

Price Setting with Customer Retention Concerns*

Luigi Paciello

Einaudi Institute for Economics and Finance and CEPR

Andrea Pozzi

Einaudi Institute for Economics and Finance

Nicholas Trachter

Federal Reserve Bank of Richmond

October 4, 2013

Abstract

We consider a model where customers incur search costs when they decide to change the firm they buy from, thus generating sluggishness in the firm's customer base. In this setting, firms care about retaining customers and this affects their pricing strategy. The model delivers two important results consistent with a growing body of empirical evidence. First, pass-through of cost shocks is incomplete as firms are wary of raising prices for fear of losing customers. Second, the elasticity of demand faced by a firm depends on its productivity with better performing firms facing a less elastic demand. Therefore, pass-through is heterogeneous across firms. We exploit scanner data on households purchases at a large US supermarket chain to produce novel direct empirical evidence on the relationship between customer dynamics and pricing. We estimate the key parameters of the model and perform counterfactual exercises to quantify the importance of the customer retention channel in generating low pass-through of cost and its dispersion.

JEL classification:

Keywords:

*A previous version of this paper circulated under the title "Price Setting with Customer Retention Concerns". We thank Mark Aguiar, George Alessandria, Fernando Alvarez, Lukasz Drozd, 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, Leena Rudanko, Randall Wright, and participants at the 2nd Rome Junior Macroeconomics conference, 10th Philadelphia Search and Matching conference, ESSET 2013, the Minnesota Macro Workshop 2013, the Macro Faculty Lunch at Stanford, the MBF-Bicocca conference, and University of Tor Vergata for useful comments. 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 are important determinants of a firm's performance (Gourio and Rudanko (2011)), and that a firm's pricing policy and the evolution of its customer base are intimately related (Foster et al. (2012)). In this paper we study a general equilibrium setup where the build up of the customer base of a firm depends on its current and future prices. Using novel micro evidence we quantify the relevance of this channel and explore its pricing implications, all of them consistent with several documented facts.

The setup we study is one where customers react to the pricing policy of a firm by not only demanding more or less units of the good, but also by exiting the customer base of the firm. The introduction of *customer retention concerns* to an otherwise standard model delivers a set of predictions supported by an increasing stack of empirical evidence. The literature documents that (i) markups are pro-cyclical (Petrin and Warzynski (2012)), (ii) the response of firms to cost shocks is incomplete (Nakamura and Zerom (2010), Goldberg and Hellerstein (2013)), (iii) heterogeneous across firms (Berman et al. (2012)), and (iv) asymmetric with respect to the sign of the shock (Goldberg (1995), Pollard and Coughlin (2004)). In a model where firms are wary of raising prices for fear of losing patrons, firms are willing to swallow part of the cost fluctuations rather than passing them to consumers. Moreover, the dynamic aspect of the problem implies that the demand elasticity faced by the firm depends on its productivity through its effect on future expected prices. Therefore, response to cost shocks will depend on the firm characteristics.

Motivated by the evidence in Kaplan and Menzio (2013) we incorporate customer base concerns into the model by assuming that customers need to incur in a search cost if they want to change the firm they patronize. This friction generates persistence in the customer base -i.e. the set of customers currently purchasing from a firm- and makes it valuable to try and keep customers around. We use data from the retail industry in the US to estimate the parameters of the distribution of this search cost, which allows us to document the effect of customer retention on markups and pass-through of cost shocks.

We construct a search model where customers have demand $d(p)$ for an homogeneous good sold in different locations, each of them populated by a single firm. Each customer is matched to a particular firm at any point in time and must decide whether to stay in the same location, and purchase from the same supplier as in the last period, or to look for a new provider. Firms are characterized by different productivity levels, which evolve stochastically through time. The price set by a firm depends on its productivity, implying price dispersion and thus incentivizing customers to search. In this setting firms face two

margins of demand. The first one is the usual *intensive margin*, governed by the elasticity $\varepsilon_d(p)$. The intensive margin concerns *how much* of the good the firm is able to sell to each of its customers, conditional on the customer buying goods from the firm. The second margin, which we label *extensive margin* of demand, concerns the decision of customers to remain in the customer base. We show that this margin is governed by the elasticity $\varepsilon_m(p, z)$ and relates to *how many* individuals decide to remain and buy from their old firm. The price p enters the extensive margin because customers care about the current price when deciding if to exit the customer base. Firm productivity z enters the extensive margin because it carries information about future prices and customers are more reluctant to leave firms that are likely to have lower prices in the future. For example, if productivity is persistent and prices are monotonic in productivity, the extensive margin elasticity will be lower for more productive firms.

On the empirical side, despite the growing attention towards the implications of sluggishness in the customer base, evidence in that regards is still mostly based on anecdotal or survey sources. Our study, instead, provides novel direct evidence of this fact by looking at the evolution of the customer base of a large U.S. supermarket chain. Supermarket scanner data are particularly suited to study customer base dynamics. We document that there is persistence in a customer's choice of a retailer and that the decision to exit the customer base of a firm responds to the price faced by the customer. We observe that shoppers tend to shop at the same supermarket chain for long spells of time, confirming that the buyer-supplier relationship is long lasting. However, the firm experiences dynamic in its customer base as some habitual shoppers in our sample stop buying at the chain, presumably moving their business to a competing retailer. We investigate the relationship between *breaks* in the customer-retailer match and pricing policy of the retailer. We construct a measure of customer retention and evaluate how the customer base of a retailer responds to prices: 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 0.014 percentage points.

We calibrate the model combining the microdata on households' shopping decisions and some standard macroeconomic targets. The quantitative exploration of the model stresses the relevance of the effect of the presence of the extensive margin of demand on firms' pricing strategies, affecting both markups and pass-through of shocks. At the parameter estimates the implied average markup is 14%, slightly above half of the markup that would be obtained if the retention frictions were to be eliminated. Moreover, the implied pass-through of cost on prices is 20%, a quarter of that one obtained without the retention concerns.

The presence of customer base concerns provides a real world motivation for a model where the demand elasticity faced by a firm depends on its productivity. The dependence of

the elasticity on productivity is crucial in order to generate two documented facts: low pass-through of cost shocks and dispersion of pass-through across firms with different productive efficiency. The same result can be achieved making specific assumptions on primitives of the model. For example, [Atkeson and Burstein \(2008\)](#) assumes Cournot competition with a nested CES demand system.¹ In this case the dependence of markups on productivity takes place through the *contemporaneous* price, as both the pricing and demand decision are static choices. The larger the dispersion in productivity, the larger the dispersion in markups. Our model nests these frameworks by allowing the intensive margin elasticity $\varepsilon_d(p)$ to depend on p : everything else being equal, when $\varepsilon'_d(p) > 0$ markups increase with productivity as more productive firms charge lower prices. Crucially, our mechanism makes the markups to depend also on the path of *future* prices because the customers' exit decision is a dynamic problem trading-off search costs with expected gains in surplus: markups are larger for those firms characterized by lower expected future prices. Moreover, the entire path of profits matters for optimal markups because the firm's pricing decision is a dynamic problem trading-off current and future expected profits: markups are larger for those firms characterized by lower expected future profits relatively to current profits. Given that the path of prices and profits is determined by the path of productivity, our mechanism exacerbates the dependence of markups on productivity. Equilibrium markups depend both on the cross-sectional dispersion as well as on the dynamics of productivity over time.

