

Markups 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

February 14, 2014

(Preliminary and Incomplete)

Abstract

We study a model where customers face frictions when changing their supplier, generating sluggishness in the firm's customer base. Firms care about expanding their customer base and this affects their pricing strategy. We characterize optimal pricing in this model and estimate it using data on the evolution of the customer base of a large US retailer. The introduction of customer markets reduces average markups, more markedly for less productive firms. We use the model to perform a counterfactual exercise and investigate the cyclical behaviour of markups in response to both aggregate supply and demand shocks.

JEL classification:

Keywords:

*We benefited from the comments at the Minnesota Workshop in Macroeconomic Theory, 2nd Rome Junior Macroeconomics conference, 2nd Annual UTDT conference in Advances in Economics, 10th Philadelphia Search and Matching conference, ESSET 2013, MBF-Bicocca conference, and seminars at Columbia University, Federal Reserve Bank of Richmond, the Ohio State University, University of Pennsylvania, Macro Faculty Lunch at Stanford, and University of Tor Vergata. We thank Mark Aguiar, George Alessandria, Fernando Alvarez, Lukasz Drozd, Huberto Ennis, Stefania Garetto, Mike Golosov, Bob Hall, Jonathan Heathcote, Hugo Hopenhayn, Eric Hurst, Pat Kehoe, Pete Klenow, Francesco Lippi, Ricardo Lagos, Erzo Luttmer, Kiminori Matsuyama, Guido Menzio, Dale Mortensen, Juan Pablo Nicolini, Ezra Oberfield, Guillermo Ordonez, Facundo Piguillem, Valerie Ramey, Leena Rudanko, Stephanie Schmitt-Grohe, Martin Uribe, Randall Wright. Luigi Paciello thanks Stanford University for 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

There is ample consensus that dynamics in the firms' customer base -i.e. the set of customers currently purchasing from a firm- are important determinants of performance, and that firms act to influence its evolution (Foster et al. (2012)). However, how relevant customer markets are for the firms' optimal pricing policy is still an open question (Blanchard (2009)). As in Bils (1989), we refer to customer markets to markets where consumers are attached to a firm, store, or product, that they previously purchased. In this paper we study a model of price setting with customer markets. We characterize the equilibrium of the model and estimate it using novel micro data. Using the estimated model, we study the implications of customer markets for firm's pricing policy. We show that introducing a friction that increases customers' attachment to their current supplier into a standard price setting model has relevant consequences for equilibrium price dynamics. In particular, we show that it impacts the cyclicity of markups in response to both idiosyncratic and aggregate shocks.

We build on the seminal work on customer markets by Phelps and Winter (1970) and explicitly model the game between a customer and a firm. In particular, we model inertia in customer dynamics by introducing search frictions along the lines of Burdett and Coles (1997).¹ Customers demand a homogeneous good sold in different locations, each of them populated by a single firm. Every customer is matched to a particular firm at any point in time and decides whether to look for a new supplier based on the price posted by her current one. If the customer decides to search, she incurs in a cost and draws a price quote from the population of firms, where the probability of matching is proportional to a firm's market share. After observing the price of the potential new supplier, the customer decides if she wants to switch or stay in the customer base of the firm she currently patronizes. The presence of a search cost increases inertia in the customer base.

The stickiness in the customer base leads firms to consider their customers as an asset. They price taking into account the impact on consumer dynamics; we assume that there is no possibility for commitment nor price discrimination. The dynamic problem of a firm is linked to two different margins of demand. The first one is the static *intensive margin* and concerns *how much* of the good the firm is able to sell to each of its customers, conditional on the customer buying from the firm. The second margin, which we label *extensive margin* of demand, concerns the decision of customers to patronize a particular firm, and therefore whether to belong to its customer base. This dynamic margin relates to *how many* individuals decide to remain and buy from their old firm. Firms are characterized by different productivity levels, which evolve stochastically through time according to a Markovian

¹This approach is also used in the context of labor markets. See for instance Burdett and Mortensen (1998), Coles (2001) and Coles and Mortensen (2013).

process. The combination of search frictions on the consumer side and idiosyncratic firm productivity gives rise to equilibrium price dispersion, incentivizing customers to search.

We obtain several noteworthy results from the characterization of our model. First, we show that the optimal markup is a function of both the intensive and the extensive margin elasticity of demand. When customer concerns bind, the optimal markup is lower than the markup that maximizes static profits. The higher the extensive margin elasticity, the lower the optimal markup. Moreover, the extensive margin is linked to a change in demand that last for several periods, as lost customers are hard to win back. Therefore, it weights more in the determination of prices than the intensive margin elasticity which only affects current demand. Even a small extensive margin elasticity can, therefore, have a large impact on optimal markups.

Looking at the cross-section of sellers within an industry, we show that shifts in productivity change the demand faced by the firm. In a Markovian equilibrium with persistence in productivity, current productivity carries information about future prices. If higher productivity reflects in lower prices, customers expect highly productive firms to offer better prices in the future. As a consequence, they are more reluctant to leave them and firms with higher productivity face lower extensive margin elasticity of demand and charge a higher markup.

We complement our modelling effort with a two-pronged empirical analysis. From a descriptive point of view, we provide novel direct evidence of sluggishness in the customer base by looking at the evolution of the customer base of a large U.S. supermarket chain. Despite the growing attention towards the implications of customer retention, in fact, evidence in that regards is still mostly based on anecdotal or survey sources. Supermarket scanner data are particularly well suited to study customer base dynamics because they allow to track precisely customers' choice of the store where to buy. This allows us to document that there is persistence in a customer's choice of a retailer, and to show that a shopper decision to exit the customer base responds to the relevant price. On average, a one percent change in the price of the customer's typical basket of goods raises her likelihood of leaving the retailer by 0.015 percentage points. This estimate is important as it identifies the size of the friction determining customer dynamics in our model. Therefore, we target the extensive margin elasticity implied from our descriptive regression, as well as other standard macroeconomic targets, to quantify the model.

We use the estimated model to quantify the relevance of customer markets for price dynamics. As a benchmark, we compare to an alternative economy where we shut down the customer retention concerns, so that it approximates a model with standard constant elasticity of substitution (CES) preferences. At the parameter estimates the implied average markup is half of the equivalent statistic in the CES economy. Our model also generates

more dispersion in markups and, therefore, less dispersion in prices. Finally, in our model markups are highly responsive to idiosyncratic productivity (the elasticity is 0.85), consistent with evidence in [Petrin and Warzynski \(2012\)](#).

Our estimated model provides a laboratory to study the implications of customer markets for the response of markups to aggregate shocks. Although we are not the first to study the implications of frictions in customers' ability to change supplier for markup cyclicalities, we introduce a novel channel of transmission of aggregate shocks to markups which in our setting depend on the extensive elasticity of demand. Everything else being equal, a shock that increases households consumption also decreases the benefits to search and reduces the extensive margin elasticity.

Customer markets were first analyzed in the context of macroeconomics quantifiable models by [Phelps and Winter \(1970\)](#), and by [Rotemberg and Woodford \(1991\)](#), who modeled the flow of customers as a function of the price posted by the firm. We microfound these approaches by having customer dynamics arising endogenously by solving the game between firms and customers. The literature on "deep habits" ([Ravn et al. \(2006\)](#), [Nakamura and Steinsson \(2011\)](#)) represents an alternative way to generate persistence in demand by introducing habits in consumption.

Our analysis of the implications of build-up of a customer base for pricing and markup ties into a growing body of literature using models where the market share of the firm is sluggish to study a number of issues such as pricing-to-market ([Alessandria \(2009\)](#), [Drozd and Nosal \(2012\)](#)) and firm investment ([Gourio and Rudanko \(2011\)](#)). We focus on the influence of customer base concerns on firm price setting as in [Bils \(1989\)](#), [Burdett and Coles \(1997\)](#), [Menzio \(2007\)](#) and [Kleshchelski and Vincent \(2009\)](#). [Bils \(1989\)](#) studies a perfect foresight economy where a monopolist is faced with an exogenous stream of potential buyers, while we study an economy without perfect foresight and where firms compete for customers. [Burdett and Coles \(1997\)](#) studies the relationship of equilibrium price dispersion with firm size in a model with endogenous entry of firms, while we allow for idiosyncratic shocks to productivity and study its implications for markups. [Menzio \(2007\)](#) focuses on optimal price setting with commitment and asymmetric information between sellers and buyers, whereas we study a model with symmetric information and no commitment. Differently from [Kleshchelski and Vincent \(2009\)](#), we study an equilibrium where firms are not symmetric allowing us to explore the relationship between markups and productivity. None of these papers analyze the reaction to aggregate shocks.

By showing that consumers are influenced by prices in their decision of breaking long term relationships with suppliers, we add to the literature using scanner data to document empirical regularities in pricing and shopping behavior. A series of contributions ([Aguilar](#)

and Hurst (2007), Coibion et al. (2012), and Kaplan and Menzio (2013)) integrates store and customer scanner data to show that intensity of search for lower prices is depends on income and opportunity cost of time. We view our contribution as complementary to theirs as we instead focus on the elasticity of search to prices.

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. In Section 4 we discuss the calibration of the model. Results on markup dynamics, both with respect to idiosyncratic and aggregate shocks, are presented in Section 5. Section 7 concludes.

2 The model

The economy is populated by a measure one of firms producing an homogeneous good, and a measure Γ of customers.

Customers. We use the superscript $i \in [0, \Gamma]$, to index a customer. Let I denote the income level of a customer each period. Let $d(p)$ and $v(p)$ denote the static demand and customer surplus functions respectively which, naturally, only depend on the current price p . We assume that: (i) $d'(p) < 0$, $d''(p) \geq 0$, with $\lim_{p \rightarrow \infty} d(p) = 0$; (ii) $\varepsilon_d(p) \equiv -\partial \ln d(p) / \partial \ln p \geq 1$; (iii) $\varepsilon'_d(p) \geq 0$, and (iv) $v'(p) < 0$; $v''(p) \leq 0$. Assumption (i) states that the demand function is decreasing and convex in prices and that the demand approaches zero as the price diverges, (ii) is required for positive markups, assumption (iii) implies that the demand elasticity increases with prices, while assumption (iv) simply states that the surplus of the customer is decreasing and concave in the price. Assumptions (i)-(iv) are typical properties arising in standard macro models with CRRA utility functions and CES demand across a large set of varieties. In Appendix A.1, we derive properties (i)-(iv) from first principles. Notice that both demand and customer surplus functions do not depend on the identity of the customer i . Nevertheless, customers face different realizations of both demand and surplus given they face different prices p because customers are matched to different producers. Moreover, notice that we allow for the elasticity of demand $\varepsilon_d(p)$ to depend on the price p to nest recent models of variable markups and incomplete cost pass-through.² This allows to compare our mechanism to the existing literature.

Firms. We use the superscript j , $j \in [0, 1]$, to index a firm. The production technology of the good is linear in the unique production input, ℓ , and depends on the firm specific productivity z^j . That is, $y^j = z^j \ell$. We let w denote the marginal cost of the input ℓ , p denote the price of the good, and $\pi(p, z) \equiv d(p)(p - w/z)$ denote the profit per customer.

²See Atkeson and Burstein (2008), Melitz and Ottaviano (2008), and Berman et al. (2012) for a review.

We assume that $\pi(p, z)$ is single-peaked. We assume that firm specific productivity z^j is i.i.d. across firms. Let $z \in [\underline{z}, \bar{z}]$ with $\bar{z} > \underline{z} > 0$, and $f(z)$ denote the unconditional density of productivity z . Also let $f(z'|z)$ denote the conditional density function of z' given z . We assume that the conditional distribution of the productivity shock, defined as $F(z'|z) \equiv \int_{\underline{z}}^{z'} f(s|z)ds$, satisfies that $F(z'|z_h) < F(z'|z_l)$ for any z' and $z_h > z_l$, i.e. $F(z'|z_h)$ first order stochastically dominates $F(z'|z_l)$. Finally, let $m^j \in [0, \Gamma]$ denote the mass of customers that are currently buying from firm j .

Search and matching. Customers draw a random search cost $\psi \in [0, \infty)$ each period from the same distribution $G(\psi)$, with $g(\psi) \equiv \partial G(\psi)/\partial \psi$. We restrict our attention to density functions that are continuous on all the support of ψ . A relatively flexible distribution guaranteeing such properties is the Gamma distribution, characterized by a shape parameter, ζ , and a scale parameter, λ . The Gamma distribution satisfies the continuity requirement as long as $\zeta > 1$. If the customer incurs in the utility cost ψ , she draws a price quote from the distribution of firms, where the probability of receiving a price quote from firm j depends on the size of the firm's customer base. This captures the idea that larger firms attract more customers. The customer then compares both firms and decides between remaining with the firm she started the period matched to, or to exit the customer base of this firm and join that one of the new firm. We allow customers to search at most once per period, and assume that once the customer has exited the customer base of a firm she cannot return.

Timing of events. A period starts with all customers distributed across firms; we let m_{t-1}^j denote the mass of customers matched with firm j at the end of period $t-1$ and that therefore start period t matched with firm j . The timing of events is the following: (i) productivity shocks z_t^j are realized and firm j posts price p_t^j , (ii) each customer draws her search cost ψ_t^i and observes p_t^j as well as the relevant state of firm j (i.e. z_t^j and m_{t-1}^j), (iii) each customer decides between searching for a new firm -after incurring in search cost ψ_t^i -, or remaining matched to firm j -and not incurring in cost ψ_t^i -, (iv) if the customer incurs the search cost she draws a new supplier j' with probability $m_{t-1}^{j'}/\Gamma$, observes the price and state of firm j' , and decides if to exit the customer base of firm j , and (v) customer surplus $v(p_t^j)$ and profits $\pi(p_t^j, z_t^j)$ are realized.

2.1 The game between customers and firms

In this section we discuss the solution to the customers' search and firms' price-setting problems. Given the timing of the model, a firm and the customers matched to it play an anonymous sequential game, where the firm plays first by posting a price and customers best respond to it. We look for a Markov Perfect equilibrium where strategies are a function of the current state. Because neither the firm nor its customers can commit to future

actions, they take as given the continuation value. This continuation value is encoded in the customer base size m and productivity level z , where we omitted the time index t and identity index j to ease on notation. Notice that there would be time inconsistency if instead firms were setting a path for prices. This would happen as a firm would promise in the current period to charge low prices in the future to increase current profits and attract more customers today, but when the future arrives the firm would renege on its promise. Let $\tilde{\mathcal{P}}(m, z) : [0, \infty) \times [\underline{z}, \bar{z}] \rightarrow [0, \infty)$ be the pricing strategy of the firm mapping m and z to the space of prices. Also, let $\tilde{s}(p, z, m, \psi) : [0, \infty) \times [\underline{z}, \bar{z}] \times [0, \infty) \times [0, \infty) \rightarrow \{0, 1\}$ be the search strategy of the customer mapping the price p , the productivity z , the customer base m , as well as the search cost ψ to $\{0, 1\}$, where $\tilde{s} = 1$ denotes the case of searching. Let $\tilde{e}(p, z, m, p', z', m') : [0, \infty) \times [\underline{z}, \bar{z}] \times [0, \infty) \times [0, \infty) \times [\underline{z}, \bar{z}] \times [0, \infty) \rightarrow \{0, 1\}$ denote the exit strategy of a customer who, conditional on searching, draws a firm with price p' , productivity z' , and customer base size m' , where $\tilde{e} = 1$ denotes the case of exiting.

