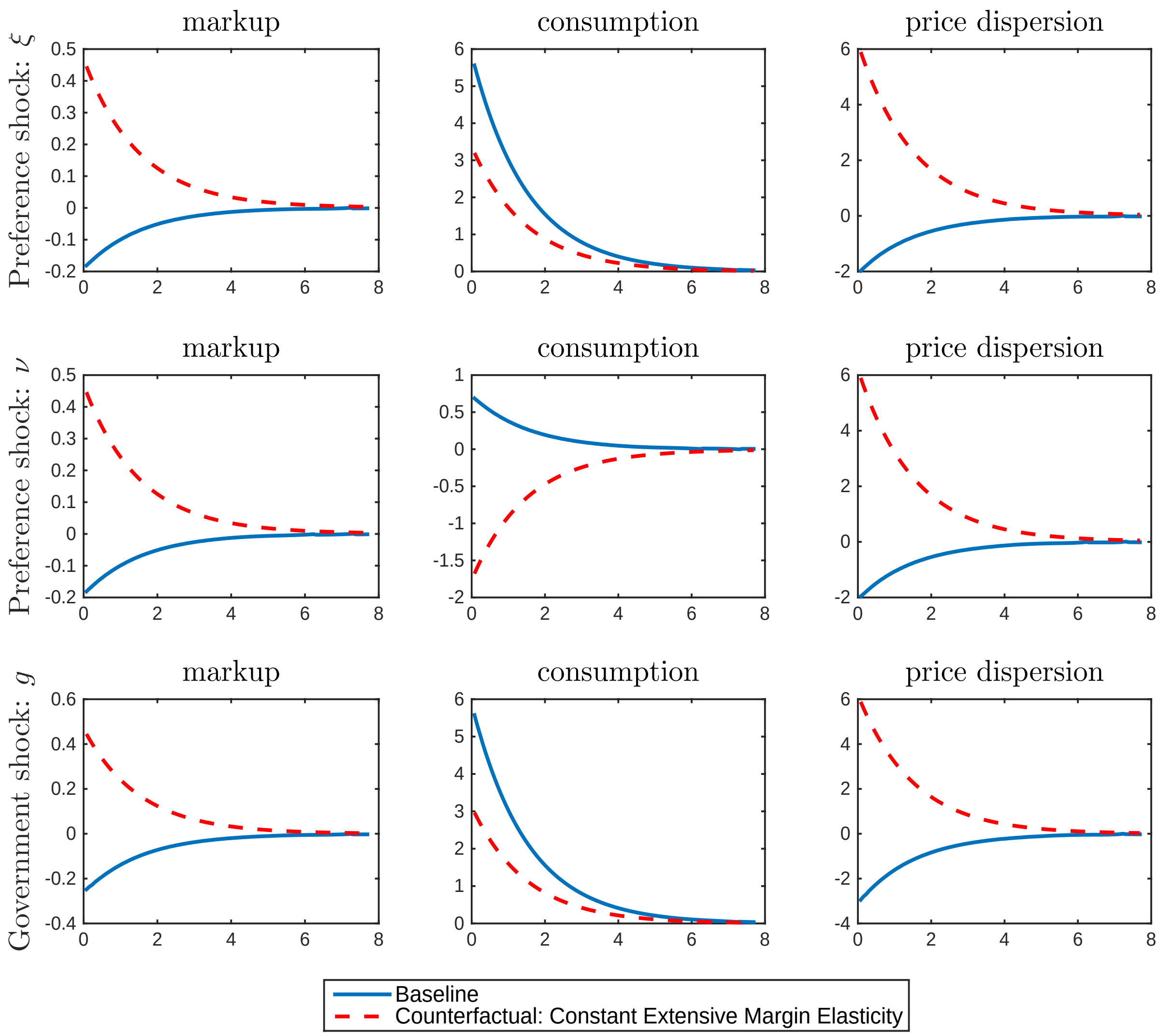
where ξ_t is the utility shifter and $v_t(p)$ is given by equation (19). This is the shock we study in Section 5.1 of the main text.

The second type of demand shock is a preference shock to the consumption-leisure/saving margin. This shock is a shift ν_t to the overall utility flow from both leisure and consumption. In this case the utility flow of consumption is given by $\nu_t \, v_t(p)$ where $v_t(p)$ is given by equation (19), whereas the disutility from labor in equation (14) becomes $\nu_t \, \ell_t^{1+\phi}/(1+\phi)$.