In our model the attenuated response of prices to cost shocks is delivered without needing to assume nominal price rigidities. Firms are reluctant to change prices since consumers have to pay a cost to search for a new provider, which makes them less keen on moving. Therefore, it is hard for firms to regain lost business. In this sense, our paper links to a classic literature trying to explain price stickiness purely through real rigidities, for example assuming a kinked demand function ([Sweezy \(1939\)](#)). However, in our setup, the rigidity generating price stickiness arises from first principles, whereas in the kinked demand literature it is assumed into the model.

Our paper ties into a growing body of literature of quantifiable macroeconomic models where the market share of the firm is sluggish, such as [Ravn et al. \(2006\)](#), [Alessandria \(2009\)](#), [Kleshchelski and Vincent \(2009\)](#), [Gourio and Rudanko \(2011\)](#), [Shi \(2011\)](#), or [Drozd and Nosal \(2012\)](#). Our contribution to this literature is twofold. First, we point out the link between customer base dynamics and asymmetry and heterogeneity of pass-through. Second, we put together micro data and model to quantify the importance of this channel.

We also add to the literature using scanner data to document empirical regularities in

¹Another alternative is assuming a linear demand system with horizontal product differentiation, as in [Melitz and Ottaviano \(2008\)](#).

pricing and shopping behavior. [Coibion et al. \(2012\)](#) integrates store and customer scanner data to show that customers search for lower prices as a result of income shocks. A similar result is found in [Kaplan and Menzio \(2013\)](#). We view our contribution as complementary to these papers: they explore the elasticity of search to income whereas we focus on the elasticity of search to prices.

Finally, our paper relates to [Coles and Mortensen \(2011\)](#), which studies a dynamic equilibrium where workers search for firms, and firms post wages in order to attract these workers. While the application is clearly different, the search models in the two papers share several features and predictions.

2 The model

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

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 $\pi(p, z)$ the profit function. We assume that $\pi(p, z)$ is continuously differentiable in p and 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 $\mu(z)$ denote the unconditional density of productivity z . Also let $\mu(z'|z)$ denote the conditional density function of z' given z . The productivity shock is persistent in the sense that the cumulative distribution function, $\int_{\underline{z}}^s \mu(z'|z) dz'$, is weakly decreasing in z .

Customers. We use the superscript i , $i \in [0, \Gamma]$, to index a customer. Let I denote the income level of a customer each period. Let $d(p)$, $v(p)$ denote the static demand and customer surplus functions which, as in standard models, only depends 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 will face different realizations of both demand

and surplus given they will face different prices p^i , and these different realizations occur 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.

Search and matching. Customers draw a random search cost $\psi \in [\underline{\psi}, \infty)$, drawn randomly each period from the same distribution $G(\psi)$, where $g(\psi) \equiv \partial G(\psi)/\partial \psi$; we restrict our attention to density functions that are continuous in $(\underline{\psi}, \infty)$. In the remaining of the paper we assume that $\underline{\psi} = 0$. If the customer incurs in the utility cost ψ , she exits the customer base of her current firm and is randomly matched to a new one, where the probability of matching with a particular type of firm is proportional to its customer base size. Matching occurs instantaneously and we allow customers to search at most once per period. Finally, we assume no recall of previously visited firms.

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 beginning of period t . 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 $\{p_t^j, z_t^j, m_{t-1}^j\}$, (iii) each customer decides between searching for a new firm -after incurring in search cost ψ_t^i - and being assigned randomly within period t , or remaining matched to firm j -and not incurring in cost ψ_t^i -, and (iv) customer surplus $v(p_t^j)$ and profits $\pi(p_t^j, z_t^j)$ are realized. We refer to (iii) as the extensive margin of demand, where each customer decides whether to remain in the customer base of firm j or exit, while we refer to (iv) as the intensive margin of demand.

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 . Let $\tilde{\mathcal{P}}(m, z) : [0, \infty) \times [z, \bar{z}] \rightarrow [0, \infty)$ be the pricing strategy of the firm mapping m and z to the space of prices. Also, let $\tilde{\chi}(p, z, m, \psi) : [0, \infty) \times [z, \bar{z}] \times [0, \infty) \times [0, \infty) \rightarrow \{0, 1\}$ be the exit strategy of the customer mapping the price p , the productivity z , the customer base m , as well as the search cost ψ

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

to $\{0, 1\}$, where $\tilde{\chi} = 1$ denotes the case of exiting the customer base of the firm.³

In our setup, as we will show later, the customer base enters multiplicatively 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. Therefore, $\mathcal{P}(z) : [z, \bar{z}] \rightarrow [0, \infty)$ denotes the pricing strategy of the firm mapping z to the space of prices. Also, $\chi(p, z, \psi) : [0, \infty) \times [z, \bar{z}] \times [0, \infty) \rightarrow \{0, 1\}$ denotes the exit strategy of the customer mapping z , p and ψ to $\{0, 1\}$.

2.2 The problem of a customer

Consider a customer buying good 1 from a given firm j . Let $\bar{V}(p, z)$ be the value function for a customer who decided to stay and buy from firm j , paying a price p for good 1 in current period, and facing a pricing policy $\mathcal{P}(\cdot)$ for future states. We have

$$\bar{V}(p, z) = v(p) + \beta \int_z^{\bar{z}} \left[\int_0^\infty \max \left\{ \bar{V}(\mathcal{P}(z'), z'), \hat{V} - \psi \right\} g(\psi) d\psi \right] \mu(z'|z) dz', \quad (1)$$

where \hat{V} is the outside option to the customer. Consider the search decision of a customer who draws a search cost ψ . The exit strategy of the customer is $\chi(z, p, \psi) = 1$ if $\bar{V}(p, z) < \hat{V} - \psi$, and $\chi(z, p, \psi) = 0$ otherwise. Because the value of exiting is decreasing in ψ and the value of not exiting does not depend on ψ , the exit strategy takes the form of a trigger, $\bar{\psi}$, such that the customer exits if $\psi < \bar{\psi}$. We next characterize the trigger policy.

Let $\bar{\mathcal{P}}(z)$ denote the function that solves $\bar{V}(\bar{\mathcal{P}}(z), z) = \hat{V}$, so that no customer decides to exit the customer base of firm j with productivity z if $p \leq \bar{\mathcal{P}}(z)$.⁴ Using [equation \(1\)](#) we get that

$$\bar{\mathcal{P}}(z) = v^{-1} \left(\hat{V} - \beta \int_z^{\bar{z}} \left[\int_0^\infty \max \left\{ \bar{V}(\mathcal{P}(z'), z'), \hat{V} - \psi \right\} g(\psi) d\psi \right] \mu(z'|z) dz' \right). \quad (2)$$

Notice that the variable $\bar{\mathcal{P}}(z)$ ranks the different firms according to the continuation value for the customer: when comparing two firms selling the good at the same current price p ,

³As noted in [Coles \(2001\)](#), supporting this type of equilibria requires imposing assumptions on off-equilibrium paths. To see this notice that in equilibrium firms are supposed to choose the same price whenever the productivity is the same. However, because future prices are taken as given in the current period, a firm has incentives to deviate from this price marginally and customers should not react by exiting the customer base of the firm as this small price change only lasts for one period. To avoid this type of behavior we assume that if customers observe a price which is not consistent with the equilibrium price they exist the customer base of the firm.