In our setup, as we will show later, the customer base enters multiplicatively in the firm's problem, i.e. the value function of the firm is homogeneous in m . As a result, we will construct an equilibrium where strategies do not depend on the customer base size. Furthermore, we will study equilibria where the present discounted value to a customer of being matched and buying from a firm with productivity z , an object that we denote by $\hat{V}(z)$, is increasing in z . Studying equilibria that satisfy this property is natural as firms with higher productivity can charge lower prices and there is persistence in the process for productivity, so that firms with higher productivity can provide higher value to customers matched to it. In the end, strategies reduce to the following: (i) $\mathcal{P}(z) : [\underline{z}, \bar{z}] \rightarrow [0, \infty)$ denotes the pricing strategy of the firm mapping z to the space of prices, (ii) $s(p, z, \psi) : [0, \infty) \times [\underline{z}, \bar{z}] \times [0, \infty) \rightarrow \{0, 1\}$ denotes the search strategy of the customer mapping z , p and ψ to $\{0, 1\}$, and (iii) $e(p, z, z') : [0, \infty) \times [\underline{z}, \bar{z}] \times [\underline{z}, \bar{z}] \rightarrow \{0, 1\}$ denotes the exit strategy of the customer mapping p , z , and z' to $\{0, 1\}$.

It is useful to define some equilibrium objects. Let $H(z) : [\underline{z}, \bar{z}] \rightarrow [0, \Gamma]$ the equilibrium invariant measure of customers over productivity levels, i.e. the mass of customers matched to firms with productivity smaller or equal than z . Let $Q(z) : [\underline{z}, \bar{z}] \rightarrow [0, \Gamma]$ denote the equilibrium invariant measure of *searching* customers over firms productivity, i.e. the mass of searching customers coming from a firm with productivity smaller or equal than z . While H and Q are equilibrium objects, we will assume throughout the paper that $H(z)$ and $Q(z)$ are continuously differentiable with densities $h(z) \equiv \partial H(z)/\partial z$ and $q(z) \equiv \partial Q(z)/\partial z$. We later verify that, in the equilibrium we explore, this is true.

2.2 The problem of a customer

Consider a customer buying goods from a given firm j . Let $V(p, z, \psi)$ be the value function for a customer matched to firm j , characterized by productivity z and price p , and that has drawn a search cost equal to ψ . We have,

$$V(p_t^j, z_t^j, \psi_t^i) = \max \left\{ \bar{V}(p_t^j, z_t^j), \tilde{V}(p_t^j, z_t^j) - \psi_t^i \right\} ,$$

where $\bar{V}(p, z)$ is the customer's value if she does not search, and $\tilde{V}(p, z)$ is the value, not including the search cost, if she does search. The value in the case of not searching is given by

$$\bar{V}(p_t^j, z_t^j) = v(p_t^j) + \beta \int_0^\infty \int_{\tilde{z}}^{\bar{z}} V(\mathcal{P}(z'), z', \psi') dF(z'|z_t^j) dG(\psi') ,$$

which follows from the assumptions of Markov equilibria where future prices are given by a function $\mathcal{P}(\cdot)$ mapping future productivity into prices, and of search costs ψ being i.i.d. over time. The value when searching is given by

$$\tilde{V}(p_t, z_t) = \int_{\tilde{z}}^{\bar{z}} \max \left\{ \bar{V}(p_t, z_t), \bar{V}(\mathcal{P}(z), z) \right\} \frac{dH(z)}{\Gamma} ,$$

where the customer takes expectations over all possible draws of potential new firms which, given the assumed Markovian structure of prices, only depend on the firm's productivity level $\tilde{z}$, and conditional on each of the possible draws decides if leaving her current firm and joining the customer base of the new one.

The following proposition describes the customer's optimal search and exit policy rules.

Proposition 1 *Let $\hat{V}(z) \equiv \bar{V}(\mathcal{P}(z), z)$ be increasing in z . The exit strategy takes the form of a trigger such that the customer matched to a firm with productivity z_t^j charging price p_t^j exits if she draws a new firm with productivity higher than a threshold, i.e.,*

$$e(p_t^j, z_t^j, z_t^{j'}) = \begin{cases} 1 & \text{if } z_t^{j'} \geq \hat{z}(p_t^j, z_t^j) , \\ 0 & \text{otherwise} , \end{cases} \quad (1)$$

where the threshold is given by $\hat{z}(p, z) = \hat{V}^{-1}(\bar{V}(p, z))$. Similarly, the search strategy takes the form of a trigger such that the customer matched to a firm with productivity z_t^j charging price p_t^j searches if she draws a search cost ψ_t^i smaller than a threshold, i.e.

$$s(p_t^j, z_t^j, \psi_t^i) = \begin{cases} 1 & \text{if } \psi_t^i \leq \hat{\psi}(p_t^j, z_t^j) , \\ 0 & \text{otherwise} , \end{cases} \quad (2)$$

where the threshold is given by $\hat{\psi}(p, z) = \tilde{V}(p, z) - \bar{V}(p, z) \geq 0$.

The proof of the proposition is in [Appendix A.2](#). The first part of the proposition describes the exit behavior of customers. Because the value of being matched with a given firm increases with its productivity, the customer exits the customer base of its current firm if it finds a new firm which productivity is high enough. We characterized this optimal behavior by the function $\hat{z}(p, z)$. The second part of the proposition describes the search behavior of customers. Given that search is costly, not all customers currently matched to a given firm exercise the search option; only those with a low search cost ψ do so. We characterized the threshold defining who searches by the function $\hat{\psi}(p, z)$. Notice the thresholds dependence on both the price of the firm, p , and its productivity, z . 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 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, a customer's conjecture about the firm's expected future prices are linked to the firm's current productivity through $F(z'|z)$. That is, current productivity z conveys information about the expected value of remaining in the customer base of a given firm.

A particularly useful way of describing the search threshold $\hat{\psi}(p, z)$ is one where the impact of the price level p on the search threshold is separable from the impact of the firm's productivity z . To this end, let $\bar{p}(z)$ denote the price level at which no customer would search at a firm with productivity z . This price solves $\bar{V}(\bar{p}(z), z) = \tilde{V}(\bar{p}(z), z)$ for each z . Notice that, when $\bar{V}(\mathcal{P}(z), z)$ is increasing in z , the previous condition reduces to $\bar{V}(\bar{p}(z), z) = \bar{V}(\mathcal{P}(\bar{z}), \bar{z})$ for all z . Moreover, because $\bar{V}(p, z)$ is monotonous in p , the value $\bar{p}(z)$ is unique, and it is increasing in z because $\hat{V}(z)$ is increasing in z . Using the definition of $\bar{p}(z)$, we can express the search threshold $\hat{\psi}(p, z)$ as

$$\hat{\psi}(p, z) = \begin{cases} 0 & \text{if } p \leq \bar{p}(z) , \\ v(\bar{p}(z)) - v(p) & \text{if } p > \bar{p}(z) . \end{cases} \quad (3)$$

The next lemma characterizes the thresholds $\hat{z}(p, z)$ and $\hat{\psi}(p, z)$.

Lemma 1 *The threshold $\hat{z}(p, z)$ is decreasing in p and increasing in z . The threshold $\hat{\psi}(p, z)$ is increasing in p and decreasing in z .*

The proof of [Lemma 1](#) is in [Appendix A.3](#) and follows from the assumptions of $v(p)$ decreasing in p , $\hat{V}(z)$ increasing in z and the persistence of the productivity process. An important

implication of the lemma is that, not only customers are more likely to search and exit from firms charging higher prices, but also that they are more likely to do so from firms with lower productivity. As previously discussed, this follows from the dependence of the expected future path of prices on the firm's current productivity, and those firms with lower productivity offering lower value to customers, something that follows from $\hat{V}(z)$ being increasing in z .

2.3 The problem of the firm

In [Section 2.2](#) we concluded that a customer searches for a new firm to patronize when, for a given price p and productivity level z , her search cost is low enough; we described this optimal behavior by the threshold function $\hat{\psi}(p, z)$. We also concluded that, upon searching for a new firm, the customer would exit the customer base of the firm she is currently patronizing if, for a given price p and productivity level z , the new firm has a high productivity level; we described this optimal behavior by the threshold function $\hat{z}(p, z)$. In this section we use that customers follow this policy to derive the optimal price set by firms.

The customer base of a firm; m . Recall that m_{t-1}^j denotes the mass of customers at the beginning of a period currently buying from firm j , and that $H(\cdot)$ and $Q(\cdot)$ denote the distributions of all customers and searching customers across productivity levels, respectively. The evolution of the customer base is

$$m_t^j = m_{t-1}^j \left[1 - G\left(\hat{\psi}(p_t^j, z_t^j)\right) \left(1 - \frac{H(\hat{z}(p_t^j, z_t^j))}{\Gamma}\right) \right] + \frac{m_{t-1}^j}{\Gamma} Q(\hat{z}(p_t^j, z_t^j)) , \quad (4)$$

The term m_{t-1}^j/Γ represents the fraction of searching customers that are matched to firm j , while $Q(\hat{z}(p_t^j, z_t^j))$ is the total mass of searching customers coming from a firm with productivity smaller than $\hat{z}(p_t^j, z_t^j)$. The term $G(\hat{\psi}(p_t^j, z_t^j))$ represents the fraction of customers at firm j which are searching, a fraction $1 - H(\hat{z}(p_t^j, z_t^j))/\Gamma$ of which will be matched to a firm with productivity larger than $\hat{z}(p_t^j, z_t^j)$ and therefore exit the customer base of their current firm.

We can express the customer base growth of the firm as $m_t^j = m_{t-1}^j \Delta(p_t^j, z_t^j)$, where the function $\Delta(\cdot)$ is given by

$$\Delta(p, z) \equiv 1 - G\left(\hat{\psi}(p, z)\right) \left(1 - \frac{1}{\Gamma} H(\hat{z}(p, z))\right) + \frac{1}{\Gamma} Q(\hat{z}(p, z)) , \quad (5)$$

Notice that the firm's growth is independent of its size, consistent with empirical evidence about Gibrat's Law (see [Luttmer \(2010\)](#)). This result originates from the linearity of the production function, coupled with the proportionality of the matching technology. The next lemma discusses the properties of the customer base growth with respect to prices and pro-

ductivity.

Lemma 2 *The customer base growth $\Delta(p, z)$ is decreasing in p and increasing in z .*

The proof of [Lemma 2](#) follows directly from [Lemma 1](#). The lemma states that the growth rate of the customer base of a firm is decreasing in the price it charges and increasing in its productivity. The fact that it is decreasing in the current price is straightforward: all else equal, a higher price induces customers to match with other firms. The fact that the growth rate of the firm increases with its productivity follows from the fact that customers expect this firm to be able to charge lower prices in the future, therefore increasing the value of being matched to it.

Profits. Recall that $\pi(p, z)$ denotes the period profits per customer of a firm charging price p , with productivity z ,

$$\pi(p, z) = d(p) \left(p - \frac{w}{z} \right), \quad (6)$$

from where it is immediate to see that $\pi(p, z)$ is strictly increasing and strictly concave in z . Given the assumptions we made regarding the demand function $d(p)$, the profit function is continuously differentiable in p . We also made assumptions regarding the single-peakness of the profit function. With this in mind, let $p^*(z)$ denote the price that maximizes static profits for a firm with current productivity z , i.e., $\partial\pi(p^*(z), z)/\partial p = 0$. We have that

$$p^*(z) = \frac{\varepsilon_d(p)}{\varepsilon_d(p) - 1} \frac{w}{z}, \quad (7)$$

where $\varepsilon_d(p)$ is the *intensive* margin demand elasticity as it controls the expenditure of those customers that decided to stay. In [Appendix A.1](#) we show that the existence of a unique maximizer $p^*(z)$ follows when the utility function of customers is CRRA with CES demand across a large set of varieties.

The problem of a firm. Let $\tilde{W}(z_t^j, m_{t-1}^j)$ be the value function for a firm with current productivity z_t^j and customer base m_{t-1}^j , so that $\tilde{W}(z_t^j, m_{t-1}^j)$ solves

$$\tilde{W}(z_t^j, m_{t-1}^j) = \max_p m_t^j \pi(p, z_t^j) + \beta \int_{\underline{z}}^{\bar{z}} \tilde{W}(z', m_t^j) f(z' | z_t) dz',$$

subject to [equation \(4\)](#) which determines customer dynamics and [equation \(6\)](#) which determines profits per customer. To ease on notation, we let $z = z_t^j$. We conjecture that the value function for 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 $W(z)$ solves

$$W(z) = \max_p \Delta(p, z) \left(\pi(p, z) + \beta \int_{\underline{z}}^{\bar{z}} W(z') f(z'|z) dz' \right), \quad (8)$$

where we used [equation \(4\)](#). Notice that, because the size of customer base does not appear in the problem presented in [equation \(8\)](#), two firms with the same productivity z but different customer base size face the same problem and therefore charge the same price; this happens because firms care about the growth rate of the customer base, not about its size.

Two important issues arise from inspection of the firm's problem presented in [equation \(8\)](#). The first one is that the effective discount rate, $\Delta(p, z)\beta$, is such that the operator for $W(z)$ is not necessarily a contraction; this happens because, for some z , the effective discount rate can be above one. To address this we make sure that the discount factor is low enough such that the effective discount rate is always below one; this restriction guarantees that the problem is a contraction. Notice that under the restriction, because the problem is a contraction, the homogeneity conjecture is also verified. The second issue is that the problem of the firm is not necessarily continuously differentiable nor globally concave in p . This implies that the first order condition is not necessary nor sufficient for the optimal price schedule; this stems from $\hat{\psi}(p, z)$, and its effect on the effective discount rate and period payoff. We later explore conditions under which the first order condition is necessary and sufficient in order to characterize the optimal price.

Let $\Pi(p, z) \equiv \pi(p, z) + \beta \int_{\underline{z}}^{\bar{z}} W(z') f(z'|z) dz'$ denote the value of a customer for a firm with productivity z . It is composed of two terms. The first term, $\pi(p, z)$ accounts for the profits the customer brings to the firm in the current period. The second term, $\beta \int_{\underline{z}}^{\bar{z}} W(z') f(z'|z) dz'$, accounts for the value that the customer brings to the firm in the future. Rewriting [equation \(8\)](#) we get that

$$W(z) = \max_p \Delta(p, z) \Pi(p, z),$$

which shows that the firm's optimal price trades-off the size of the customer base $\Delta(p, z)$ which is decreasing in p (*extensive margin*), with the value per customer $\Pi(p, z)$ which attains its maximum at $p^*(z)$ (*intensive margin*).

The next proposition characterizes the solution to the firm's maximization problem, and discusses cases in which the first order condition applies.

Proposition 2 *Let $\hat{\mathcal{P}}(z)$ denote the set of prices solving the first order condition,*

$$\frac{\partial \pi(p, z)}{\partial p} \frac{p}{\Pi(p, z)} = - \underbrace{\frac{p}{\Delta(p, z)} \frac{\partial \Delta(p, z)}{\partial p}}_{\equiv \varepsilon_m(p, z)}, \quad (9)$$

for each z . If $g(\psi)$ is continuous at $\psi = 0$ (i.e. $\zeta > 1$), the first order condition is necessary for an optimum and we have $\hat{\mathcal{P}}(z) = p^*(z)$ if $p^*(z) \leq \bar{p}(z)$, and $\hat{\mathcal{P}}(z) \subset [\bar{p}(z), p^*(z))$ otherwise. Moreover, if $\varepsilon_m(p, z)$ is non-decreasing in p , for all $p \leq p^*(z)$, the first order condition is also sufficient for an optimum, and $\hat{\mathcal{P}}(z) \equiv \hat{p}(z)$ is a singleton for each $z \in [\underline{z}, \bar{z}]$.