The third type of shock is a government spending shock g_t. The government shock is financed with lump-sum taxes in units of the final consumption good c_t. For simplicity, the government expenditure is assumed to follow the same structure of private expenditure and be allocated across the different suppliers according to the market share of each of them. Then, in terms of the equations determining the equilibrium dynamics of our economy, the main change concerns the disposable income to households which reduces from I_t to $I_t - g_t \, P_t$, where P_t is the price of the consumption basket in units of the numeraire. Thus while demand from private agents is now $m_t \, d_t(p) = m_t \, (I_t/P_t - g_t) \, (p/P_t)^{-\theta}$, whereas total demand for a firm also includes government spending and it is given by $m_t \, d_t(p) = m_t \, I_t/P_t \, (p/P_t)^{-\theta}$. Consumption of the basket by private agents reduces from I_t/P_t to $I_t/P_t - g_t$, so that the utility flow from consumption is given by $v_t(p) = (I_t/P_t - g_t)^{1-\gamma}/(1-\gamma)$.

In Figure 9 we report impulse responses of average markup, consumption and price dispersion to the three demand shocks in our baseline economy (blue solid line) against the counterfactual economy with invariant extensive margin elasticity (red dashed line). The first row reports impulse responses to the preference shock ξ_t, where $\xi_t = 1$ in steady state, $\xi_t = 1.05$on impact of the shock and $\log(\xi_t)$ mean reverts to steady state with an AR(1) coefficient of 0.95. The second row uses the same parametrization but with respect to the preference shock ν_t, i.e. the case considered in the main text. The third row reports the impulse responses to the government spending shock, where the shock is equivalent to 5% of steady state consumption and dies out in logs with an AR(1) coefficient of 0.95. We notice that all three shocks deliver similar impulse responses of markups and price dispersion in the baseline economy and in the counterfactual economy. The response of consumption to the government spending shock (g) and the preference shock (ξ) is stronger than to the preference shock (ν) because the latter scales up both the utility of consumption and the disutility of working, so that aggregate consumption increases only because markup decrease in the baseline economy. In contrast, g and ξ have a positive direct effect on demand that is magnified (dampened) by the reduction (increase) in markup of our baseline (counterfactual) model.

Figure 9: Impulse responses to an aggregate demand shock, in % deviations from s.s.



Notes: All plots report the impulse response to a same 5% shock for different outcomes of the model; the figures are averages where firms are weighted by the size their customer base. The blue solid line documents the response in our baseline model; the dashed red line shows the results in the counterfactual economy, where the extensive margin elasticity is invariant to the aggregate shock.

G Extension: search cost as leisure time

We consider the same framework described in Appendix E, where however now the representative household is composed by a mass one of customers/workers. All customers/workers supply the same amount of labor in any period t, ℓ_t, which they take as given and is decided by the representative household who then splits labor income equally across the shoppers. The value function of each customer/worker is given by

$$V_t(p, z, \psi) = \max\left\{\bar{V}_t(p, z)\,,\, \hat{V}_t(p, z) - \Psi(\psi, \ell_t)\right\}\ , \tag{24}$$

where

$$\Psi(\psi, \ell_t) = \frac{(\ell_T + \psi)^{1+\phi}}{1+\phi} - \frac{\ell_t^{1+\phi}}{1+\phi}$$

is the search cost in terms of extra disutility from less leisure,

$$\hat{V}_t(p, z) = \int_{-\infty}^{+\infty} \max\left\{\bar{V}_t(p, z)\,,\, x\right\} dH_t(x)\ , \tag{25}$$

and

$$\bar{V}_t(p, z) = v_t(p) - \frac{\ell_t^{1+\phi}}{1+\phi} + \beta\,(1-q)\,\mathbb{E}_G\left[\int_{\underline{z}}^{\bar{z}} V_{t+1}(\hat{p}_{t+1}(x), x, \psi')\, dF(x|z)\right] + \tag{26}$$
$$+\,\beta\, q\,\mathbb{E}_G\left[\int_{\underline{z}}^{\bar{z}} V_{t+1}(\hat{p}_{t+1}(x), x, \psi')\, d\bar{F}(x)\right]\ .$$

The solution to the shopper search problem gives a threshold in terms of labor disutility

$$\hat{\Psi}_t(p, z) \equiv \int_{\bar{V}_t(p,z)}^{\infty} \left(x - \bar{V}_t(p, z)\right) dH_t(x) \geq 0\ ,$$

which can be mapped into leisure time

$$\hat{\psi}_t(p, z) = \left[\ell_t^{1+\phi} + (1+\phi)\,\hat{\Psi}_t(p, z)\right]^{\frac{1}{1+\phi}} - \ell_t\ .$$

The remaining equations are as in Appendix E.

H Cost pass-through

Seeing as the model links the distribution of search cost to the extent of cost pass-through exercised by firms, we use a measure of pass-through computed in our data as a target moment in the estimation. Our estimates are based on the retailer Price Data which include both prices and cost for every item. To measure pass-through of idiosyncratic shocks, we regress the log-price index of each store in a given week on its log-cost index. The price index p_t^j is constructed as described in Section 3.3 and the cost index is analogously computed using the data on replacement cost provided by the retailer. To avoid inflating the short-term (weekly) pass-through due to the persistence of both price and cost variables, we include in the specification lagged values of the independent variable. We experiment with an alternative way to deal with the persistence of the dependent variable by measuring the short-term pass-through using first differences. Finally, we include time and market fixed effects to control for aggregate trend (e.g. demand shocks) that can move prices independently from cost shifts. The pass-through we recover ranges between 13% and 24%.