⁴The existence of $\bar{\mathcal{P}}(z)$ is a direct implication of $\bar{V}(p, z)$ being strictly decreasing in p . Because $v(p)$ is strictly decreasing in p , a direct application of the Contraction Mapping Theorem implies that $\bar{V}(p, z)$ is a strictly decreasing in p . The fact that no customer decides to exit if $p \leq \bar{\mathcal{P}}(z)$ follows immediately.

the customer prefers the firm characterized by higher $\bar{P}(z)$. Likewise, notice that $\bar{P}(z)$ is the highest price the firm can charge so that no customer searches for a new firm.

Next we define the threshold for the search cost ψ below which customers decide to exit the customer base of a given firm. Let $\bar{\psi}(p, z)$ denote the threshold rule such that a customer decides to exit if $\psi \leq \bar{\psi}(p, z)$, so that we have

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

and therefore the exit strategy is

$$\chi(p, z, \psi) = \begin{cases} 1 & \text{if } \psi \leq \bar{\psi}(p, z) , \\ 0 & \text{otherwise} . \end{cases}$$

The next lemma characterizes the threshold rule $\bar{\psi}(p, z)$.

Lemma 1 *For all $p > \bar{P}(z)$, the threshold rule $\bar{\psi}(p, z)$ is strictly increasing in p . Moreover, (i) if $\bar{P}(z) = \bar{P}$ then $\bar{\psi}(p, z) = \bar{\psi}(p)$, (ii) if $\bar{P}(z)$ is increasing in z then $\bar{\psi}(p, z)$ is increasing in z .*

The proof of **Lemma 1** follows from $v(p)$ being strictly decreasing in p . The first part of the Lemma shows that the fraction of customers exiting the base of a given firm increases with its price p . Moreover, it implies that, when comparing two firms charging the same price p in the current period, a larger fraction of customers will exit the customer base of the firm with lower value of $\bar{P}(z)$.

2.3 The problem of the firm

In **Section 2.2** we concluded that a customer exits the customer base of the firm she currently patronizes when the price of the good is high enough; we described this optimal behavior by the function $\bar{\psi}(p, z)$. 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} denotes the mass of customers at the beginning of a period currently buying from firm j . Also, let δ denote the arrival rate of new customers. The evolution of the customer base is

$$m_t = m_{t-1} [1 + \delta - G(\bar{\psi}(p, z))] \equiv m_{t-1} \Delta(p, z) , \quad (4)$$

where $z = z_t^j$ and $p = p_t^j$. The term $m_{t-1}\delta$ accounts for the inflow of new customers, while the term $m_{t-1}(1 - G(\bar{\psi}(p, z)))$ accounts for the mass of customers that decided to stay with the

firm. In the end, $\Delta(p, z)$ measures the customer base growth of the firm. The linearity of the production function, coupled with the arrival rate of new customers being proportional to the customer base size of the firm, implies Gibrat's Law. That is, firm's growth is independent of its size (Luttmer (2010)).

Let

$$\varepsilon_m(p, z) \equiv -\frac{p}{\Delta(p, z)} \frac{\partial \Delta(p, z)}{\partial p} = -v'(p)p \frac{G'(\bar{\psi}(p, z))}{1 - G(\bar{\psi}(p, z)) + \delta} \geq 0, \quad (5)$$

denote the elasticity of the growth rate of the customer base to prices, where the last equality follows from equation (3) and the definition of $\Delta(p, z)$. We refer to this object as the *extensive* margin demand elasticity as it governs the amount of customers within the firm's customer base. Likewise, we refer to $\varepsilon_d(p)$ as the *intensive* margin demand elasticity as it controls the expenditure of those customers who decided to stay. Notice that aggregate quantity sold by firm j is $q_t(p, z) \equiv m_{t-1} \Delta(p, z) d(p)$ so that overall demand elasticity is

$$\varepsilon_q(p, z) = \varepsilon_d(p) + \varepsilon_m(p, z) .$$

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

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

from where it is immediate to see that $\pi(p, z)$ is strictly increasing and strictly concave in z . Given the assumptions made regarding the demand function $d(p)$, the profit function is continuously differentiable in p . Moreover, we assume that the profit function is strictly concave, so that it attains its unique maximum at a unique value, $\hat{\mathcal{P}}(z)$. This value denotes the price that maximizes static profits for a firm with current productivity z , i.e. $\partial \pi(\hat{\mathcal{P}}(z), z) / \partial p = 0$. We have that

$$\hat{\mathcal{P}}(z) = \frac{\varepsilon_d(p)}{\varepsilon_d(p) - 1} \frac{w}{z}. \quad (7)$$

In Appendix A.1 we show that concavity of the profit function with respect to p , and the existence of a unique maximizer $\hat{\mathcal{P}}(z)$ follows when the utility function of customers is CRRA with CES demand across a large set of varieties.

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

$$\tilde{F}(z_t, m_{t-1}^j) = \max_p m_t^j \pi(p, z_t) + \beta \int_z^{\bar{z}} \tilde{F}(z_{t+1}, m_t^j) \mu(z_{t+1} | z_t) dz_{t+1},$$

An application of the Contraction Mapping Theorem implies that the value function for a firm is homogeneous of degree one in m , i.e. $\tilde{F}(z, m) = m \tilde{F}(z, 1) \equiv m F(z)$, where $F(z)$ solves

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

where the profit per customer function $\pi(p, z)$ satisfies [equation \(6\)](#) and we used [equation \(4\)](#). Also, to ease on notation, we let $z = z_t$ and $z' = z_{t+1}$. Notice that, because the size of customer base does not appear in the problem presented in [equation \(8\)](#), two firms with 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 $\hat{F}(z)$ 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 $\beta(1 + \delta) < 1$; this restriction guarantees that the problem is a contraction. 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 sufficient for the optimal price schedule; this stems from $\bar{\psi}(p, z)$, and its effect on the effective discount rate and period payoff.

Let $\Pi(p, z) \equiv \pi(p, z) + \hat{\beta} \int_{\underline{z}}^{\bar{z}} F(z') \mu(z'|z) dz'$ denote the value of a customer for the firm. 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, $\hat{\beta} \int_{\underline{z}}^{\bar{z}} F(z') \mu(z'|z) dz'$, accounts for the value that the customer brings to the firm in the future. Rewriting [equation \(8\)](#) we get that

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

which shows that the firm's optimal price trades-off size of the customer base $\Delta(p, z)$, which is decreasing in p , with the value per customer $\Pi(p, z)$, which attains its maximum at $\hat{\mathcal{P}}(z)$. Moreover, $\Delta(p, z)$ is regulated by the extensive margin demand elasticity $\varepsilon_m(p, z)$ while $\Pi(p, z)$ is regulated by the intensive margin demand elasticity $\varepsilon_d(p)$.

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

Proposition 1 *Let $\tilde{\mathcal{P}}(z)$ denote the set of prices solving the following expression,*

$$\frac{\partial \pi(p, z)}{\partial p} \Delta(p, z) = -\frac{\partial \Delta(p, z)}{\partial p} \Pi(p, z). \quad (9)$$

Then $\mathcal{P}(z) = \tilde{\mathcal{P}}(z) = \{\hat{\mathcal{P}}(z)\}$ if and only if $\hat{\mathcal{P}}(z) \leq \bar{\mathcal{P}}(z)$. If, instead, $\hat{\mathcal{P}}(z) > \bar{\mathcal{P}}(z)$, then $\mathcal{P}(z) = \tilde{\mathcal{P}}(z) \subset [\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$. Moreover, if $\partial\Delta(p, z)/\partial p$ is continuous and $\bar{\mathcal{P}}(z) < \hat{\mathcal{P}}(z)$, then $\mathcal{P}(z) = \tilde{\mathcal{P}}(z)$ and $\tilde{\mathcal{P}}(z) \subset (\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$.