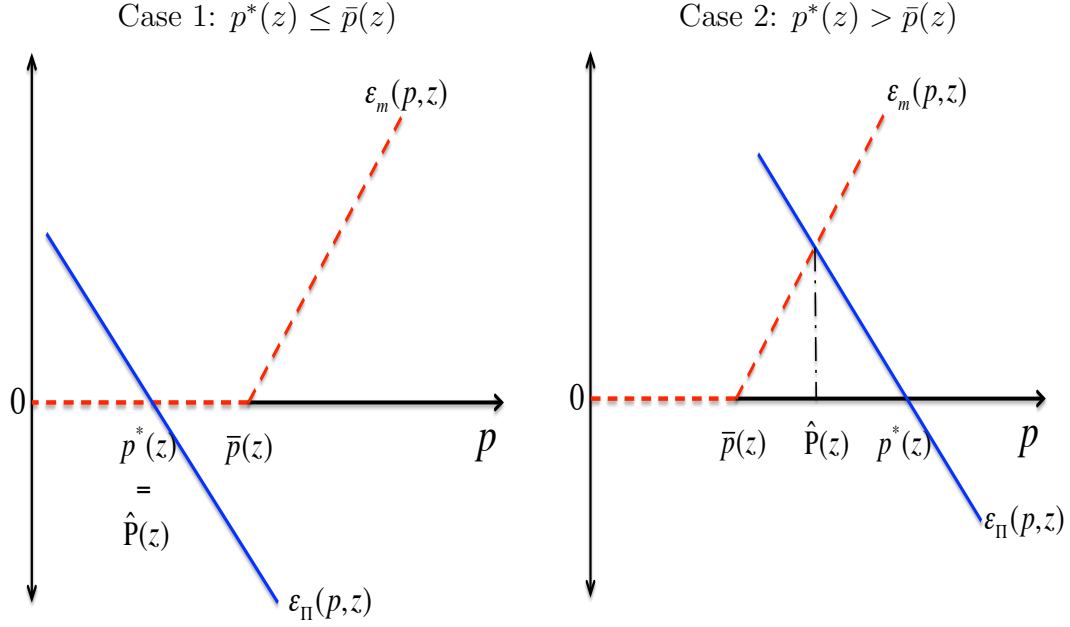
A proof of the proposition can be found in [Appendix A.4](#), and a graphical representation of the optimal price schedule, under the required assumptions for the first order condition to be necessary and sufficient, is presented in [Figure 1](#). The requirement that the density of search costs, i.e., $g(\psi)$, is continuous at zero follows because otherwise there can be a discontinuous change in the amount of customers exiting the customer base of a firm if it the prices changes marginally from below the maximum price at which no customer searches, i.e., $\bar{p}(z)$, to above this price. This would imply that the value function exhibits a kink at this price, therefore making the first order condition not necessary for the optimal price characterization. Under the continuity restriction, the proposition states that a firm's optimal price is never above the price that maximizes per-customer static profits, $p^*(z)$. It further argues that, only when productivity is high enough, which implies a high $\bar{p}(z)$, the optimal price is that one that maximizes static profits; this is depicted in the left panel of [Figure 1](#), where $p^*(z) \leq \bar{p}(z)$. The intuition behind this result is that customers do not find it optimal to exit highly productive firms and, as a result, these firms are not constrained by the extensive margin of demand and therefore price as in a static setup. Whenever $p^*(z) > \bar{p}(z)$, which occurs when productivity is low, the optimal price is below that one that maximizes static profits. This occurs because $\varepsilon_m(p, z)$, which represents the *extensive* margin demand elasticity, is positive so that firms are constrained by the extensive margin of demand. If, on top of the continuity restriction, we have that $\varepsilon_m(p, z)$ is increasing in p , we conclude that when the optimal price is unique. This situation, for the case where $p^*(z) \leq \bar{p}(z)$, is depicted in the right panel of [Figure 1](#). The plot shows that, given that the extensive margin elasticity is positive at the intersection, the optimal price lies in the open set $(\bar{p}(z), p^*(z))$.

The next lemma characterizes $\varepsilon_m(p, z)$. In particular, it provides sufficient conditions under which it is increasing on p and decreasing in z .

Lemma 3 $\varepsilon_m(p, z) \geq 0$. If $g(\psi)/(1 - G(\psi))$ is strictly increasing in ψ for all $\psi \geq 0$ (i.e. $\zeta > 1$), there exists a value $\bar{\lambda}$, such that for any scale of search cost satisfying $\lambda > \bar{\lambda}$ we have that $\varepsilon_m(p, z)$ is increasing in p for all $p \leq p^*(z)$ and decreasing in z .

A proof of the lemma can be found in [Appendix A.5](#). [Lemma 3](#) ensures that the firm's maximization problem has a unique solution, characterized by the first order condition, and that the optimal price is a differentiable function of productivity (see [Proposition 2](#)). In

Figure 1: Optimal price



Note: Both plots assume that $g(\psi)$ is continuous at zero and that $\varepsilon_m(p, z)$ is non-decreasing in p .

Section 4 we will estimate the model and show that estimated parameters do support $\varepsilon_m(p, z)$ increasing in p and decreasing in z .

A salient feature of our model is that productivity z has two distinct effects on prices. First, productivity has the standard *direct effect* of productivity on prices: higher productivity induces lower prices as the marginal benefit on per-customer profits of a price increase is decreasing in both p and z . Second, productivity has also an *indirect effect* on prices. This follows as current productivity determines the probability distribution of future prices in that firm. Thus productivity in our model affect both costs and demand structure of a firm. As a result, when the extensive margin elasticity $\varepsilon_m(p, z)$ is decreasing in z , $\hat{P}(z)$ is not necessarily a monotonic function of current productivity as more productive firms might face a more inelastic demand function than less productive firms. By this argument, everything else being equal, more productive firms would like to charge a higher price. Thus, depending on which effects prevails, the optimal pricing policy might not be monotonic in productivity.

The variable $\varepsilon_m(p, z)$ describes the *extensive* margin elasticity of demand, and plays a crucial role in the optimal price determination. Using the demand elasticities definitions into

equation (9), we obtain that the optimal price $\hat{\mathcal{P}}(z)$ is the solution to the following expression,

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

which highlights the importance of the different margins of demand elasticity for optimal markups, $\mu(p, z)$. The equation shows that the optimal markup is affected by (i) the intensive margin demand elasticity $\varepsilon_d(p)$, and (ii) the extensive margin elasticity $\varepsilon_m(p, z)$ weighted by the relative importance of the value of retaining the marginal customer with respect to a measure of its static value, $\Pi(p, z)/(d(p)p)$. As a result, this expression shows that the optimal price trades-off short term profits vs. long term value accrued from the marginal customer: losing a customer in the current period means not only losing revenues for that period, but the loss in demand is persistent as customers are hard to win back. Therefore, the optimal markup solves a dynamic problem. This type of trade-off is not present in standard macro models where markups are governed by the intensive margin elasticity, as there a change in price only affects current demand but does not have consequences for future demand. The next lemma presents a set of comparative statics.

Lemma 4 *For a given productivity level z , the optimal markup $\mu(\hat{\mathcal{P}}(z), z)$ is (i) decreasing in the level of the intensive margin elasticity $\varepsilon_d(\cdot)$, (ii) decreasing in the level of the extensive margin elasticity $\varepsilon_m(\cdot)$, (iii) decreasing in $\Pi(p, z)/(d(p)p)$. Finally, (iv) fixing a given total demand elasticity $\varepsilon_q(p, z) \equiv \varepsilon_m(p, z) + \varepsilon_d(p)$, the higher $\varepsilon_m(p, z)$ the lower the optimal markup.*

The proof of the lemma is straightforward and therefore omitted. Part (i) and (ii) of the lemma simply states that the higher either the intensive or extensive margin elasticities, the lower the price charged by a firm. Notice that when customer base concerns are not binding, the optimal price depends only on the intra-temporal elasticity of demand and the markup coincides with that one chosen by a monopolist facing elasticity $\varepsilon_d(p)$. When customer base concerns are present, i.e., when $\varepsilon_m(p, z) > 0$, the optimal markup is below that one following from static profit maximization; this happens because firms value retaining customers, and customers are hard to retain. Part (iii) of the lemma states that the optimal price decreases with the relative importance of the long term value of a customer relative to the short term value. Part (iv) of the lemma states that the larger the extensive margin elasticity share of the aggregate demand elasticity, the lower the optimal markup. As a result, if we compare two firms with the same total demand elasticity $\varepsilon_q(p, z)$, they might charge very different prices depending on the combination of intensive and extensive margin demand elasticities.

2.4 Equilibrium

In this section we define a stationary equilibrium and we discuss some of its general properties. We next describe the equilibrium.

Definition 1 *A stationary equilibrium is (i) two functions for the threshold rules $\hat{\psi}(\cdot)$ and $\hat{z}(\cdot)$ that solve the customer problem in [equations \(1\) and \(2\)](#) given firms' optimal pricing, (ii) a pricing policy $\mathcal{P}(\cdot) \equiv \hat{p}(\cdot)$ that solves firm's optimal pricing problem in [equation \(8\)](#) given customers' optimal threshold rules, and (iii) invariant distributions $h(\cdot)$ and $q(\cdot)$ that solve*

$$h(z') = \int_{\underline{z}}^{\bar{z}} \Delta(\hat{p}(x), x) f(z'|x) h(x) dx, \quad (11)$$

$$q(z) = G(\hat{\psi}(\hat{p}(z), z)) h(z), \quad (12)$$

for each $z \in [\underline{z}, \bar{z}]$ with boundary condition $\int_{\underline{z}}^{\bar{z}} h(z) dz = \Gamma$.

We now start discussing equilibrium characterization. Notice that an implication of [Lemma 3](#) is that, given that $\hat{p}(z)$ is continuous in z , also do $h(z)$ and $q(z)$ (see [equations \(11\) and \(12\)](#)); this verifies the initial conjecture on these measures.

Proposition 3 *If the assumptions of [Lemma 3](#) are satisfied, there exists an equilibrium where $\hat{V}(z)$ is increasing in z with the following properties:*

- (i) *the equilibrium exhibits price dispersion, that is, there exists two productivity levels z_1 and z_2 belonging to $[\underline{z}, \bar{z}]$ for which $\hat{p}(z_1) \neq \hat{p}(z_2)$;*
- (ii) *$\hat{z}(\hat{p}(z), z) = z$ for each $z \in [\underline{z}, \bar{z}]$;*
- (iii) *$\hat{p}(\bar{z}) = p^*(\bar{z})$, while $\hat{p}(z) < p^*(z)$ for all $z \in [\underline{z}, \bar{z})$;*
- (iv) *$\hat{\psi}(\hat{p}(\bar{z}), \bar{z}) = 0$, while $\hat{\psi}(\hat{p}(z), z) > 0$ and decreasing in z for all $z \in [\underline{z}, \bar{z})$;*
- (v) *customer growth $\Delta(\hat{p}(z), z)$ is increasing in z for all $z \in [\underline{z}, \bar{z}]$ with $\Delta(\hat{p}(\underline{z}), \underline{z}) < 1$ and $\Delta(\hat{p}(\bar{z}), \bar{z}) > 1$;*
- (vi) *if $\varepsilon_d(\hat{p}(z))$ is not increasing in z , then the equilibrium markups $\mu(\hat{p}(z), z)$ increase with z , for any $z \in [\underline{z}, \bar{z}]$.*

A proof of the proposition can be found in [Appendix A.6](#). The existence of an equilibrium, with $\hat{V}(z)$ increasing in z , follows by applying Brouwer fixed point theorem. Part (i) states

that the equilibrium exhibits price dispersion which, ulteriorly, motivate customers to search for lower prices. Parts (ii) to (iv) characterizes the optimal policy of customers and firms. In equilibrium only firms with the highest productivity, i.e., those firms with productivity $\bar{z}$, do not face an active extensive margin elasticity (which can be expressed analogously as $\hat{\psi}(\hat{p}(\bar{z}), \bar{z}) = 0$) and therefore can charge the price that maximizes static profits $p^*(\bar{z})$; this is an immediate consequence of the value for customers $\hat{V}(z)$ being increasing in firm's productivity z . Likewise, this implies that, for any productivity level $z < \bar{z}$, firms do experience a positive extensive margin elasticity (could also be expressed as $\hat{\psi}(\hat{p}(z), z) > 0$ for all $z < \bar{z}$), which distorts the optimal price so that $\hat{p}(z) < p^*(z)$ for all $z < \bar{z}$. The same argument explains part (v) of the proposition. That is, it explains why the growth rate of the customer base of a firm increases with its current productivity level, and explains why firms with low productivity experience a shrinkage of their customer base, while those with high productivity enjoy an expansion. Part (vi) states that when, in equilibrium, the intensive margin elasticity does not increase with z for any z , then the optimal markup increases with the firm's productivity. Sufficient conditions for this requirement are either that $\varepsilon_d(p)$ is constant in p (as it is the case in models of monopolistic competition with infinitely many varieties), or that the optimal price $\hat{P}(z)$ monotonically decreases as the productivity level increases. Under the restrictions imposed in the proposition, the monotonicity of markups follows from the fact that, in equilibrium, the extensive margin elasticity $\varepsilon_m(p, z)$ is decreasing in z , and that the relative value of a customer is also a decreasing function of productivity as customer demand increases with productivity but the latter is mean reverting. Notice that the proposition is silent about the monotonicity of prices, something that we sketched in previous sections. Non-monotonicities can arise because more productive firms face a lower extensive margin elasticity and therefore, as described in [Lemma 4](#), might find it optimal to charge higher prices than less productive firms.³

In the next two Remarks we characterize two useful limiting cases. In the first Remark we analyze the equilibrium where customer base concerns are no present, which happens when $\lambda \rightarrow \infty$. Because of no customer base concerns, the model in this case reduces to the standard price setting problem under monopolistic competition. In the second Remark we discuss the equilibrium where there is no productivity dispersion.

Remark 1 *Suppose that the scale of search cost is divergent, i.e. $\lambda \rightarrow \infty$. Then, in equilibrium: (i) the optimal price maximizes static profits, $\hat{p}(z) = p^*(z)$ for all z , and (ii) customers do not find it optimal to exit the customer base of their current firm. The equilibrium is unique.*

³Nevertheless, an increasing price schedule for all productivity levels is not possible as this would violate the equilibrium condition that $\hat{V}(z)$ increases with z .

The proof of the Remark follows immediately since, because the search cost diverges, firms do not face customer base concerns and therefore find it optimal to charge the price that maximizes static profits. Formally, when $\lambda \rightarrow \infty$, we have that $\bar{p}(z) \rightarrow \infty$, so that $p^*(z) \leq \bar{p}(z)$ for all z . Then, $\hat{p}(z) = p^*(z)$.

Remark 2 *Let $\underline{z} = \bar{z} = z_0$, $f(z'|z)$ degenerate at $z = z_0$, and the scale of search cost arbitrarily larger than zero, i.e. $\lambda > 0$. Then, there is a unique equilibrium where each firm charges the price that maximizes static monopoly profits, i.e. $\hat{p}(z) = p^*(z)$.*

A proof of the Remark can be found in [Appendix A.7](#). The Remark shows how our model relates to [Diamond \(1971\)](#). It shows that when firms are homogeneous in their productivity level they all charge the price that maximizes static profits. Moreover, because every firm charges the same price, customers have no incentives to search so that the customer base is constant.