Table 6: Pass-through of idiosyncratic cost shocks

	(1)	**(2)**	**(3)**	**(4)**
Dep. variable	$\log(p_t^j)$	$\log(p_t^j)$	$\Delta\log(p_t^j)$	$\Delta\log(p_t^j)$
$\log(c_t^j)$	0.17***	0.24***		
	(0.04)	(0.09)		
$\log(c_{t-1}^j)$	0.04	0.06		
	(0.03)	(0.04)		
$\log(c_{t-2}^j)$	-0.01	0.02		
	(0.05)	(0.07)		
$\log(c_{t-3}^j)$	0.06***	0.05*		
	(0.02)	(0.03)		
$\log(c_{t-4}^j)$	0.07	0.07		
	(0.05)	(0.05)		
$\Delta\log(c_t^j)$			0.13*	0.11*
			(0.07)	(0.06)
$\Delta\log(c_t^j)$				0.40**
×High productivity				(0.15)
High productivity				-0.00
			(0.00)	
Observations	12,915	8,295	8,295	8,295
MSA f.e.	No	Yes	Yes	Yes
Time f.e.	No	Yes	Yes	Yes

Notes: An observation is a store(j)-week(t) pair. The dependent variable is the price index of the store and the independent variables are the cost index of the store and its lags. Standard errors are in parenthesis and are clustered at the store level. ***: Significant at 1% **: Significant at 5% *: Significant at 10%.

In column (4), we experiment with a specification that allows to test the model prediction that firms with higher productivity (i.e. lower cost) should display higher pass-through. We consider the distribution of store price index cost for each store and construct an indicator variable "High productivity" that takes value 1 when a store's cost is in the bottom tercile of the distribution.[40] We include the "High productivity" dummy and its interaction with the store cost variable in the specification to estimate short run pass-through; we cannot replicate the same exercise for long run pass-through because the lagged realization of cost may lie in different terciles of the cost distribution making it ambiguous when we could characterize a firm as being high productivity.

The average pass-through we recover is similar to that obtained in columns (1)-(3). The positive large and significant coefficient on the interaction term, however, means that cost pass-through is much higher when firms are more productive. The pass-through of a high productivity store is still incomplete but it is over three times larger than the average pass-through of stores in the top two terciles of their cost distribution.

I Price dispersion and demand shocks

The measure of price dispersion we use in Section 5.3 is constructed based on the IRI data. We exploit two measures of price dispersion: at the level of the UPC and at the level of the store basket.

UPC price dispersion

We focus on 100 UPCs for each of the 30 product categories in the IRI data. The UPC selected are among the highest seller so that there is no missing price information for any of the stores in the sample in any of the weeks. To remove the effect of staggered adjustment (possibly induced by the presence of nominal rigidities) we only retain prices that result from a price adjustment. Formally we define these prices as:

$$\mathring{p}^j_{umt} = \begin{cases} p^j_{umt} & \text{if } p^j_{umt} \neq p^j_{umt-1} \ , \\ . & \text{otherwise} \ . \end{cases} \tag{27}$$

where u denotes the UPC, j the store, m the market and t the week.

To remove the effect of inflation, we standardize the price of a UPC in a particular store using the average price of the same UPC across all the stores in the same market in the week. This implies that our measure of UPC dispersion amounts to calculating the standard deviation of the following object: $\mathring{p}^j_{umt} \backslash (\frac{1}{\# s \in m(j)} \sum_{s \in m(j)} \mathring{p}^j_{umt})$.

[40]This is an imperfect test of our model implication, which refers to the productivity of a firm with respect to the cross-sectional distribution in a given week and not to the cross-time distribution for a given store.

Basket price dispersion

The price dispersion of the store basket (calculated as described in Section 3.3) is calculated in a way analogous to that just described for UPC-level price dispersion.

Correlation between price dispersion and demand

The results in Table 4 regard a regression of the price dispersion whose construction we just discussed on a measure of demand. We use the average expenditure in food reported in the CEX and are therefore limiting the analysis to Metropolitan Statistical Areas for which the CEX report such information (and that appear in the IRI database): Atlanta, Boston, Chicago, Cleveland, Dallas, Detroit, Houston, Kansas City, Los Angeles, Milwaukee, Minneapolis\St. Paul, New York, Philadelphia, San Diego, San Francisco, Seattle\Tacoma, St. Louis and Washington, DC. Finally, we limit the time series to the years 2003-2011 to ensure that the set of stores active within each MSA is stable and entry and exit dynamics are not affecting our measure of price dispersion.