A proof of the proposition can be found in [Appendix A.2](#). The next lemma gives conditions under which the first order condition in [equation \(9\)](#) is necessary and sufficient for an optimum.

Lemma 2 *If $g(\psi) = 0$, $\Delta(p, z)$ is continuously differentiable, and the first order condition in [equation \(9\)](#) is necessary for an optimum. If furthermore, the hazard rate of ψ is increasing and $\delta \rightarrow 0$, then $\varepsilon_m(p, z)$ is increasing in p , and $\tilde{\mathcal{P}}(z)$ is a singleton.*

From now on we consider the case where the assumptions of [Lemma 2](#) are satisfied. Notice that productivity z has two distinct effects on prices in our model. An important result of [Proposition 1](#) is that the optimal price of a firm is never above the price that maximizes per-customer static profits, $\hat{\mathcal{P}}(z)$. First, productivity has the standard *direct effect* 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 through $\bar{\mathcal{P}}(z)$: in equilibrium $\bar{\mathcal{P}}(z)$ depends on current productivity of a given firm, as it 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, if $\bar{\mathcal{P}}(z)$ increases with productivity z , then $\mathcal{P}(z)$ is not necessarily a monotonic function of current productivity level as more productive firms might face a less elastic 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. This is possible in those cases where customers value more firms with higher productivity, i.e. $\bar{\mathcal{P}}(z)$ increasing in z , and the sensitivity in the growth in the customer base to prices decreases fast when $\bar{\mathcal{P}}(z)$ increases.

The optimal price $\mathcal{P}(z)$ solves [equation \(9\)](#). Using the demand elasticities definitions we can rewrite this expression as,

$$\frac{p}{w/z} = \frac{\varepsilon_d(p) + \alpha_\pi(p, z)\varepsilon_m(p, z)}{\varepsilon_d(p) - 1 + \alpha_\pi(p, z)\varepsilon_m(p, z)}, \quad (10)$$

where $\alpha_\pi(p, z) \equiv \Pi(p, z)/\pi(p, z) \in (1, \infty)$. The equation shows that the optimal markup, $p/(w/z)$, 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 her static value, $\alpha_\pi(p, z)$. As a result, this expression

shows that the optimal price trades-off short term profits vs. long term value accrued from the marginal customer. The next proposition presents a set of comparative statics.

Proposition 2 *The optimal price of a firm is (i) decreasing in the intensive margin elasticity $\varepsilon_d(p)$, (ii) decreasing in the extensive margin elasticity $\varepsilon_m(p, z)$, (iii) decreasing in $\alpha_\pi(p, z)$, and (iv) $p = \frac{\varepsilon_d(p)}{\varepsilon_d(p)-1} \frac{w}{z} = \hat{\mathcal{P}}(z)$ if and only if $\varepsilon_m(p, z) = 0$ and $p < \frac{\varepsilon_d(p)}{\varepsilon_d(p)-1} \frac{w}{z} = \hat{\mathcal{P}}(z)$ if $\varepsilon_m(p, z) > 0$. Finally, (v) for 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 price.*

The proof of the proposition is straightforward and therefore omitted. Part (i) and (ii) of the proposition simply states that the higher either the intensive or extensive margin elasticities, the lower the price charged by a firm. Part (iii) of the proposition states that the optimal price decreases with the relative importance of the long term value of a customer relative to the short term value, which is mostly governed by the discount factor β . Part (iv) states 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 (v) of the proposition states that the larger the extensive margin elasticity share of the aggregate demand elasticity, then the lower optimal price.

An immediate result is that if a firm with productivity z choosing the price that maximizes static profits $\hat{\mathcal{P}}(z)$ faces no customer base concerns, then the optimal price for this firm is $\hat{\mathcal{P}}(z)$.

As discussed, the extensive demand margin elasticity $\varepsilon_m(p, z)$ plays a crucial role in the optimal price determination. Therefore, understanding what shapes up this object is useful. We already established that $\varepsilon_m(p, z) \geq 0$, and discussed conditions under which $\varepsilon_m(p, z)$ is increasing in p . The next lemma discusses conditions under which $\varepsilon_m(p, z)$ is increasing in z .

Lemma 3 *If the assumptions of Lemma 2 are satisfied, and $\partial \bar{\psi}(p, z) / \partial z \leq 0$, then $\varepsilon_m(p, z)$ is decreasing in z .*

The proof of the lemma is straightforward and therefore omitted. The lemma states that if the threshold rule $\bar{\psi}(p, z)$ decreases with z , then the extensive margin elasticity is decreasing in the productivity level z . Notice that $\bar{\psi}(p, z)$ decreases with z when customers expect firms with higher productivity today to charge lower prices in the future and, as a result, under this assumption the extensive margin elasticity falls as productivity increases. An example is when optimal prices are monotonic in productivity, and productivity has some

persistence. Finally, notice that both derivatives of $\varepsilon_m(p, z)$ with respect to p and z crucially depend on the hazard rate of ψ being increasing. This assumption implies that, in relative terms, the density of customers exiting to the fraction remaining in the customer base of the firm increases with $\bar{\psi}$.

2.4 Equilibrium

In this section we define a stationary equilibrium and we discuss some of its general properties. Let $\mathcal{P}^*(z) \subseteq \mathcal{P}(z)$ denote the candidate equilibrium price at productivity level z . Let $M(z)$ measure the amount of customers currently buying good 1 from firms with productivity z . The endogenous inflow of customers per period δ is such that the measure $M(z)$ is stationary and satisfies the following equations:

$$M(z') = \int_{\underline{z}}^{\bar{z}} M(z) \Delta(\mathcal{P}^*(z), z) \mu(z'|z) dz , \quad (11)$$

$$\Gamma = \int_{\underline{z}}^{\bar{z}} M(z) dz . \quad (12)$$

Definition 1 *A stationary equilibrium is (i) a pricing policy $\mathcal{P}^*(z)$, (ii) an invariant distribution $M(z)$ and arrival rate δ , (iv) a function for the threshold rule $\bar{\psi}(\cdot, \cdot)$, (v) the customer outside option satisfying $\hat{V} = \int_{\underline{z}}^{\bar{z}} \bar{V}(\mathcal{P}^*(z), z) \mu(z) dz$, customers search optimally (equation (3)), firm's optimal pricing solves equation (8), and dynamics of customer base satisfy equations (11).*

We now start discussing equilibrium characterization, which we will do in more detail in [Section 2.5](#). To this end, let \mathcal{Z}_1 denote the set of productivity levels at which the firm's optimal price maximizes static profits and let \mathcal{Z}_2 denote the set of productivity levels at which the firm's optimal price is below the price that maximizes static profits. That is, $\mathcal{Z}_1 = \{z \text{ such that } \mathcal{P}^*(z) = \hat{\mathcal{P}}(z)\}$ and $\mathcal{Z}_2 = \{z \text{ such that } \mathcal{P}^*(z) < \hat{\mathcal{P}}(z)\}$.