Remark 3 *Let the assumptions of [Lemma 3](#) be satisfied and the cumulative distribution function of future productivity z' conditional on current productivity z be given by*

$$F(z'|z) = \rho \mathbf{1}_{z' \leq z} + (1 - \rho) \hat{F}(z') , \quad (13)$$

for some $\rho \in [0, 1)$, where $\hat{F}(z')$ is a CDF with support $[\underline{z}, \bar{z}]$. Then equilibrium prices are such that $\hat{p}(z)$ decreases monotonically with productivity z .

The result originates from the fact that, given the process in [equation \(13\)](#), the continuation value $\hat{V}(z)$ can be increasing in z if and only if $\hat{p}(z)$ is decreasing in z , as the continuation value of a customer depends on current productivity z only through the extent that the firm draws the same productivity with probability ρ , otherwise the continuation value is independent of current productivity z .

3 Empirical evidence

We complement the theoretical analysis with an empirical investigation that relies on cashier register data from a large US supermarket chain on grocery purchases of households holding the chain loyalty card. The empirical analysis has two purposes: in this section we exploit the data to document that fluctuations in the price posted by a firm influence a customer decision to remain in its customer base and measure the size of this effect. Next, we use the data to estimate the model and quantify the importance of customer base concerns in shaping firm price setting.

3.1 Data and variable construction

For every trip made at the supermarket chain by a panel of households between June 2004 and June 2006, we have information on date of the trip, store visited and list of goods (identified by their Universal Product Code, UPC) purchased, as well as quantity and price paid. The data are ideal to study customer base dynamics since the loyalty card allows to track precisely when a particular household is shopping at the supermarket chain. Moreover, the fact that all the households in our sample own a loyalty card of the chain implies that they can be thought of as regular customers. This makes it more plausible that they would face some cost of switching to a different retailer even in a setting, like the supermarket industry, where there are no formal contracts binding the customer-retailer relationship.

The micro data on household purchases are complemented by data on store sales for the same time period. We use information on weekly revenues and quantities sold for a sample of stores representative of the different price areas to compute prices for each UPC in the sample. The construction of the price variable is therefore analogous to that in [Eichenbaum et al. \(2011\)](#) and is subject to the same caveats. The price variable varies weekly matching the timing of price adjustment for the retailer. For the analysis we retain only UPC's for which we have complete time series of prices for the 104 weeks in the sample span.

For the purpose of our exercises we have to construct two key variables: (i) an indicator signaling when the household is leaving 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 B](#).

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 left the customer base when she has not shopped at the chain for eight or more consecutive weeks and we date the exit event to the last time the customer visited a store of the chain. The 8-weeks window is a conservative choice since households in our sample hold a loyalty card of the chain, suggesting that they are not casual customers. To make sure of this, we retain in the sample only households shopping at the chain at least 48 times over the two years in the sample. Regular customers are unlikely to experience a 8-weeks spell without shopping for reasons other than having switched to another chain (e.g. consuming their inventory). Indeed, customers in our sample make an average of 157 shopping trips at 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 number of days elapsed between consecutive trips is close to four and the 99th percentile is 24 days, roughly half the length of the absence we require before inferring that a household is buying its grocery at a competing chain.

We construct the price of the basket of grocery goods usually purchased by the households

in a fashion similar to [Dubois and Jodar Rosell \(2010\)](#). We identify the goods belonging to a household’s basket using scanner data on items the household purchased over the two years in the sample. The price of its basket in a particular week is then computed as the average of the weekly prices of the goods included in the basket, weighted by their expenditure share in the household budget. Since households differ in their choice of grocery products and in the weight such goods have in their budget, the price of the basket is household specific.

3.2 Price and customer base dynamics

We estimate a linear probability model where the dependent variable is an indicator for whether the household has left the customer base of the retail chain in a particular week. Our regressor of interest is the logarithm of the price of the basket of grocery goods usually purchased by the households at the chain ($p^{retailer}$).

In [Table 1](#), we report results of regressions of the following form,

$$Exit_{it} = b_0 + b_1 p_{it}^{retailer} + X_i' b_2 + \varepsilon_{it} \quad (14)$$

In the regression we include year-week fixed effects to account for time-varying drivers of the decision of exiting the customer base common across households and we control for observable characteristics, such as age, income, and education, through inclusion of household’s demographics matched from Census 2000. We add the number of competing grocery retailers in the zipcode, as well as the distance (in miles) from the closest store of the chain and that from the closest store of the competition to account for the fact that households living closer to outlets of the chain and far away from alternative options will be less likely to leave the customer base of the chain. Finally, we include as regressors the logarithm of the price of the basket in the first week in the sample and the standard deviation of price changes for each household over the sample period. These are meant to control for differences in the composition of the basket across shoppers. For example, some customers may purchase product categories more prone to promotions than others and experience more intense price fluctuations as a result.

Our approach differs from the standard discrete choice studies of demand. We are not modeling the household’s choice of the preferred retailer among a set of potential alternatives, but rather the decision of whether to leave a given retailer or to stick with it. In other words, our results only speak to how relevant the price of the basket is in driving separations between previously formed matches between customers and retailers. They convey no information on whether and how much this variable matters for the formation of such matches. This distinction is key because it implies that the usual concern of price endogeneity driven by

unobserved store characteristics is not relevant in this context. However, price could still be endogenous if: i) the retailer knows about household idiosyncratic shocks -unobservable to us- that affect the decision of changing retailer; ii) the retailer can use this knowledge to adjust accordingly the price faced by the household. Whereas it is conceivable that the supermarket observes variables predicting the exit from the customer base of groups of households, it is unlikely that it can react with targeted prices. In fact the basket of different households will at least partially overlap making it impossible to fine tune the basket price faced by some households without affecting the price of others.

A final important difference with standard demand estimation is in the interpretation of the price coefficient. According to our model, the price plays a dual role: it represents the cost of purchasing the basket at the retailer today but, as stressed by its dependence on productivity, it also signals the convenience of purchasing from the retailer in the future. Therefore, the coefficient cannot be used to predict the customer reaction to a random -i.e. unrelated to productivity shocks- change in prices.

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

	Exiting: Missing at least 8 consecutive weeks	
	(1)	(2)
$\log(P_i^{retailer})$	0.015** (0.007)	0.015* (0.008)
$\log(P_i^{competitors})$		0.001 (0.001)
Observations	71,049	52,670

Notes: An observation is a household-week pair. The sample only includes households who prominently shop at stores for which we have complete price data for all the UPCs they purchase. We exclude from the sample the top and bottom 1 % in the distribution of the number of trips over the two years. Demographic controls rely on a subsample of households for which information on the block-group of residence was provided and include as regressors ethnicity, family status, age, income, education, and time spent commuting (all matched from Census 2000) 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. ***: Significant at 1% **: Significant at 5% *: Significant at 10%.

Table 1 reports the results of a regression whose dependent variable is an indicator that takes value one if in that week the customer decides to leave the retailer, and zero otherwise. Column (1), documents that a 1% increase in the weekly price of the consumer-specific grocery basket is associated with 0.015 percentage points increase in the probability that the customer leaves the chain to patronize a rival firm. The coefficient on the price of the

basket is identified by good-retailer specific shocks as those triggered, for example, by the expiration of a contract between the chain and a manufacturer. Since we use household level data to construct a consumer specific basket, even chain-wide shifts in the cost of all goods generate useful variation in price because different customers give different weight to the various goods in their baskets. Furthermore, we observe in our data a fair amount of good-store specific shocks: within the chain, the price of a same good moves differently in different stores. This can be due, for instance, to variation in the cost of supplying the store due to logistics (e.g. distance from the warehouse) which will hit differently goods with different intensity in delivery cost (e.g. refrigerated vs. non refrigerated goods).

The results in column (1) rely on the variation in the cost of the household grocery basket at the chain and do not control for the pricing behavior of the competitors. This raises natural concerns on the precision of our estimate of the elasticity. If we fear that a large fraction of the variation in $p_{it}^{retailer}$ came from aggregate cost shocks affecting all the retailers (e.g. a spike in the price of aluminium that makes soda cans more expensive for every retailer), failing to control for the general level of prices would lead to underestimate b_1 . Only shock idiosyncratic to the chain should be expected to affect the probability of leaving the retailer. Aggregate cost shocks do not change the relative price and, therefore, should not trigger exit from the customer base. Furthermore, our retailer is a major player in the markets included in our sample and it is reasonable to assume that the competition takes its prices into consideration when deciding their own. This introduces correlation between price variations at the chain and price variations at the alternative outlets the customer may visit. Disregarding the prices of the competitors may lead to biases in the magnitude and even the sign of the own-price elasticity.

In column (2) we address both of these concerns by directly controlling for the prices posted by competitors of the chain using the IRI Marketing data set. This source includes weekly prices UPC's prices for 30 major product categories for a representative sample of chain stores across 64 markets in the US.⁴ Using this data, we can compute the price of each UPC in the Metropolitan Statistical Area of residence of a customer by averaging the price posted for the item by all the chains sampled by IRI. Then, we construct the average market price of the basket bought by the consumer in the same fashion described for the price of the basket at the retailer.⁵ Even after controlling for the general level of prices, the coefficient on the price of the basket at the chain stays negative and significant and nearly unaffected

⁴A 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.

⁵We define this variable average market price of the basket, rather than price of the basket at competitors because it includes the price posted by our chain as well. In fact, chain identity is masked in the IRI data, preventing us from excluding the prices of our retailer from the average.

in magnitude, suggesting that most of the variation in the index comes from chain, or even good-chain, specific shocks.

4 Calibration

In this section we discuss the procedure we follow to calibrate the model.

First, we need to choose a functional form for the utility of customers. We assume that their utility function is $u(c) = c^{1-\gamma}/(1-\gamma)$ and we interpret c as a composite of two types of goods: one (that we label d) is the product produced by the type of firms described in our model; the other (n) is produced by a competitive representative firm with linear production function and unitary labor productivity. The composite good is then defined as $c = \left(\omega d^{\frac{\theta-1}{\theta}} + (1-\omega) n^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}$, $\theta > 1$. We set n as the numeraire good so that the budget constraint faced by the agents is $pd + n = I$, where I is the agent's income.⁶ In [Appendix A.1](#) we show that moving from these assumptions we can derive a demand function ($d(p)$) and a consumer surplus function ($v(p)$) consistent with the assumptions made in [Section 2](#)

We assume that productivity follows an AR(1) process of the following form: $\log(z_t^j) = \rho \log(z_{t-1}^j) + \sigma \varepsilon_t^j$; where ε is distributed according to a standardized normal. For consistency with the theoretical model, where productivity is assumed to follow a Markov process, we then use the procedure described in [Tauchen \(1986\)](#) to construct the Markov chain that approximate the AR(1) process we estimated for productivity.

We assume that a period in the model corresponds to a week. It is traditionally hard to identify the discount factor from data; therefore we have to fix its value. We assume the firm discount rate to be $\beta = \hat{\beta} \cdot (1 - \delta)$, where $\hat{\beta} = 0.999$ and δ is a rescaling factor representing the “death probability” of the firm decision maker. We choose $\delta = 0.005$ to match the average tenure of CEOs in the retail food industry reported in [Henderson et al. \(2006\)](#). Rescaling the discount factor with the death probability helps ensuring that the effective discount factor faced by the firms stays below one, which is necessary the operator $W(z)$ in [equation \(8\)](#) to be a contraction. For simplicity, we set the customers' discount rate at the same level as firms'. The relative risk aversion parameter needs also to be fix; we choose a value of 2, which is standard in the literature.

Finally, the data lack the information to identify some of the parameters of the utility function. We assume that goods d and n have the same weight in the consumer's preferences

⁶We assume that agents derive their income from working and from their ownership of shares of the firms, which entitles them to part of their profits. In [Appendix A.9](#) we solve for the equilibrium labor supply of the agents. The assumptions of perfect competition and linear technology in the production of the competitive good n pin down the wage in the economy.

($\omega = 0.5$). The parameter θ is not directly fixed but it is selected so that the implied average intensive margin elasticity is 5, a value in the range of those used in the macro literature.⁷

All the remaining parameters are chosen by matching moments from our data with the analogue computed from the numerical solution of the model. Below we explain how the choice of the moments contributes to identifying particular parameters. This argument 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.

We assume that the search cost is drawn from a Gamma distribution with shape parameter ζ , and scale parameter λ . To identify the parameters of the search cost distribution we exploit the estimates of the relationship between price and probability of exiting the customer base discussed in [Section 3](#). The marginal effect of price on exit probability is informative about the level of the search costs. Thus we identify the scale parameter λ by matching the average effect of p on the exit probability

$$\int_z \frac{\partial(G(\hat{\psi}(p, z))(1 - H(\hat{z}(p, z))))}{\partial p} dz$$

to the parameter b_1 in [equation \(14\)](#). In particular, we choose the specification in column (2) of [Table 1](#), which controls for the prices posted by the competing chains. The shape parameter ζ is proportional to the coefficient of variation of the search cost distribution. In the model, higher dispersion of search costs implies higher cross-firms variation in the extensive margin elasticity. In the data, we measure this variation by fitting a spline to [equation \(14\)](#), allowing the price marginal effect on the probability of exit to vary for different terciles of price levels. Then ζ is identified by matching this estimated to the dispersion generated by the model. More specifically, if $z_{(1)}$ is the first and $z_{(2)}$ the second tercile, we use the following model generated moments

$$E \left[\frac{\partial(G(\hat{\psi}(p, z))(1 - H(\hat{z}(p, z))))}{\partial p} \mid z \geq z_{(2)} \right] - E \left[\frac{\partial(G(\hat{\psi}(p, z))(1 - H(\hat{z}(p, z))))}{\partial p} \mid z \leq z_{(1)} \right]$$

The persistence, ρ and volatility σ of the productivity shock are chosen so to match those of the resulting posted prices matches, measured using the store-level prices provided by our chain.⁸

⁷See [Nakamura and Steinsson \(2010\)](#) for a discussion.

⁸These statistics are obtained fitting an AR(1) process to the time series of prices separately for each of the 126 stores for which we have store level price data. This step delivers 126 estimates of the persistence parameters and 126 estimates of the volatility of the residuals. We match the median among those values for each statistic.

The minimum distance estimator is constructed defining $\Omega \equiv [\zeta \ \lambda \ \rho \ \sigma]'$ as the vector of parameters to be estimated, and denoting by $v(\Omega)$ the vector of the theoretical moments evaluated at Ω , and by v_d their empirical counterparts. We search for Ω that minimizes the quadratic form $(v_d - v(\Omega))' (v_d - v(\Omega))$. [Table 2](#) reports the results.

Table 2: Parameter estimates

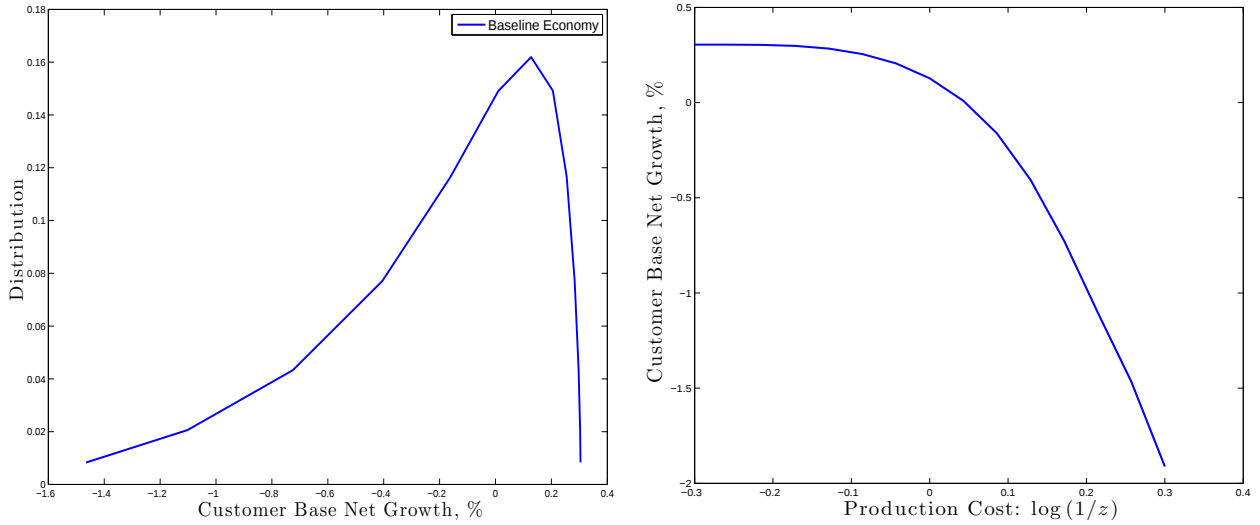
	Value	Target
Persistence of productivity innovations, ρ	0.195	Price process persistence: 0.31
Volatility of productivity innovations, σ	6.72	Price process dispersion: 0.2
Distribution of cost, $g(\psi) \sim \text{Gamma}(\zeta, \lambda)$		
Shape parameter, ζ	3	Inter-tercile difference in marginal effect: 1.8%
Scale parameter, λ	0.17	Average marginal effect: 1.4%

5 Growth, markups and idiosyncratic cost

We use the estimates obtained in the previous section to understand the behavior of the model. We begin by looking at the evolution of the customer base, that is the mass of customers purchasing from a particular firm. The left panel of [Figure 2](#) displays the invariant distribution of the net growth rate of the customer base of the firms in our economy. The picture that emerges is one of high activity on the extensive margin. The distribution of growth rates is fairly dispersed with a fraction of firms increasing the number of their customers up to half a percentage point and a long tail of firms that are losing customers. The right panel relates customer base dynamics to cost, computed as the inverse of productivity. Firms with cost below average experience positive net growth of their customers base; the growth rate does not decline steeply when the cost raises. High cost firms see their customer base shrink as their customer are more likely to search and to find a better match when they do so. In this case, the rate of contraction of the customer base grows fast with cost; small differences in efficiency translate into significantly more negative growth rates.

We next shift our focus to the price setting behavior and analyze the distribution of markups in our economy. As a benchmark, we compare those resulting from our model with customer base stickiness (henceforth, “Baseline economy”) with the ones obtained from a model where we shut down that channel. We do so by letting search costs diverge to infinity

Figure 2: Equilibrium growth of firms



(i.e., $\lambda \rightarrow \infty$) so that customers would never want to search for a new firm. It is important to stress that we make sure to fix θ so that the resulting average total elasticity of demand (i.e., $\int_{\bar{z}} \varepsilon_q(\hat{P}(z), z) f(z) dz$) is the same as in our Baseline economy. Since this alternative model is analogous to the standard CES preferences widely used in the macro literature, we will refer to it as “CES economy”.

The left panel of [Figure 3](#) compares the equilibrium markups in the Baseline and in the CES economy. Firms with the lowest production cost offer better value to their customers and face lower risk of losing them. They are therefore unconstrained by the presence of customer markets and their markups coincide with those of the CES economy. Firms who are not at the top of the productive efficiency scale face an actual risk of losing customers, giving them incentive to keep the price low at the expenses of the markup. In the frictionless CES economy losing customers is less concerning and this results in higher markups being charged by the firms.

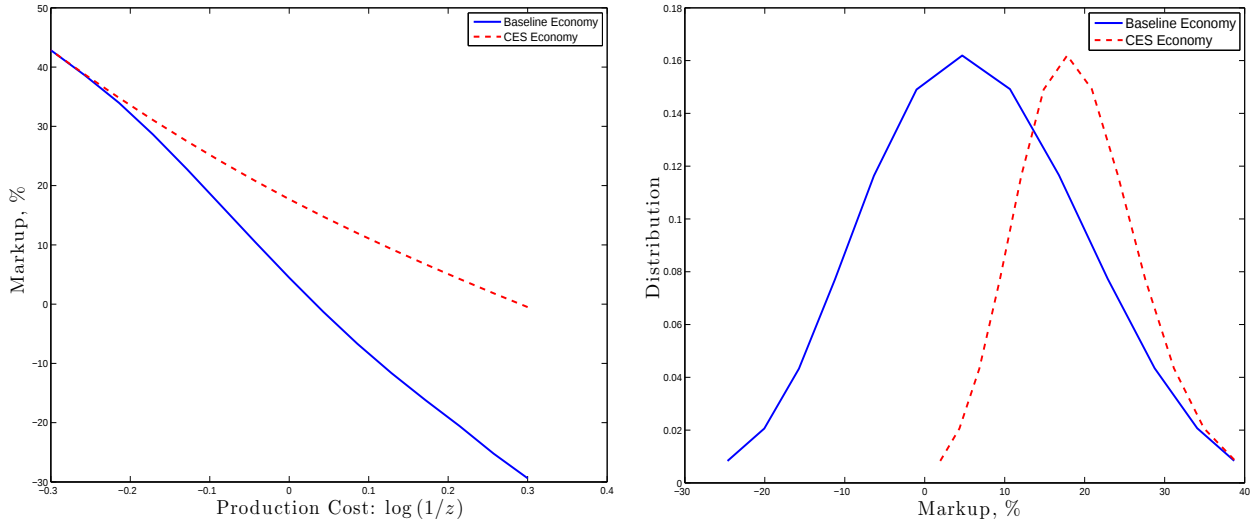
The figure also suggests that the optimal markup is strictly decreasing in production cost in both economies but it decreases faster in the Baseline economy than in the CES economy. In fact, in both models the intensive elasticity of demand, $\varepsilon_d(p)$, is increasing in p , giving rise to a negative comovement between markups and production cost.⁹ Given that equilibrium

⁹With CES preferences the demand of good i depends on the relative price p_i/P . With a finite number of goods in the basket of the consumer, an increase in p_i , also increases the price of the basket, P , thus reducing the overall increase in p_i/P and effect on demand. The effect on P is larger, the higher the weight of good i in the basket, that is the lower the price p_i and the higher its demand. Therefore, the elasticity of demand

prices are monotonically increasing in production cost, it follows that firms with higher production cost face higher elasticity of demand, so that optimal markups are decreasing in production cost. However, in the baseline economy the extensive margin elasticity also plays a role. Increases in price lead to the loss of customers, taming the incentive to adjust prices after cost spikes. Moreover, a persistent increase in firm specific cost of production is associated to higher expected future prices, and thus a worsened ability to retain and attract customers. This means that our model delivers markups that are procyclical with respect to productivity shocks, consistent with the empirical evidence provided by [Petrin and Warzynski \(2012\)](#), and more strongly so than in a CES economy.

The right panel of [Figure 3](#) documents the implications for the distribution of markups. The distribution of markups in our model displays substantial negative skewness, with almost 42% of firms charging a negative markup, while no firms charge a negative markup in the CES economy. The average markup in the Baseline economy is 5%, substantially lower than the 18% average markup predicted by the CES economy. Finally, the fact that the function mapping production cost to markups is steeper in the Baseline economy implies an increase in the markups dispersion from 7% to 14%.

Figure 3: Equilibrium Markups



$\varepsilon_d(p)$ increases in p .

6 Aggregate shocks and markup cyclicalities

TBA

7 Conclusions

The customer base is an important determinant of firm performance. Introducing customer retention consideration into standard models can improve our understanding of firm pricing behavior. We setup and estimate a model where firms face sticky demand and use it to explore the implications of this feature for the cyclicalities of markups and cost-passthrough.

In our setting, there are two margins of demand adjustment. Customers can adjust both the quantity they purchase in response to price fluctuations and decide to leave and shop elsewhere. Because of search frictions, lost customers are hard to gain back and firms have an incentive to care about retaining their customers. We characterize the equilibrium of the model and show that customer retention concerns introduce a dynamic element into optimal markup so that markups depend on firm productivity.

We use scanner data on households' purchases at a U.S. supermarket chain to provide direct evidence that customers do respond to variation in the price of their consumption basket. We also exploit the data to estimate the key parameters of the model and provide a quantification of the effect of customer retention concerns on firm pricing.

We show that our model implies pro-cyclicalities of markups to idiosyncratic shocks, which is more accentuated than in a standard CES model. We also explore the dynamics of markups in response to aggregate shocks in a setting where customer markets play a role.

References

- Aguiar, M. and Hurst, E. (2007). Life-cycle prices and production, *American Economic Review* **97**(5): 1533–1559.
- Alessandria, G. (2009). Consumer search, price dispersion, and international relative price fluctuations, *International Economic Review* **50**(3): 803–829.
- Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices, *American Economic Review* **98**(5): 1998–2031.
- Berman, N., Martin, P. and Mayer, T. (2012). How do different exporters react to exchange rate changes?, *The Quarterly Journal of Economics* **127**(1): 437–492.

- Bils, M. (1989). Pricing in a customer market, *Quarterly Journal of Economics* **104**(4): 699–718.
- Blanchard, O. (2009). The state of macro, *Annual Review of Economics* **1**(1): 209–228.
- Bronnenberg, B. J., W., K. M. 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.
- Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment, *International Economic Review* **39**(2): 257–73.
- Coibion, O., Gorodnichenko, Y. and Hong, G. H. (2012). The cyclicalty of sales, regular and effective prices: Business cycle and policy implications, *IMF Working Papers 12/207*, International Monetary Fund.
- Coles, M. G. (2001). Equilibrium wage dispersion, firm size and growth, *Review of Economic Dynamics* **4**(1): 159–187.
- Coles, M. G. and Mortensen, D. T. (2013). Equilibrium labor turnover, firm growth and unemployment, *Working paper*.
- Diamond, P. A. (1971). A model of price adjustment, *Journal of Economic Theory* **3**(2): 156–168.
- Drozd, L. A. and Nosal, J. B. (2012). Understanding international prices: Customers as capital, *American Economic Review* **102**(1): 364–95.
- Dubois, P. and Jodar Rosell, S. (2010). Price and brand competition between differentiated retailers: A structural econometric model, *CEPR Discussion Papers 7847*, C.E.P.R. Discussion Papers.
- 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.
- Foster, L., Haltiwanger, J. and Syverson, C. (2012). The slow growth of new plants: Learning about demand?, *Working Papers 12-06*, Center for Economic Studies, U.S. Census Bureau.

- Gourio, F. and Rudanko, L. (2011). Customer capital, *NBER Working Papers 17191*, National Bureau of Economic Research, Inc.
- Henderson, A. D., Miller, D. and Hambrick, D. C. (2006). How quickly do CEOs become obsolete? Industry dynamism, CEO tenure, and company performance, *Strategic Management Journal* **27**: 447–460.
- Kaplan, G. and Menzio, G. (2013). Five facts about price dispersion, *Working paper*, University of Pennsylvania.
- Kleshchelski, I. and Vincent, N. (2009). Market share and price rigidity, *Journal of Monetary Economics* **56**(3): 344–352.
- Luttmer, E. G. (2010). Models of growth and firm heterogeneity, *Annual Review of Economics* **2**(1): 547–576.
- Melitz, M. J. and Ottaviano, G. I. P. (2008). Market size, trade, and productivity, *Review of Economic Studies* **75**(1): 295–316.
- Menzio, G. (2007). A search theory of rigid prices, *PIER Working Paper Archive 07-031*, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.
- Nakamura, E. and Steinsson, J. (2010). Monetary non-neutrality in a multisector menu cost model, *The Quarterly Journal of Economics* **125**(3): 961–1013.
- Nakamura, E. and Steinsson, J. (2011). Price setting in forward-looking customer markets, *Journal of Monetary Economics* **58**(3): 220–233.
- Petrin, A. and Warzynski, F. (2012). The impact of research and development on quality, productivity and welfare, *mimeo*.
- Phelps, E. S. and Winter, S. G. (1970). Optimal price policy under atomistic competition, in G. Archibald, A. A. A. and E. S. Phelps (eds), *Microeconomic Foundations of Employment and Inflation Theory*, New York: Norton.
- Pistaferri, L. (2003). Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply, *Journal of Labor Economics* **21**(3): 729–754.
- Ravn, M., Schmitt-Grohe, S. and Uribe, M. (2006). Deep habits, *Review of Economic Studies* **73**(1): 195–218.

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

A Technical Appendix

A.1 Derivation of assumptions from first principles

In this section we propose a micro founded model that can give rise to the assumptions we made on the paper regarding customer's demand $d(p)$ and surplus $v(p)$.

Let customers derive utility from a large number of varieties $N > 1$ according to $u(C) = C^{1-\gamma}/(1-\gamma)$, where $C = \left(\sum_{n=1}^N c_n^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}$ is a CES consumption aggregator, where c_n denotes the consumption level of each variety, and $\theta > 1$. The way to connect this setup with the one used in the paper is to consider $d(p) = c_n(p_n)$ and $v(p) = u(C(p_n))$, where p_n denotes the price of variety n . That is, our model can be interpreted as evaluating the competition in the industry producing variety n , where firms differ in their productivity.

Each period the customer maximizes,

$$\max_{\{c_n\}_{n=1}^N} u(C) \text{ subject to } \sum_{n=1}^N p_n c_n = I ,$$

from where we obtain the standard first order condition for variety n ,

$$u'(C) \frac{\partial C}{\partial c_n} \frac{\theta-1}{\theta} c_n^{\frac{\theta-1}{\theta}-1} = \lambda p_n \text{ for all } n,$$

where λ denotes the Lagrange multiplier of the budget constraint. Operating with the first order conditions we obtain that

$$c_n = \frac{I}{P} \left(\frac{p_n}{P}\right)^{-\theta} , \quad P \equiv \left(\sum_{n=1}^N p_n^{1-\theta}\right)^{\frac{1}{1-\theta}} ,$$

where we have written the demand for variety n , c_n , as a function of its own price, p_n , and the price level, P . Notice that the price level is such that $P C = I$.

We start by discussing the properties of the demand function $d(p)$, which in this setup maps to evaluating the properties of $c_n(p_n)$. It is immediate to see that the demand for variety n converges to zero as its price diverges to infinity. That is, $\lim_{p_n \rightarrow \infty} c_n = 0$, which follows directly from the expression for c_n . We now show that when the number of varieties is large, the demand for variety n is decreasing and convex in its price. To this end, it proves

useful to compute the following derivatives,

$$\begin{aligned}\frac{\partial P}{\partial p_n} &= \left(\frac{P}{p_n}\right)^\theta \equiv A(p_n), \\ \frac{\partial c_n}{\partial p_n} &= \frac{c_n}{p_n} \left[-\theta + (\theta - 1) A(p_n)^{\frac{\theta-1}{\theta}}\right],\end{aligned}$$

where $A(p_n) > 0$ and $A(p_n) = N^\theta$ in a symmetric equilibrium. Notice that in the symmetric equilibrium, if N is large, we have that $\frac{\partial c_n}{\partial p_n} < 0$ and $\frac{\partial^2 c_n}{\partial p_n^2} > 0$, consistent with the demand function $d(p)$ being decreasing and convex in p . Moreover, because c_n and the price index P are twice continuously differentiable in prices and number of varieties N , the result also applies more generally away from the symmetric equilibrium.