Proposition 3 *In equilibrium we have (i) $\mathcal{P}^*(z) \leq \hat{\mathcal{P}}(z)$ for all z , and (ii) $\mathcal{Z}_1 \neq \emptyset$ and $\mathcal{Z}_2 \neq \emptyset$. Then, for any two productivity levels $\{z, y\}$, $z \in \mathcal{Z}_1$ and $y \in \mathcal{Z}_2$, we have that $\Delta(\mathcal{P}^*(z), z) = 1 + \delta > \Delta(\mathcal{P}^*(y), y) > \delta$, and for some y , $\Delta(\mathcal{P}^*(y), y) < 1$.*

A proof of the proposition can be found in [Appendix A.3](#). Part (i) of the proposition states that no firm prices above the price that maximizes static profits; this is an immediate implication of [Proposition 1](#). Part (ii) of the proposition shows that there is a set of productivity levels where the optimal price maximizes static profits; this happens because for

productivity levels lying in the set \mathcal{Z} , customer retention concerns are not present and as a result $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z) < \bar{\mathcal{P}}(z)$. It also shows that there is a set of productivity levels where the optimal price is strictly below the price that maximizes static profits. The latter follows from the assumption that $\underline{\psi} = 0$, so that price dispersion necessarily implies that some customers will exit the customer base of firms with highest prices. As a result, relative to the standard model of price setting under monopolistic competition where the optimal price is $\hat{\mathcal{P}}(z)$, the presence of customer retention concerns generate lower average prices and markups.

Proposition 4 *If the assumptions of Lemma 2 are satisfied, an equilibrium exists.*

A proof of the proposition can be found in [Appendix A.4](#). The assumptions of [Lemma 2](#) ensures that the firm's maximization problem has a unique solution, and that the optimal price is a continuous function of productivity. The proof follows by applying Brouwer fixed point theorem.

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 $\underline{\psi} \rightarrow \infty$. 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 search cost is divergent, i.e. $\underline{\psi} \rightarrow \infty$. Then, in equilibrium: (i) the optimal price maximizes static profits, $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$ for all z , and (ii) customers do not find it optimal to exit the customer base of their current firm.*

The proof of the Remark follows immediately since, because the search cost diverges, firms do not face customer retention concerns and therefore find it optimal to charge the price that maximizes static profits. Formally, when $\underline{\psi} \rightarrow \infty$, we have that $\bar{\mathcal{P}}(z) \rightarrow \infty$. It is immediate that $\hat{\mathcal{P}}(z) \leq \bar{\mathcal{P}}(z)$ so that $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$.

Remark 2 *Let $\underline{z} = \bar{z} = z_0$, and $\mu(z_0) = 1$. Then, $\mathcal{P}^*(z) = \bar{\mathcal{P}}(z) = \hat{\mathcal{P}}(z)$. As a result, $\Delta(\mathcal{P}^*(\bar{z}), z) = 1$ so that firms enjoy a constant customer base. Moreover, if $G(0) = 1$, the equilibrium price is $\mathcal{P}^*(\bar{z}) \in [W/\bar{z}, \hat{\mathcal{P}}(\bar{z})]$.*

A proof of the Remark can be found in [Appendix A.5](#). 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.

The next proposition discusses the equilibrium properties when optimal prices are monotonic in productivity.

Proposition 5 *If $\mathcal{P}^*(z)$ decreases monotonically with z the equilibrium is such that $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$ for all $z \geq \hat{z}$, and $\mathcal{P}^*(z) < \hat{\mathcal{P}}(z)$ for all $z < \hat{z}$, where $\hat{z} \in (\underline{z}, \bar{z})$ is unique.*

A proof can be found in [Appendix A.6](#). A special case where the assumptions of the Proposition are satisfied is when productivity shocks are i.i.d..

In the next section we provide examples where the assumptions of [Proposition 4](#) are satisfied, and explore the implications of our mechanism for markups.

2.5 Equilibrium prices when productivity follows a simple process

In this section we explore some properties of the model using a simple productivity process. The process we consider is the following

$$z' = \begin{cases} z & \text{with probability } \rho \\ \varepsilon & \text{with probability } 1 - \rho \end{cases} \quad (13)$$

where ε is a random draw from $\mu(\varepsilon)$. This process encompasses the case of i.i.d. productivity shocks, when $\rho = 0$, and entails some persistence when $\rho > 0$. The next proposition characterizes the properties of optimal prices under [equation \(13\)](#).

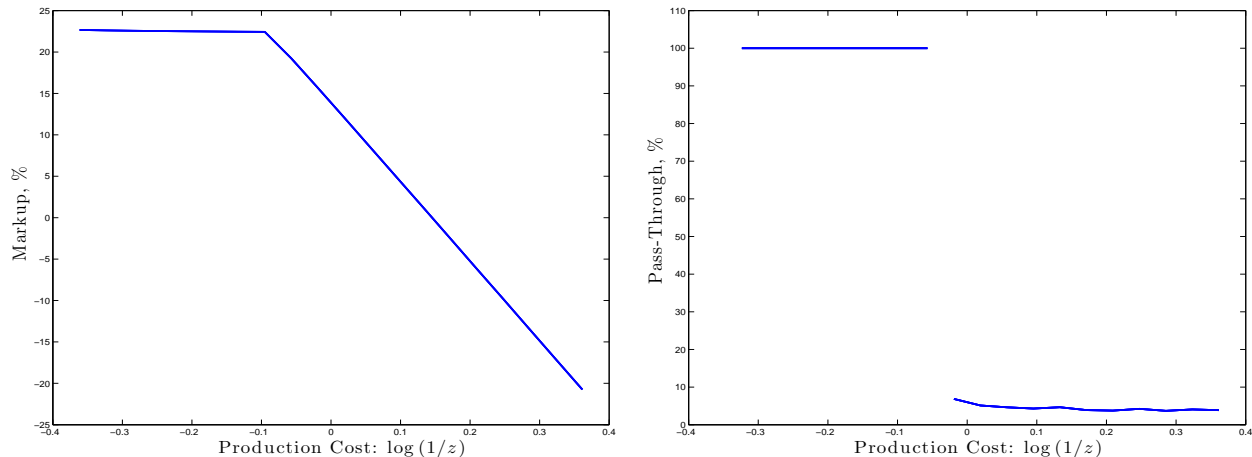
Proposition 6 *Let z follow the process given by [equation \(13\)](#), then the equilibrium prices are such that: (i) $\mathcal{P}^*(z)$ decreases monotonically with productivity z ; (ii) $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$ for all $z \geq \hat{z}$, and $\mathcal{P}^*(z) < \hat{\mathcal{P}}(z)$ for all $z < \hat{z}$; (iii) markups, i.e. $\mathcal{P}^*(z)/(w/z)$, increase with z .*

Point (i) stems from the “promised” continuation value to customers being necessarily non-decreasing in the current productivity z , given the process for z . Point (ii) follows by applying [Proposition 4](#). Point (iii) follows because monotonicity of equilibrium prices guarantees that $\varepsilon_m(\mathcal{P}^*(z), z)$ and $\varepsilon_d(\mathcal{P}^*(z))$ decrease with z , and the process in [equation \(13\)](#) makes $\alpha_\pi(\mathcal{P}^*(z), z)$ inversely proportional to equilibrium profits $\pi(\mathcal{P}^*(z), z)$, and thus decreasing with z . Then by using [equation \(10\)](#) (iii) follows. For a detailed proof see [Appendix A.7](#).