Direct computations using the definition of c_n and price index P provide that the elasticity of demand of variety n is given by

$$\varepsilon_d(p_n) = -\frac{\partial \ln c_n}{\partial \ln p_n} = \theta - (\theta - 1) \frac{c_n p_n}{I},$$

where $c_n p_n = \left(\sum_{i=1}^N \left(\frac{p_i}{p_n}\right)^{1-\theta}\right)^{-1}$ which, in a symmetric equilibrium, is positive when $N\theta > \theta - 1$. For example, this condition is guaranteed to apply for large N . In particular, as N diverges to infinity we get that $\varepsilon_d(p_n) = \theta$, so that when there are infinite many varieties the demand elasticity is constant. Moreover, notice that

$$\frac{\partial \varepsilon_d(p_n)}{\partial p_n} = (\theta - 1)^2 \frac{1}{p_n} \left(\sum_{i=1}^N \left(\frac{p_i}{p_n}\right)^{1-\theta}\right)^{-1} \left[1 - \left(\sum_{i=1}^N \left(\frac{p_i}{p_n}\right)^{1-\theta}\right)^{-1}\right],$$

which in a symmetric equilibrium is equal to $(1/p_n)(\theta - 1)^2(1 - 1/N)/N > 0$.

Now we evaluate $v(p)$. In this micro funded setup this maps into exploring the effect of p_n on $u(C)$. Recall that, from the construction of the price index P , $P C = I$, so that

$$\begin{aligned}\frac{\partial C}{\partial p_n} &= -\frac{c_n}{P}, \\ \frac{\partial^2 C}{\partial p_n^2} &= -\left[\frac{\partial c_n}{\partial p_n} \frac{1}{P} - \frac{c_n}{P^2} A(p_n)\right],\end{aligned}$$

Then,

$$\begin{aligned}
\frac{\partial u(C)}{\partial p_n} &= C^{-\gamma} \frac{\partial C}{\partial p_n} < 0, \\
\frac{\partial^2 u(C)}{\partial p_n^2} &= -C^{-\gamma-1} \left(\frac{\partial C}{\partial p_n} \right)^2 \left[\gamma - \frac{C}{\frac{\partial C}{\partial p_n}} \frac{\partial^2 C}{\partial p_n^2} \right] \\
&= -C^{-\gamma-1} \left(\frac{\partial C}{\partial p_n} \right)^2 \left[\gamma + \theta \left(1 - A(p_n)^{\frac{1-\theta}{\theta}} \right) - 2 \right]
\end{aligned}$$

so that $\frac{\partial^2 u(C)}{\partial p_n^2} \leq 0$ if $\gamma + \theta \left(1 - A(p_n)^{\frac{1-\theta}{\theta}} \right) - 2 \geq 0$. For example, in the symmetric equilibrium, where $A(p_n) = N^\theta$, the required condition can be rewritten as $\gamma + \theta (1 - N^{1-\theta}) \geq 2$, which shows that a sufficient condition (in the symmetric equilibrium) is that $\gamma \geq 2$.

We now explore the existence of a unique solution that maximizes the profit function of the firm. This involves proving two different things. First, that there exists a unique solution to $\partial \pi(p, z)/\partial p = 0$. Second, that this solution is a maximum (i.e., that the profit function is strictly concave).

The first derivative of the profit function with respect to the price reads,

$$\frac{\partial \pi(p_n, z)}{\partial p_n} = c_n \left[1 - \varepsilon_d(p_n) \left(1 - \frac{w/z}{p_n} \right) \right],$$

where a solution to $\frac{\partial \pi(p_n, z)}{\partial p_n} = 0$ exists and it is unique if $\frac{p_n}{w/z} = \frac{\varepsilon_d(p_n)}{\varepsilon_d(p_n)-1}$ has a unique solution. Let $h_1(p_n) \equiv \frac{p_n}{w/z}$ and $h_2(p_n) \equiv \frac{\varepsilon_d(p_n)}{\varepsilon_d(p_n)-1}$. Notice that $h_1(p_n)$ is continuous, strictly increasing, with $h_1(0) = 0$ and $\lim_{p_n \rightarrow \infty} h_1(p_n) = \infty$. Also, because $\varepsilon_d(p_n)$ is continuous and increasing, $h_2(p_n)$ is continuous, decreasing, with $\lim_{p_n \rightarrow \infty} h_2(p_n) = \theta$. It follows that, for any number of varieties N , there exists a unique price solving $\frac{\partial \pi(p_n, z)}{\partial p_n} = 0$.

We now show that this unique price maximizes the firm's profits. To this end we show that in a symmetric equilibrium, for large N , the profit function evaluated at this price is concave. The second derivative of the profit function with respect to p_n reads,

$$\frac{\partial^2 \pi(p_n, z)}{\partial p_n^2} = -\frac{c_n}{p_n} \left[\varepsilon_d(p_n) \left(1 - \varepsilon_d(p_n) \left(1 - \frac{w/z}{p_n} \right) \right) + p_n \varepsilon'_d(p_n) \left(1 - \frac{w/z}{p_n} \right) + \varepsilon_d(p_n) \frac{w/z}{p_n} \right].$$

Notice that, in a symmetric equilibrium, c_n , p_n , $\varepsilon_d(p_n)$, and $\varepsilon'_d(p_n)$ are continuous in N . We will use this fact to prove that for large N the profit function is concave at the price maximizing static profits. Notice that when N diverges to infinity the second derivative reduces to

$$\left. \frac{\partial^2 \pi(p_n, z)}{\partial p_n^2} \right|_{N \rightarrow \infty} = -\frac{c_n}{p_n} \left[\theta \left(1 - \theta \left(1 - \frac{w/z}{p_n} \right) \right) + \theta \frac{w/z}{p_n} \right],$$

because $\lim_{N \rightarrow \infty} \varepsilon_d(p_n) = \theta$ and $\lim_{N \rightarrow \infty} \varepsilon'_d(p_n) = 0$. Moreover, the markup $p_n/(w/z)$ can be obtained from equalizing the first derivative to zero. The markup in this case is $\theta/(\theta - 1)$ and, as previously discussed, it is unique. Therefore,

$$\left. \frac{\partial^2 \pi(p_n, z)}{\partial p_n^2} \right|_{N \rightarrow \infty} = -\frac{c_n}{p_n}(\theta - 1) < 0 ,$$

so that when there are infinite many varieties, under the symmetric equilibrium the profit function has a unique maximizer, and it equalized the first derivative of the profit function to zero. Moreover, because c_n , p_n , $\varepsilon_d(p_n)$, and $\varepsilon'_d(p_n)$ are continuous in N , it is also the case that, in a symmetric equilibrium, $\frac{\partial^2 \pi(p_n, z)}{\partial p_n^2} < 0$ for large N . In the end, we concluded that if there is a large number of varieties, the profit function is concave, and $\partial \pi(p, z)/\partial p = 0$ characterizes its maximizer.

A.2 Proof of Proposition 1

Customer's decisions are sequential: first she decides if to incur in the search cost ψ and then, conditional on searching, she decides between staying and exiting depending on the draw of the new potential firm. We solve the customer's problem backwards, and thus determine first her optimal exit rule, conditional on searching. The exit strategy of the customer is $e(z, p, z') = 1$ if $\bar{V}(p, z) \leq \bar{V}(\mathcal{P}(z'), z')$, and $e(z, p, z') = 0$ otherwise. If $\hat{V}(z)$ is increasing in z , then $\bar{V}(p, z)$ is also increasing in z . As a result, the exit strategy takes the form of a trigger, $\hat{z}$, such that the customer exits if draws a firm with productivity $z' \geq \hat{z}$, where the threshold solves $\hat{V}(\hat{z}) = \bar{V}(p, z)$. Consider now the search decision of a customer who draws a search cost ψ . Because the value function in the case of searching is decreasing in ψ and the value function in the case of not searching does not depend on ψ , the search strategy takes the form of a trigger, $\hat{\psi}$, such that the customer searches if $\psi < \hat{\psi}$. The search strategy of the customer is $s(z, p, \psi) = 1$ if $\bar{V}(p, z) \leq \tilde{V}(p, z) - \psi$, and $s(z, p, \psi) = 0$ otherwise.

A.3 Proof of Lemma 1

The proof of Lemma 1 follows from the assumption of $v(p)$ being strictly decreasing in p so that $\bar{V}(p, z)$ is decreasing in p ; the threshold $\hat{z}(p, z)$ is increasing in z because of the assumptions that $\hat{V}(z)$ is increasing in z and the productivity process assumed to exhibit persistence, so that $\bar{V}(p, z)$ increases with z . The assumption that $\hat{V}(z)$ is increasing in z also implies that $\bar{p}(z)$ is decreasing in z , so that threshold $\hat{\psi}(p, z)$ is decreasing in z .

A.4 Proof of Proposition 2

First we prove that if $g(\psi)$ is continuous at zero and $p^*(z) \leq \bar{p}(z)$ then $\hat{\mathcal{P}}(z) = p^*(z)$ follows. Continuity of $g(\psi)$ at zero implies that $\Delta(p, z) \Pi(p, z)$ is continuously differentiable in p , so that [equation \(9\)](#) is necessary for an optimum. By definition of $p^*(z)$ and $\bar{p}(z)$, and given that firm's value is increasing in both profits per customer and customer base, $p^*(z) \leq \bar{p}(z)$ implies $\hat{\mathcal{P}}(z) = p^*(z)$. Given that for all $p \leq \bar{p}(z)$, $\frac{\partial \Delta(p, z)}{\partial p} = 0$, the first order condition in [equation \(9\)](#) reduces to $\frac{\partial \pi(p, z)}{\partial p} = 0$, which is uniquely satisfied at $p = p^*(z)$. Therefore, $p^*(z) \in \hat{\mathcal{P}}(z)$. Finally, given that $\pi(p, z)$ is strictly concave in p , attaining its maximum at $p^*(z)$, and $\Delta(p, z)$ non-increasing in p , we have that $\hat{\mathcal{P}}(z) = p^*(z)$.

We next show that if $p^*(z) > \bar{p}(z)$, then $\hat{\mathcal{P}}(z) \subset [\bar{p}(z), p^*(z))$. First we show $\hat{\mathcal{P}}(z) < p^*(z)$. Suppose in the contradiction that $\hat{\mathcal{P}}(z) = \hat{p} > p^*(z)$, then $\Delta(p, z)$ bein strictly decreasing in p for $p > \bar{p}(z)$ (from [Lemma 2](#)) and $\frac{\partial \pi(p, z)}{\partial p} < 0$ at any $p > p^*(z)$ and continuous, imply that there exists another value of $p < \hat{p}$ that gives higher value to the firm, contradicting optimality of $\hat{p} > p^*(z)$. Moreover, given results above, we have that $\hat{\mathcal{P}}(z) < p^*(z)$ (that is, $p^*(z) \notin \hat{\mathcal{P}}(z)$). Finally, notice that for $p < \bar{p}(z)$, $\Delta(p, z)$ is constant and profits per customer are strictly increasing in p , so that $\hat{\mathcal{P}}(z) \geq \bar{p}(z)$.

We next show that if $\varepsilon_m(p, z)$ is increasing in p for all $p \leq p^*(z)$ then [equation \(9\)](#) is also sufficient for an optimum, i.e. $\hat{\mathcal{P}}(z)$ is a singleton. It is useful to rewrite [equation \(9\)](#) as $\varepsilon_\Pi(p, z) = \varepsilon_m(p, z)$, where $\varepsilon_\Pi(p, z) \equiv \frac{\partial \Pi(p, z)}{\partial p} \frac{p}{\Pi(p, z)}$ is the elasticity of $\Pi(p, z)$ with respect to the price. Notice that $\varepsilon_\Pi(p, z)$ is strictly decreasing, with $\varepsilon_\Pi(p^*(z), z) = 0$. This follows from the fact that $\frac{\partial \varepsilon_\Pi(p, z)}{\partial p}$ is strictly decreasing and $\Pi(p, z)$ attaining its unique maximum at $p^*(z)$. Moreover, we have that $\varepsilon_m(p, z)$ is continuous at $\bar{p}(z)$ under the assumption that $g(\psi)$ is continuous at zero, and $\varepsilon_m(p, z) = 0$ for any $p < \bar{p}(z)$ by construction of $\Delta(p, z)$. Finally, given that we concluded that $\varepsilon_\Pi(p, z)$ is strictly decreasing, if $\varepsilon_m(p, z)$ is weakly increasing for any $p \leq p^*(z)$, we obtain that $\varepsilon_\Pi(p, z)$ and $\varepsilon_m(p, z)$ have to cross only once.

A.5 Proof of Lemma 3

It is useful to provide the expression for $\varepsilon_m(p, z)$,

$$\varepsilon_m(p, z) = -v'(p) \left(1 - \frac{H(\hat{z}(p, z))}{\Gamma} \right) g(\hat{\psi}(p, z)) \frac{p}{\Delta(p, z)} + \frac{\partial \hat{z}(p, z)}{\partial z} \frac{q(\hat{z}(p, z)) + h(\hat{z}(p, z))G(\hat{\psi}(p, z))}{\Gamma} \frac{p}{\Delta(p, z)}. \quad (15)$$

Positive elasticity $\varepsilon_m(p, z) \geq 0$. Notice that, using that $v'(p) < 0$ and $\partial \hat{z}(p, z)/\partial z = -v'(p)/\hat{V}'(\hat{z}(p, z))$, we obtain that the extensive margin elasticity is non-negative.

Monotonicity of $\varepsilon_m(p, z)$. An immediate implication of [Lemma 1](#) and [Lemma 2](#) is that $H(\hat{z}(p, z))$ and $\Delta(p, z)$ are decreasing in p and increasing in z , while $G(\hat{\psi}(p, z))$ is increasing

in p and decreasing in z . Moreover, as $v'(p) < 0$ and $v''(p) < 0$, we obtain that $-v'(p)p$ increases with p .

Let $\hat{\varepsilon}_m(p, z) \equiv \frac{g(\hat{\psi}(p, z))}{1-G(\hat{\psi}(p, z))}$. By the assumptions of the lemma, we have that $\hat{\varepsilon}_m(p, z)$ is increasing in p and decreasing in z after using [Lemma 1](#). Notice that, when $\lambda \rightarrow \infty$, $q(\hat{z}(p, z))$, $H(\hat{z}(p, z))$ and $G(\hat{\psi}(p, z))$ converge to zero, so that $\varepsilon_m(p, z) \rightarrow \hat{\varepsilon}_m(p, z)$. Given that $\varepsilon_m(p, z)$ is continuous in λ and both $\varepsilon_m(p, z)$ and $\hat{\varepsilon}_m(p, z)$ are continuous in both p and z , there exists a value $\bar{\lambda} > 0$ such that if $\lambda > \bar{\lambda}$ then $\left| \frac{\partial \varepsilon_m(p, z)}{\partial p} - \frac{\partial \hat{\varepsilon}_m(p, z)}{\partial p} \right| < K$ and $\left| \frac{\partial \varepsilon_m(p, z)}{\partial z} - \frac{\partial \hat{\varepsilon}_m(p, z)}{\partial z} \right| < K$ for some $K > 0$. As a result, for $\lambda > \bar{\lambda}$ we have that $\varepsilon_m(p, z)$ is increasing in p and decreasing in z .