[Figure 1](#) illustrates the properties of equilibrium markups and pass-through for a reasonable parameterization of the model in the case of i.i.d. productivity shocks.⁵ We can group firms in two categories: highly productive firms (i.e. $z > \hat{z}$) fully pass-through shocks to production costs, while low productivity firms pass-through only a small fraction. Quantitatively, the difference in the pass-through depends on the scale of the search costs, with larger search costs inducing larger differences in the pass-through across the two groups.

⁵We consider an economy with customers having CES preferences over N varieties, each variety produced in a different industry populated by a continuum of producers differing in their productivity level. Thus, the model we study applies to each of these industries. We set $N = 100$, so that $\varepsilon_d(p)$ is roughly constant. See [Appendix A.1](#) for details.

Figure 1: Equilibrium Markups and Pass-Through with i.i.d. productivity shocks



Note: Equilibrium markups and pass-through are given by [equation \(10\)](#), where $\varepsilon_d(p) = 5$ for all p , while $\varepsilon_m(p, z)$ and $\alpha_\pi(p, z)$ are solved for in a model where z is given by [equation \(13\)](#) with $\rho = 0$, and the functions $v(p)$ and $\pi(p, z)$, as well as the distribution of search costs $G(\psi)$ are chosen as described in [Section 4](#).

Finally, it is interesting to study the role of ρ for optimal markups. When productivity shocks are i.i.d. $\varepsilon_m(p, z)$ does not depend on z directly but only through the contemporaneous price p . The direct effect of z on $\varepsilon_m(p, z)$ amplifies the impact of variation in productivity onto variation in equilibrium markups. The larger the persistence ρ , the more important this channel.

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, as the model would predict, fluctuations in the price of the good sold by a firm influence a customer decision to remain in its customer base. Next, we will 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 two years in the sample. 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 (i.e. the household stops shopping at the chain), and (ii) the price of the household basket. Here we briefly describe the procedure followed to obtain them, further details are included in [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. When a particular household has not shopped at the chain for eight or more consecutive weeks, we deduce that she has left the customer base. The 8-weeks window is a conservative choice. Every household in the data holds a loyalty card which suggests that they are not casual shoppers at the chain. We made sure of that by selecting only households shopping there at least 48 times over the two years in the sample. Regular customers are unlikely to experience that long of a 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 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 (p).

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

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

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 control for observable characteristics through inclusion of household’s demographics matched from Census 2000. 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.

The regressions aim at testing the existence of a correlation between the decision of exiting the customer base of the chain and the price of the basket at the chain, relative to what that price would be at the competition. However, since we are using price data from a single chain, we cannot exclude that some of the variation in p_{it} comes 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). Aggregate cost shocks do not change the relative price and, therefore, should not trigger exit from the customer base. Failing to remove aggregate shocks should not affect the sign of the correlation but add noise to it, biasing against finding a significant result. Moreover retailer specific shocks (i.e. a shift in the cost of all the goods sold by the chain providing the data), despite being a valid source of identification, do not contribute to it in our setting. In fact, since we only use data from one chain, their effect is absorbed by the time dummies. The correlation of interest is then identified by good-retailer specific shocks as those triggered, for example, by the expiration of a contract between the chain and a manufacturer. The availability of data from several stores helps this identification strategy by introducing 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).

Unlike in standard demand analysis we are not modeling the household’s choice of the retailer where to shop but rather the decision of whether to leave a given retailer or to stick with it. Endogeneity driven by unobserved store characteristics correlated with both price and the propensity to leave is not an issue in this context. 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 equation 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, unrelated to productivity shocks, change in prices. Furthermore, the decision to exit the customer base is also influenced by the productivity level z , which is customer-time specific. Because we do not include a measure of productivity in the specification, the price coefficient should be regarded as averages across productivity levels, which follows from a stationary interpretation of the regression.

Column (1) of [Table 1](#) documents that a 1% increase in the growth rate of individual specific weekly prices is associated with 0.012 percentage points increase in the probability that the household leaves the chain to patronize a rival grocer. In the analysis we have so far abstracted from the behavior of the competitors of the chain. In part this is justified by the fact that in the theoretical framework we introduce in [section 2](#) we assume that search is indirect. Therefore, pricing and promotion policies of competitors should not affect the customer’s decision of leaving the chain. Indeed, under the lens of the model only price shocks idiosyncratic to the chain should affect the decision of leaving implying that aggregate shock can only bias us against finding a significant price effect. Furthermore, in order to believe that this omission is driving our result we should assume that the retailer systematically raises its prices when competition increases rather than lowering them. Nevertheless, we address the issue of competition in the last two columns of [Table 1](#). The specification in column 2 includes the number of competing supermarket stores, calculated using information from Reference

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

Exiting: Missing at least 8 consecutive weeks			
	(1)	(2)	(3)
$\log(P_i)$	0.012* (0.007)	0.014* (0.007)	0.013* (0.007)
Num. competitors		-0.000 (0.001)	
Walmart entry			0.041 (0.044)
Observations	77,815	71,049	66,182

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 trim 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%.

US, in the zipcode of residence of a household. In column 3 we use data from [Holmes \(2011\)](#) to construct a dummy signaling the opening of a Walmart super-center in the zip-code of the store where the customer is shopping. In both cases the main result does not change. We also assess whether our finding is robust to defining exit from the customer base in a more restrictive fashion, only considering spells of eight or more consecutive weeks without shopping at the chain. This alternative definition reduces the number of exits observed and leads to a lower but still significant effect of prices on the decision to change supplier.

4 Calibration

In this section we discuss how we parameterize the model. Some parameters of the model we borrow them directly from external sources, while we estimate other by a method of moments using our model.

We first choose preferences that microfound the demand and surplus functions, $d(p)$ and $v(p)$. Let customers derive utility from a composite of two goods, d and n . We label the composite good by c , where $c = \left(\omega d^{\frac{\theta-1}{\theta}} + (1 - \omega) n^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}$, $\theta > 1$. We let the utility

function of customers being $u(c) = c^{1-\gamma}/(1-\gamma)$. We interpret good d as the good produced by the firms described above while good n is produced by a competitive representative firm with linear production and unitary labor productivity. We let n act as a numeraire. The budget constraint of a customer is $pd + n = I$. Details on the derivation of $d(p)$ and $v(p)$ can be found in [Appendix A.1](#).

Income comes from labor proceeds and profits. Profits are divided equally across the different customers, and we assume that each customer supplies an equal amount of hours ℓ . We determine ℓ in [Appendix A.8](#) by solving the maximization problem of customers facing disutility from working. Given perfect competition in the production of good n , and given the linear production technology, the equilibrium wage, w , has to equal marginal product of labor in such sector, and thus $w = 1$.

Imputed parameters. We set the time period to be one month and we set the discount rate to be $\beta = 0.997$ (i.e. an yearly interest rate of 4%). We set the relative risk aversion parameter to $\gamma = 2$ which is a standard value. We set $\omega = 0.5$ so that goods d and n have the same weight in preferences.⁶

As it is well known prices do not adjust every period. To address this fact we use the process presented in [equation \(13\)](#), with $\varepsilon \sim N(0, \sigma)$. This process is mean reverting, where $1 - \rho$ is the probability of mean reversion. Given that in our model the equilibrium price depends solely on the productivity z , the probability that a firm changes her price is also $1 - \rho$. [Goldberg and Hellerstein \(2009\)](#) provides, for the PPI (i.e. Producer Price Index) a measure of the average monthly frequency of price changes (0.32). Interpreting these prices as the cost of production for consumption goods allows us to calibrate the parameter ρ , obtaining $\rho = 0.68$.

Parameters requiring calibration. We assume that the search cost is drawn from a Gamma distribution with shape parameter ζ , and scale parameter λ . We have to obtain estimates for these two parameters, as well as for the parameter governing the elasticity of substitution across different types of goods, θ , and the volatility of innovations in productivity, σ .

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 (the parameter b_1 in [equation \(14\)](#)) is informative about the level of the search costs and thus we use it to identify the scale parameter λ . 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

⁶In the economic application that we will perform in the next section, we will interpret the numeraire good n as a non-tradable, so that $\omega = 0.5$ will roughly correspond to an average share of non-tradables in the consumption basket of 50%.

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 the difference between the average marginal effect between the top and the bottom terciles measured in the data and in the model.

We select volatility of innovations in productivity, σ , so that the implied price dispersion in the model matches that reported by [Kaplan and Menzio \(2013\)](#) for bundles of homogeneous packaged goods. Finally, we choose the parameter θ so that the average intensive margin elasticity is 5, a value in the range of those used in the macro literature.⁷

We define $\Omega \equiv [\zeta \ \lambda \ \theta \ \sigma]'$ as the vector of parameters to be estimated, and denote 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
Volatility of productivity innovations, σ	0.2	Price dispersion: 8%
Elasticity of substitution, θ	6.72	Avg. intensive margin elasticity: 5
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	0.4	Average marginal effect: 1.4%

5 Quantitative evaluation of markups and pass-through

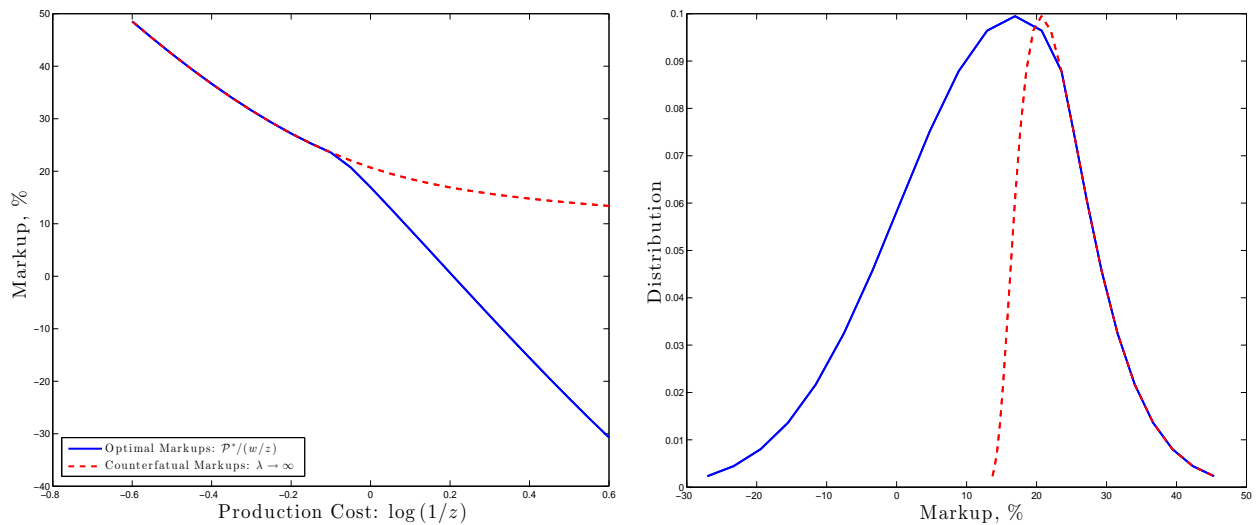
In this section we explore the properties of equilibrium markups, and how they vary as a function of idiosyncratic shocks to costs, at the stationary distribution of our economy. The exercise we perform consists in taking estimates of our model with customer base concerns for the distribution of markups and pass-through, and compare them to the counterpart statistics obtained from a model where we shut down the customer retention mechanism. In particular, we let search costs to diverge to infinity (i.e., $\lambda \rightarrow \infty$) so that customers would never find it optimal to search for a new firm. In this economy with $\lambda \rightarrow \infty$, we set θ so the resulting average elasticity of demand, $\mathbb{E}[\varepsilon_d]$, is the same of our baseline economy with active

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

customer retention concerns. Notice that this framework is analogous to the standard CES preferences model.

The left panel of **Figure 2** compares the equilibrium markups in our economy with the equilibrium markups in the CES economy. Two main features emerge from looking at the figure. First, high productivity firms, i.e. with productivity roughly above the mean, charge the same markups in our economy and in the CES economy. Low productivity firms, i.e. with productivity roughly below the mean, charge strictly lower markups in our economy with customer retention than in the CES economy where they are absent by assumption. Second, markups are strictly decreasing in production cost, both in a model with and without customer retention concerns. For low productivity firms, markups decrease faster in our economy than in the CES economy.

Figure 2: Equilibrium Markups

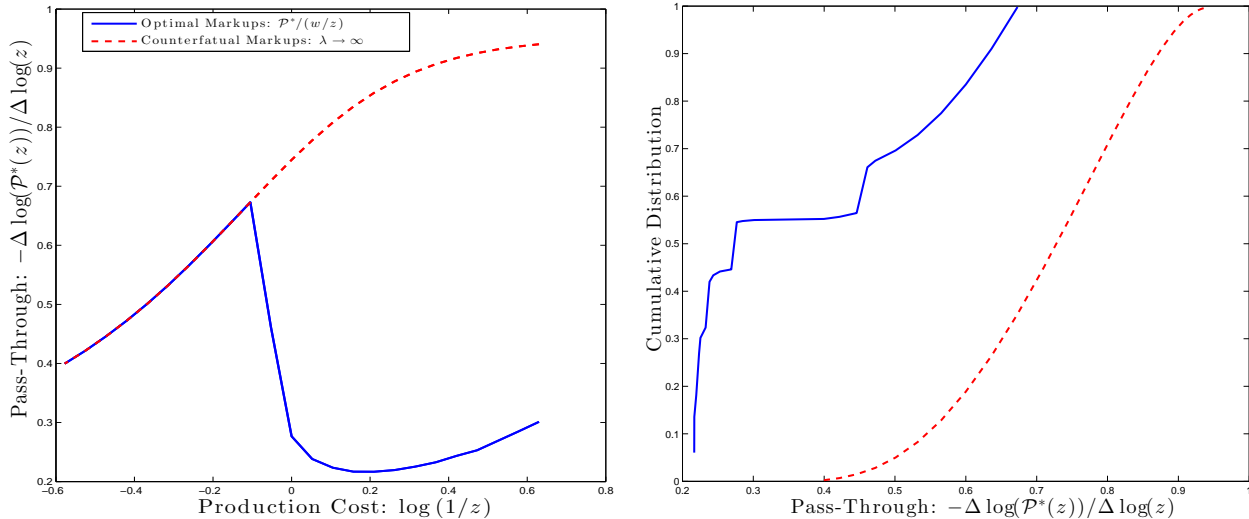


Note: The blue solid line plots the optimal equilibrium markups in the model at our baseline parameters. The red dashed line plots the pass-through in the CES economy with $\lambda \rightarrow \infty$.