A.6 Proof of [Proposition 3](#)

The outline of the proof is the following. We first conjecture that the equilibrium is such that $\hat{V}(z)$ is increasing in z . We then show that the firm pricing decision $\hat{p}(z)$ is continuous in z and in customer's thresholds, $\hat{z}(\cdot)$ and $\hat{\psi}(\cdot)$. Then, we show that, under the assumptions of the proposition, $\hat{z}(\hat{p}(z), z)$ and $\hat{\psi}(\hat{p}(z), z)$ are continuous in z and firm's optimal pricing $\hat{p}(z)$. As a consequence, we combine all these results to obtain a continuous mapping from prices $\hat{p}(z)$ into itself, and apply Brouwer's fixed-point theorem to show that there exists a solution $\hat{p}(z)$ where $\hat{V}(z)$ is increasing in z . We then show that if $\hat{V}(z)$ is increasing in z then the equilibrium is such that $\hat{p}(z) = p^*(z)$ for $z = \bar{z}$, and $\hat{p}(z) < p^*(z)$ otherwise. Finally, we characterize the properties of the prices, customer growth rate and markups.

Existence of equilibrium. Let $\mathcal{T}_1(z, \hat{z}, \hat{\psi})$ denote the mapping from productivity z and the thresholds price $\hat{z}$ and $\hat{\psi}$ to optimal pricing $\hat{p}(z)$.

Claim 1 *If the assumptions of [Lemma 3](#) are satisfied, the mapping $\mathcal{T}_1(\cdot)$ is unique and continuous in all its arguments.*

The proof follows immediately from [Proposition 2](#). Because the first order condition is continuous in p , $\hat{z}$ and $\hat{\psi}$, $\mathcal{T}_1(\cdot)$ is continuous in all its arguments.

Let $\mathcal{T}_2(z, \hat{p})$ and $\mathcal{T}_3(z, \hat{p})$ denote the mapping from productivity z and the optimal price function $\hat{p}(\cdot)$ to thresholds $\hat{z}$ and $\hat{\psi}$ respectively.

Claim 2 *The mappings $\mathcal{T}_2(\cdot)$ and $\mathcal{T}_3(\cdot)$ are unique and continuous in both arguments.*

The proof follows immediately from [Proposition 1](#), and the assumptions that $v(\cdot)$ and $f(\cdot)$ are continuous.

Let $\mathcal{T}(z, \hat{p}) \equiv \mathcal{T}_1(z, \mathcal{T}_2(z, \hat{p}), \mathcal{T}_3(z, \hat{p}))$ denote the function that maps productivity z and a pricing function $\hat{p}(\cdot)$ into the space where $\hat{p}(\cdot)$ belongs.

Claim 3 $\mathcal{T}(\cdot)$ is continuous in both arguments.

The proof follows because the composition of continuous functions is continuous.

Notice, given that the set of z is compact and given [Proposition 2](#), we have that $\hat{p}(z)$ lies in a compact set. Moreover, we know that $\mathcal{P}(z) \in [a, \hat{p}(z)]$ where a is a finite number, so that the operator $\mathcal{T}(z, \mathcal{P})$ maps prices $\mathcal{P}$ from the set $[a, \hat{p}(z)]$ to a subset of it. By applying Brouwer's fixed-point theorem, we obtain existence of a solution $\hat{p}(z) = \mathcal{T}(z, \hat{p})$ for all z .

$\hat{V}(z)$ increasing in z . We want to show that $\hat{V}(z)$ is increasing in z . This requires to show that the equilibrium operator maps increasing functions $\hat{V}(z)$ into increasing functions $\hat{V}^0(z)$. It is useful to define the following object,

$$\underline{V}(z) = \int_z^{\bar{z}} \left[\int_0^\infty \max \left\{ V(\mathcal{P}(z'), z'), \tilde{V}(\mathcal{P}(z'), z') - \psi \right\} g(\psi) d\psi \right] f(z'|z) dz',$$

so that the value for a remaining customer is $\hat{V}^0(z) = v(\mathcal{P}(z)) + \beta \underline{V}(z)$. First, notice that if $\hat{V}(z)$ is increasing in z , then $\underline{V}(z)$ is increasing in z , therefore the only way the resulting $\hat{V}^0(z)$ is decreasing in z in a given interval of z is if $\hat{p}(z)$ is increasing in z in that interval. Second, notice that this is possible only in the region of productivity z where customer concerns are binding, i.e., where $\hat{p}(z) < p^*(z)$, as where such concerns are not binding we have that $\hat{p}(z) = p^*(z)$ which is decreasing in z , and therefore the operator would map an increasing function $\hat{V}(z)$ into an increasing function. Suppose by contradiction that there exist two values $z_H > z_L$ such that $\hat{V}^0(z_H) < \hat{V}^0(z_L)$ requiring $\hat{p}(z_H) > \hat{p}(z_L)$, $\hat{p}(z_L) < p^*(z_L)$ and $\hat{p}(z_H) < p^*(z_H)$. Then $\hat{p}(z_H) > \hat{p}(z_L)$ and $\hat{V}^0(z_H) < \hat{V}^0(z_L)$ imply $\Delta(\hat{p}(z_H), z_H) < \Delta(\hat{p}(z_L), z_L)$, while $\hat{p}(z_H) > \hat{p}(z_L)$ and $z_H > z_L$ imply $\partial \pi(\hat{p}(z_H), z_H) / \partial p < \partial \pi(\hat{p}(z_L), z_L) / \partial p$. Moreover, by persistence of z , $z_H > z_L$, the definition of $\hat{p}(z)$, and noticing that firm z_H can always replicate pricing of firm z_L and have higher value, it follows that $\Pi(\hat{p}(z_H), z_H) > \Pi(\hat{p}(z_L), z_L)$; if the assumptions of [Lemma 3](#) are satisfied $\hat{p}(z_H) > \hat{p}(z_L)$ and $\hat{V}^0(z_H) < \hat{V}^0(z_L)$ imply $\hat{\psi}(\hat{p}(z_L), z_L) < \hat{\psi}(\hat{p}(z_H), z_H)$ so that $g(\psi)/(1 - G(\psi))$ increasing gives $-\partial \Delta(\hat{p}(z_H), z_H) / \partial p > -\partial \Delta(\hat{p}(z_L), z_L) / \partial p$. However, given that the first order condition in [equation \(9\)](#) is necessary and sufficient for optimal prices, we have a contradiction as the results above imply that $\hat{p}(z_H)$ and $\hat{p}(z_L)$ cannot both satisfy the first order condition.

Part (i). We first show that, in equilibrium, there is price dispersion. That is, we show that there exists at least two productivity levels z and y for which $\hat{p}(z) \neq \hat{p}(y)$. Suppose by contradiction that for any two productivity levels z and y we have that $\hat{p}(z) = \hat{p}(y) \equiv \bar{p}$. Then, because prices are independent of productivity, $\bar{p}(z) = \bar{p} = \hat{p}$ for all z . This

implies that $\tilde{V}(p, z) = \tilde{V}$ and that $\bar{p} = p^*(\bar{z})$. In fact, if $\bar{p} < p^*(\bar{z})$, by [Proposition 2](#), firms with productivity $\bar{z}$ would deviate and charge a price above $\bar{p}$. If $\bar{p} > p^*(\bar{z})$, also by [Proposition 2](#), firms with productivity $\bar{z}$ would deviate and charge a price below $\bar{p}$. In both cases, $V(\hat{p}(z), z) \neq \tilde{V}$ for all z . Because $\bar{p} = p^*(\bar{z})$ and $p^*(z)$ decreasing in productivity, it is the case that $p^*(\underline{z}) > \bar{p}$. Then, by [Proposition 2](#), firms with productivity $\underline{z}$ deviate to charge a price strictly above $\bar{p}$, so that there is price dispersion.

Parts (ii)-(iii)-(iv)-(v). It is useful to recall that a customer exits the customer base of the firm if $\bar{V}(\hat{p}(z), z) < \tilde{V}(\hat{p}(z), z) - \psi$. We first show that in equilibrium there is at least one productivity level where firms experience customers leaving. We prove this by contradiction, by assuming that no firm experiences customers exiting its customer base. If no customer exits from the customer base of any firm we need that $\bar{V}(\hat{p}(z), z) = \tilde{V}(\hat{p}(z), z)$ for all z . Conjecture that at $\hat{p}(z)$ for all z no firm experiences a customer outflow. Notice that it cannot be the case that for some z we have that $\hat{p}(z) > p^*(z)$: given that $\pi(p, z)$ is decreasing in this region, firms at this productivity level should decrease their price up to $p^*(z)$ and still no customer would exit. A similar argument also provides that it cannot be the case that for some z we have that $\hat{p}(z) < p^*(z)$: given that $\pi(p, z)$ is increasing in this region, the fact that $g(\cdot)$ continuous guarantees that $\Delta(p, z)$ is continuous implies that firms find it profitable to increase their price, and thus experience a customer outflow. As a result, if no customer would exit a firm the only possible price function is $\hat{p}(z) = p^*(z)$ for all z , a decreasing function of productivity z . Given that the productivity process exhibits persistence in the sense that $\int_{\bar{z}}^s f(z'|z)dz'$ is weakly decreasing in z , the fact that $\hat{p}(z) = p^*(z)$ for all z immediately implies that firms with the lowest productivity, $\underline{z}$, will exhibit a customer outflow. As a result, there exists at least one productivity level z at which a firm experiences a customer outflow, $\hat{p}(z) < p^*(z)$. This also implies that there is a productivity level at which no firm experiences a customer outflow. This follows because, if some firms' optimal price is such that $V(\hat{p}(z), z) < \tilde{V}(\hat{p}(z), z)$, then for another set of firms it has to be the case that $V(\hat{p}(z), z) > \tilde{V}(\hat{p}(z), z)$. Because no one exits the customer base of these firms, which implies that $p^*(z) < \bar{p}(z)$, by applying [Proposition 2](#) we get that, for firms with these productivity levels, $\hat{p}(z) = p^*(z)$.

We now show that only firms with the highest productivity level, i.e., those with $z = \bar{z}$, do not experience customer base concerns and therefore are the only ones charging the price that maximizes static profits. Notice first that the set of productivity levels at which firms do not face customer base concerns includes $\bar{z}$. The argument is the following. Because $\hat{V}(z)$ is increasing in z and the distribution of $f(z'|z_H)$ first order stochastically dominates $f(z'|z_L)$ for any $z_H > z_L$, we have that $\underline{V}(z)$ is increasing in z . By definition of $\bar{p}(z)$, the fact that

$\underline{V}(z)$ is increasing in z implies that $\bar{p}(z)$ is increasing in z . Then, because we know that at some productivity levels firms do not face customer base concerns, we know that $\bar{p}(z)$ and $p^*(z)$ intersect, and the intersection is unique because of monotonicity of both functions. This immediately implies that if firms with productivity $z < \bar{z}$ do not face customer base concerns, also firms with productivity $\bar{z}$ do not face customer base concerns. We now show that only firms with the highest productivity do not face customer base concerns. We prove this by contradiction, by assuming that there exists at least another productivity level $y < \bar{z}$ at which firms do not face customers exiting its customer base. Notice that because $\hat{V}(z)$ is increasing in z , it is more desirable to be part of the customer base of a firm with productivity $\bar{z}$ than with productivity y . Moreover, because some customers have a search cost of zero, they can search at no cost and only exit the customer base of firms with productivity y if they draw a firm with productivity $\bar{z}$. Because $h(\bar{z}) > 0$, some of these searching customers will draw firms with the highest productivity and will therefore exit the customer base of the firms with productivity level y . This is a contradiction to the conjecture. Therefore, only firms with the highest productivity, i.e., those with productivity $\bar{z}$, do not face customer base concerns. As a result, $\hat{p}(\bar{z}) = p^*(\bar{z})$ and $\hat{p}(z) < p^*(z)$ for all $z < \bar{z}$. Another implication of this result is that $\hat{\psi}(\hat{p}(\bar{z}), \bar{z}) = 0$ and $\hat{\psi}(\hat{p}(z), z) > 0$ for all $z < \bar{z}$. Furthermore, $\bar{V}(\hat{p}(z), z) = \hat{V}(z)$ increasing in z directly implies $\hat{z}(\hat{p}(z), z) = z$ by definition of $\hat{z}$. In the end, we can write $\hat{\psi}$ as $\hat{\psi}(\hat{p}(z), z) = \int_{\hat{z}}^{\bar{z}} \max\{\hat{V}(z), \hat{V}(x)\} \frac{dH(x)}{\Gamma} - \hat{V}(z)$, where $\hat{V}(\cdot)$ increasing in z directly implies that $\hat{\psi}(\hat{p}(z), z)$ decreases as z increases.

Finally, because $\hat{V}(z)$ is increasing in z , also does $\Delta(\hat{p}(z), z)$. Moreover, because firms with productivity level $\bar{z}$ do not experience a gross customer outflow and do experience a gross customer inflow (because those customer that get in contact with these firms join their customer base), we have that $\Delta(\hat{p}(\bar{z}), \bar{z}) > 1$. Moreover, because (1) it has to be the case (for consistency) that some at some productivity level firms experience a net outflow of customer, and (2) $\Delta(\hat{p}(z), z)$ increasing in z , we have that $\Delta(\hat{p}(z), z) < 1$.

Part (vi). It is useful to write optimal markups as

$$\mu(\hat{p}(z), z) = \frac{\varepsilon_d(\hat{p}(z)) + \varepsilon_m(\hat{p}(z), z) \Pi(\hat{p}(z), z) / \pi(\hat{p}(z), z)}{\varepsilon_d(\hat{p}(z)) - 1 + \varepsilon_m(\hat{p}(z), z) \Pi(\hat{p}(z), z) / \pi(\hat{p}(z), z)}.$$

Given $\hat{V}(z)$ increasing in z , as we showed before we have that $\hat{\psi}(\hat{p}(z), z)$ decreases with z and $\hat{z}(\hat{p}(z), z)$ increases with z . Then, given **Lemma 3** we have that $\varepsilon_m(\hat{p}(z), z)$ decreases with z . Moreover, mean reversion of z and $\Delta(\hat{p}(z), z) \approx 1$ because of **Lemma 3** implies that $\Pi(\hat{p}(z), z) / \pi(\hat{p}(z), z)$ decreases with z . Finally if $\varepsilon'_d(p) = 0$ or $\hat{p}(z)$ decreasing in z it follows that $\varepsilon_d(\hat{p}(z))$ is non-increasing in z so that $\mu(\hat{p}(z), z)$ increases with z .

A.7 Proof of Remark 2

We first provide a proof of the first part of the Remark. Under the assumptions of the remark every firm has the same productivity z at every point in time. As a result, as previously discussed, every firm chooses the same price $\hat{p}(z)$. We prove the statement in two steps. In the first step we conjecture a solution and then verify that it is an equilibrium. In the second step we show that there are no other solutions.

Step 1: Conjecture that firms choose $\hat{p}(z) = p^*(z)$. It is immediate to see that $\bar{p}(z) = p^*(z)$. Because $\hat{p}(z) = \bar{p}(z)$, a direct application of **Proposition 2** validates the conjecture.

Step 2: We now show by contradiction that $\hat{p}(z) = p^*(z)$ is the unique price schedule in equilibrium. There are two cases: one were firms price above $p^*(z)$ and one were they price below. For the first case, conjecture that $\hat{p}(z) > p^*(z)$. Again, it is immediate to see that $\bar{p}(z) > p^*(z)$. However, by **Proposition 2**, given that $p^*(z) < \bar{p}(z)$ firms should deviate and price at $p^*(z)$. Therefore, $\hat{p}(z) > p^*(z)$ cannot happen in equilibrium. For the second case, conjecture that $\hat{p}(z) < p^*(z)$. Now we have that $\bar{p}(z) < p^*(z)$. However, by **Proposition 2**, given that $p^*(z) > \bar{p}(z)$ a firm's optimal price should satisfy $\bar{p}(z) < \hat{p}(z) < p^*(z)$. Therefore, $\hat{p}(z) < p^*(z)$ cannot happen in equilibrium.

The proof of the second part of the corollary is straightforward and follows from Bertrand's competition.

A.8 Proof of Remark 3

It proves useful to define the following object,

$$\underline{V}(z) = \int_{\underline{z}}^{\bar{z}} \left[\int_0^\infty \max \left\{ V(\mathcal{P}(z'), z'), \tilde{V}(\mathcal{P}(z'), z') - \psi \right\} g(\psi) d\psi \right] f(z'|z) dz'.$$

So that the value for a remaining customer is $\bar{V}(p, z) = v(p) + \beta \underline{V}(z)$.

1) *Productivity shocks are i.i.d., $\rho = 0$. Points (i)-(ii).* We start by showing that $\bar{p}(z) = \bar{p}$ for all z . Under the assumptions of the proposition we have that $f(z'|z) = f(z'|y)$ for all z, y . As a result, we get that $\tilde{V}(\mathcal{P}(z), z) = \tilde{V}$ and $\underline{V}(z) = \underline{V}$. Because the continuation value of being matched with any firm equals $\underline{V}$, we get that $\bar{p}(z) = \bar{p}$. Notice that this result immediately implies that $\varepsilon_m(p, z) = \varepsilon_m(p)$.

To prove that there exists a unique value z^* recall that $p^*(z)$ is decreasing in z . This, together with **Proposition 3** and $\bar{p}(z)$ constant in z , immediately implies the existence of the unique threshold z^* where firms with productivity above $\hat{z}$ charge $p^*(z)$.

We now prove that the pricing schedule $\hat{p}(z)$ is monotonic. We do this in three steps:

Step 1: firms charging $\hat{p}(z) = p^(z)$.* For every firm with productivity $z \in \mathcal{Z}_1$ the optimal

price $\hat{p}(z)$ decreases monotonically when $z \in \mathcal{Z}_1$ because $p^*(z)$ decreases monotonically.

Step 2: firms charging $\hat{p}(z) < p^(z)$.* Consider two productivity levels z, z' with $z > z'$. Conjecture, in the contradiction, that $\hat{p}(z) > \hat{p}(z')$. There are two cases. Case 1: suppose that $\hat{p}(z') > p^*(z)$. Here, because $\partial\pi(p, z)/\partial p < 0$ for all $p > p^*(z)$ and $\bar{p}(z) = \bar{p}$ for all z , a firm with productivity z has a deviation that increases her value by charging price $\hat{p}(z')$. Case 2: suppose that $\hat{p}(z') < p^*(z)$. Here, because $\partial\pi(p, z')/\partial p > 0$ for all $p < p^*(z')$ and $\bar{p}(z) = \bar{p}$ for all z , a firm with productivity z' has a deviation that increases her value by charging $\hat{p}(z)$. Hence, in both cases we get a contradiction. Therefore, $\hat{p}(z)$ decreases monotonically when $z \in \mathcal{Z}_2$.

Step 3: the firms with lowest productivity in set $\mathcal{Z}_1$ and firm with highest productivity in set $\mathcal{Z}_2$. Consider the firm with the lowest productivity in $\mathcal{Z}_1$, which we label by z , and the firm with highest productivity belonging to set $\mathcal{Z}_2$, which we label by z' . We now show that $\hat{p}(z) = p^*(z) < \hat{p}(z')$. Consider, in the contradiction, that $\hat{p}(z') < \hat{p}(z)$. Notice that (i) the firms with productivity z do not face customer retention concerns and (ii) $\partial\pi(p, z')/\partial p > 0$ for any $p < p^*(z')$. Because of (i) and (ii), it is immediate that a firm with productivity z' has a deviation that increases her value by setting a price equal to $\hat{p}(z) = p^*(z)$.

2) *Persistent productivity shocks, $\rho \in (0, 1)$. Points (i)-(ii).* We will prove the statement by contradiction. Pick two productivity levels $z_l < z_h$. Conjecture that $\hat{p}(z_l) \leq \hat{p}(z_h)$. We rewrite $\underline{V}(z)$ as follows,

$$\underline{V}(z) = (1 - \rho) \tilde{V} + \rho \int_0^\infty \max \left\{ \bar{V}(\hat{p}(z), z), \tilde{V}(\hat{p}(z), z) - \psi \right\} g(\psi) d\psi$$

where $\tilde{V} = \int_{\bar{z}}^{\bar{z}} \left[\int_0^\infty (\bar{V}(\mathcal{P}(z'), z') - \psi) g(\psi) d\psi \right] f(z'|z) dz'$, and from where it is clear that if $\hat{p}(z_l) \leq \hat{p}(z_h)$ then it is necessary that $\bar{p}(z_l) \geq \bar{p}(z_h)$. We now show that when $\bar{p}(z_l) \geq \bar{p}(z_h)$, it is the case that $\hat{p}(z_l) > \hat{p}(z_h)$, which constitutes a contradiction.

Notice that we can rewrite [equation \(8\)](#) as $F(z) = \max_p H_1(p, z) + H_2(p, z)$ where

$$\begin{aligned} H_1(p, z) &\equiv \frac{\Delta(p, z)}{1 - \beta\rho\Delta(p, z)} \pi(p, z) , \\ H_2(p, z) &\equiv \frac{\Delta(p, z)}{1 - \beta\rho\Delta(p, z)} \beta(1 - \rho)\mathcal{F} , \end{aligned}$$

where $\mathcal{F}$ is a constant. Notice that, using [equation \(??\)](#), $\arg \max_p H_1(p, z_H) < \arg \max_p H_1(p, z_L)$ if $\bar{p}(z_l) \geq \bar{p}(z_h)$. Also, notice that $H_2(p, z)$ is constant for any $p \leq \bar{p}(z)$ and then strictly decreasing. As a result, it follows immediately that, as long as $\bar{p}(z_l) \geq \bar{p}(z_h)$, the $\arg \max H_1(p, z_H) + H_2(p, z_H) < \arg \max H_1(p, z_L) + H_2(p, z_L)$ so that $\hat{p}(z_H) < \hat{p}(z_L)$, which constitutes a contradiction. Then, $\hat{p}(z)$ decreases monotonically with z .

Because $\hat{p}(z)$ is decreasing in z , $\bar{p}(z)$ is strictly increasing in z : this follows from the definition of $\underline{V}(z)$ and $\bar{p}(z)$. To prove that there exists a unique value z^* recall that $p^*(z)$ is decreasing in z . This, together with [Proposition 3](#) and $\bar{p}(z)$ strictly increasing in z , immediately implies the existence of the unique threshold z^* where firms with productivity above z^* charge $p^*(z)$.

A.9 Closing the model: a simple model of labor choice

We describe the workings of the labor market determining the equilibrium level of income I . We assume that each period a representative household chooses labor supply to solve the following problem

$$\max_{\ell} \int_{\underline{z}}^{\bar{z}} \left(v(\hat{p}(z)) - \int_0^{v(\bar{p}(z)) - v(\hat{p}(z))} \psi g(\psi) d\psi \right) \frac{M(z)}{\Gamma} dz - \frac{\ell^{1+\phi}}{1+\phi}, \quad (16)$$

subject to

$$v(\hat{p}(z)) = \frac{\left(I (\omega^{1-\theta} \hat{p}(z)^{1-\theta} + (1-\omega)^{1-\theta})^{-\frac{1}{1-\theta}} \right)^{1-\gamma}}{1-\gamma},$$

$$I = w \ell + \int_{\underline{z}}^{\bar{z}} \pi(\hat{p}(z), z) \frac{M(z)}{\Gamma} dz.$$

The representative household takes prices and aggregate profits as given, and has the same measure of the customers, i.e Γ . The representative household chooses the labor supply ℓ and distributes labor proceeds equally across its customers. The parameter representing the disutility of labor (ϕ) is set to 1.43 so that the Frisch elasticity of labor supply is equal to 0.7 (see [Pistaferri \(2003\)](#)).

A.10 Augmenting the model with government spending

For simplicity, we assume that the government behaves exactly as the consumers in our economy, i.e. it is allocated a fraction g_t of household's steady state income, so that total government expenditure is $g_t I \Gamma$. In the steady state we assume $g_t = 0$. Government purchases are of no value to the household. Thus the only difference between the representative household and the government in our model is that the representative household also makes a labor decision, taking the lump-sum transfer by the government given, while the government's shoppers take similar decisions as the shoppers in the household. The impact of higher government spending on the household is that it decreases disposable income for given labor

proceeds and profits, while it leaves firms' demand unchanged.

The lump-sum tax on the household is such that the government budget balances every period, i.e. $\tau_t = g_t I \Gamma$. As a consequence, if $d(\hat{p}(z))$ and $n(\hat{p}(z))$ are the demand of the good d and the numeraire good by household' shoppers when matched to a producer of good d with productivity z , the corresponding demand by the government's shoppers are $d(\hat{p}(z)) g_t I / (I_t - g_t I)$ and $n(\hat{p}(z)) g_t I / (I_t - g_t I)$ respectively.

The equilibrium in the labor markets requires that labor supply equals labor demand, i.e.

$$\ell = \frac{I_t}{I_t - g_t I} \left(\int_{\underline{z}}^{\bar{z}} \frac{d(\hat{p}(z))}{z} \frac{M(z)}{\Gamma} dz + n(\hat{p}(z)) \right). \quad (17)$$

Finally, given that the representative household owns the firms, the relevant T-periods ahead stochastic discount factor in the firm maximization problem is given by

$$Q_{t+T|t} = \beta^T \frac{\int_{\underline{z}}^{\bar{z}} (C(\mathcal{P}_{t+T}^*(z)))^{-\gamma} M(z) dz}{\int_{\underline{z}}^{\bar{z}} (C(\mathcal{P}_t^*(z)))^{-\gamma} M(z) dz}, \quad (18)$$

where $U(C) = C^{1-\gamma}/(1-\gamma)$, and $C(\mathcal{P}_t^*(z)) = (I_t - g_t I) (\omega^{1-\theta} \mathcal{P}_t^*(z)^{1-\theta} + (1-\omega)^{1-\theta})^{-\frac{1}{1-\theta}}$ is the consumption basket of consumers buying good d at price $\mathcal{P}_t^*(z)$.

B Data sources and variables construction

B.1 Data sources

The empirical evidence presented in [Section 3](#) is based on two data sources provided by a large supermarket chain that operates over 1500 stores across the US. We exploit information on weekly store sales between January 2004 and December 2006 for a panel of over 200 stores located in 10 different states. For each good (identified by its UPC) carried by the stores in those weeks, the data report total amount grossed and quantity sold.

In addition to store level data, we have information on grocery purchases at the chain between June 2004 and June 2006 for a panel of over 11,000 households. For each grocery trip made by a household, we observe date and store where the trip occurred, the collection of all the UPC's purchased with quantity and price paid. The data include information on the presence and size of price discounts but do not generally report redemption of manufacturer coupons. The geographical dispersion of the households mirrors that of the store data: our customers live in some 1,500 different zipcodes across 10 states. Data are recorded through usage of the loyalty card; the retailer is able to link loyalty cards belonging to different membebr

of the family to a single household identifier. Purchases made without using the card are not recorded. However, the chain ensures that the loyalty card has a high penetration by keeping to a minimum the effort needed to register for one. Furthermore, nearly all promotional discount are tied to ownership of a loyalty card, which provides a strong incentive to use it. Another potential drawback of the data is that we only follow households when purchasing at stores of a single, albeit large, supermarket chain. Other data sources on the same industry, like the Nielsen Homescan database, rely on households themselves scanning the barcodes of the items purchased once they return home after a trip and can therefore track them shopping at a plurality of competing firms. On the other hand, cash register data contain significantly less measurement error than databases relying on home scanning (Einav et al. (2010)).

B.2 Variables construction

Exit from customer base. The dependent variable in the regression presented in equation (14) 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 household has abandoned the retailer to shop elsewhere or it is simply not purchasing grocery in a particular week, for instance because it is leaving off its inventory. In fact, we observe households when they buy grocery at the chain but do not have any information on their shopping at competing grocers. To circumvent this problem, we focus on a subsample of households who shop frequently at the chain. For them we can plausibly assume that sudden long spells without trips represent instances in which the household has left the chain and is fulfilling grocery needs shopping at one of its competitors. Operationally, we select households who made at least 48 trips at the chain over the two years spanned in the sample, implying that they would shop on average twice per month at the chain. When such households do not visit any supermarket store of the chain over at least eight consecutive weeks, we assume that the customer is shopping elsewhere. The *Exit* dummy is 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 3 summarizes shopping behavior for households in our sample. It is immediate to notice that a 8-weeks spell without purchase is unusual, as customers tends to show up frequently at the stores. This strengthens our confidence that customer missing for such a long period have indeed switched to a different retailer.

Price of the basket. The household level scanner data report 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 would not be able to infer the price of the item in that store-week from the household

Table 3: Descriptive statistics on customer shopping behavior

	<i>Mean</i>	<i>Std.dev.</i>	<i>25th pctl</i>	<i>75th pctl</i>
Number of trips	157	141	65	208
Days elapsed between consecutive trips	4.1	7.4	1	5
Frequency of exits	0.004	0.065		
Items in the basket	289.5	172.4		

panel data. However, the store level data allow us to calculate unit value prices every week for every item in sale in a given store, whether or not that particular UPC was bought by one of the households in our data. Unit value prices are computed using data on revenues and quantities sold as

$$UVP_{stu} = \frac{TR_{stu}}{Q_{stu}} ,$$

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

As explained in [Eichenbaum et al. \(2011\)](#) this only allows 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 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 sales, 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, an unfrequent circumstance and involves only rarely purchased UPC's, which are unlikely to represent important shares of the basket for any of the households in the sample. For the analysis, we only retain UPC's with at most two non consecutive missing price observations and impute price for the missing observation interpolating the prices of the contiguous weeks.

The retail chain applies different prices in different geographic areas and supplied weekly data on revenues and quantities sold by UPC for 270 stores that are representative of the different price areas. Households shop in one (or a subset) of some 1,500 stores and we have to devise a way to match the store a household visits to the price areas to which it belongs. However, we have no information on how the chain divides its markets into price areas. 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 attention to the 1,336 households whose most frequently visited store is one of the representative stores. Since the 270 representative stores are not selected following any particular criterion, the resulting subsample of households is not subject to any type of selection.

We are interested in observing whether households change their grocery supplier in response to fluctuations in the price of the basket of goods they purchase. To this end, we construct a price index summarizing for each customer in every week the price of the collection of goods she regularly buys. We include in a customer basket all the UPC's she purchased over the two years of data and construct the price of the basket for household i in week t by taking the average of the weekly prices of all the UPC's the customer purchased over the two years weighted by the share of her expenditure they represent. Namely:

$$p_{it} = \sum_{u \in U^i} w_{iu} p_{ut}, \quad w_{iu} = \frac{\sum_t E_{iut}}{\sum_{u \in U^i} \sum_t E_{iut}}$$

where U^i is the set of all the UPC's (u) purchased by household i during the sample period.

We choose to calculate the weights using the expenditure share of the UPC 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 twelve 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.

Table [Table 4](#) reports descriptive statistics on the change in price of the basket.

Table 4: Descriptive statistics on basket price changes

	<i>Mean</i>	<i>Std.dev.</i>
Δp	-0.0001	0.043
$ \Delta p $	0.029	0.031
$\% \Delta p > 1\%$	73.5	
$\% \Delta p > 5\%$	16.8	
$\% \Delta p > 10\%$	4.1	