The right panel of **Figure 2** displays the distribution of markups for both our economy and the CES one without customer retention concerns. The average markup in our economy is close to 17%, substantially smaller than the one in the CES economy, which is about 27%. The fall in average markups is a natural feature of our setup: because firms value their customers, they have to decrease their prices -and markups- so as to retain them. Interestingly, under the model with customer retention concerns, the dispersion in markups increases dramatically, from 6% in the CES economy to 14% in our economy with customer

retention concerns. The increase in dispersion is entirely driven by the behaviour of low productivity firms which populate the left tail of the markups distribution. The implied distribution of markups in our model displays substantial negative skew: 13% of firms charge negative markups.

Figure 3: Equilibrium Pass-through



Note: The blue solid line plots the equilibrium pass-through in the model at our baseline parameters. The red dashed line plots the pass-through in the CES economy with $\lambda \rightarrow \infty$.

The left panel of **Figure 3** displays the equilibrium pass-through in the two economies. It is useful to define two regions: high productivity firms, i.e. with productivity roughly above the mean, and low productivity firms, i.e. with productivity roughly below the mean. There are two main results. First, pass-through of high productivity firms is identical in the two economies, whereas low productivity firms pass-through substantially less in our economy. This results in a lower average pass-through in our economy, as suggested by the distribution displayed in the right panel of **Figure 3**, that drops from 79% in the CES economy to 38% in our economy. Moreover, while in the CES economy all firms pass-through more than 40%, less than a half do so in our economy. The second main insight relates to the differences in pass-through between high and low productivity firms in the two economies. In the standard CES economy, firms with low productivity pass-through more than firms with high productivity. In stark contrast, in our setting, it is the high productivity firms that pass-through more. For instance, firms in the bottom 10% of the productivity distribution pass-through on average 52% of an idiosyncratic shock to cost, both in our economy and in

the standard CES economy, whereas firms in the top 10% of the productivity distribution pass-through on average 30% in our economy, against 94% in the standard CES economy. [ADD ECONOMIC EXPLANATIONS]

6 The response to aggregate shocks: tariffs vs subsidies

[TBC] (In particular, we need to update the entries on the tables)

In this section we study the response of the economy to an unexpected change in production cost affecting a subset of the firms in the economy.

The experiment is the following. We consider an economy in steady state at period t_0 . We hit a subset $\tilde{J} \subset [0, 1]$ of all the firms in the economy with an unexpected and unforeseen shock to production cost in the form of a scaling factor Z_{t_0} , so that the marginal cost of production of a generic firm $j \in \tilde{J}$ jumps from $w/z_{t_0}^j$ to $w/(z_{t_0}^j Z_{t_0})$. For simplicity the process of Z_t for $t \geq t_0$ is such that Z_t reverts linearly to 1 in T periods, so that $Z_T = 1$. The timing of the shock is such that it hits the firm after idiosyncratic shocks z realize but before pricing and customer's exit decisions are taken.

One can interpret such shock, for instance, as a real exchange rate shock in an economy where both domestic and foreign producers sell to local customers.⁸ Everything else being equal, an appreciation of the real exchange rate is equivalent to an increase in the production cost of foreign producers relatively to domestic producers. Such application of our model is interesting because [Berman et al. \(2012\)](#) provide detailed empirical evidence on the pass-through of real exchange rate shocks to French exporters prices. Thus we can compare the qualitative predictions of our model with available empirical evidence. Another application of the shock is the following. Consider an economy where firms produce in local markets and sell their goods in a centralized market. For example, firms are located in different states. The shock can be interpreted as an increase in local taxes.

We are interested in studying how the pass-through of the shock (i) differs in response to negative and positive shocks, (ii) varies across firms with different idiosyncratic productivities at t_0 (before the common shock hits), (iii) differs between firms $j \in \tilde{J}$ and firms $j \notin \tilde{J}$, and (iv) depends on the size and composition of \tilde{J} . Our baseline experiment is the following. We consider a one period 10% common increase in costs, i.e. $Z_{t_0} = 0.9$ and $T = 1$, for all firms $j \in \tilde{J}$. The mass of all firms $\in \tilde{J}$ amount to 1/4 of all the firms in the economy. The firms $\in \tilde{J}$ are randomly selected so that the distribution of idiosyncratic shocks z among the selected firms and the non-selected firms is identical.⁹

⁸The origin of the movement of the real exchange rate can be justified for instance by a country specific productivity shock.

⁹The fraction 1/4 is approximately equal to the share of imports into U.S. personal consumption expen-

Table 3 displays the pass-through upon impact at t_0 of the common increase in cost to firms $\in \tilde{J}$. The pass through is computed as the log-price difference between the price set in period t_0 after the common shock and the counterfactual price the firm would have chosen if no shock would have occurred computed relatively to the common log-change in cost. We report statistics distinguishing between firms directly hit by the shock, i.e. $j \in \tilde{J}$, and firms not directly hit, i.e. $j \notin \tilde{J}$, but only affected through equilibrium feedbacks. We report the unconditional average for each group, i.e. $j \in \tilde{J}$ and $j \notin \tilde{J}$, as well as averages for different idiosyncratic productivity percentiles within each group. We find that: i) firms directly hit by the increase in cost pass-through on average more (roughly ?? times more) than firms not directly hit; ii) within the group of firms directly hit by the shock, i.e. $j \in \tilde{J}$, more productive firms passthrough more, with the top quartile passing through roughly ?? times more than the bottom quartile; iii) within the group of firms not directly hit by the shock, i.e. $j \notin \tilde{J}$, less productive firms passthrough more, with the bottom quartile passing through roughly ?? times more than the top quartile. The reason for i) is that firms directly hit by the shock have two reasons why changing prices, first because of a change in cost and second because the other firms are changing prices on average in the same direction thus affecting the customer retention problem, while firms not directly hit by the shock only change prices because of the latter. The reason for ii) is that more productive firms face lower customer retention concerns, and thus pass-through relatively more the increase in production cost. The reason for iii) is that more productive firms $j \notin \tilde{J}$ do not change prices both because their costs are not affected, and because they face no change in their demand elasticity either, as the absence of customer retention concerns prevents equilibrium feedback effects to affect their demand.

Table 4 displays the average pass-through upon impact at t_0 for firms $j \in \tilde{J}$ and $j \notin \tilde{J}$ under alternative specifications of our experiment. We consider the following changes (one at the time) from our baseline experiment: a 10% decrease in production cost, i.e. $Z_{t_0} = 1.1$; firms $j \in \tilde{J}$ are not selected randomly but are on average 10% more productive;¹⁰ the fraction of firms hit by the shock is 1/3 and 1/20.¹¹

We obtain the following results: i) the pass-through is asymmetric, so that the pass-through of a decrease in production cost is larger than the one of an (equal size) increase in production cost; ii) the average pass-through is larger when firms in the subset \tilde{J} are on average more productive than firms not in the subset \tilde{J} ; iii) the pass-through is larger when

diture for non-durable goods. Source: **Hale and Hobijn (2011)**.

¹⁰Firms $\in \tilde{J}$ have the same idiosyncratic productivity distribution of firms $\notin \tilde{J}$ but shifted by a constant so that the former have 10% higher mean productivity.

¹¹These particular numbers are interesting if one interprets the common shock as an exchange rate shock because they are approximately equal to the share of imports into U.S. personal consumption expenditure for durable goods and services respectively.

Table 3: The avg. % pass-through to a 10% increase in production cost: baseline experiment

	Firms $\in \tilde{J}$	Firms $\notin \tilde{J}$
All	16.1	5.4
Most productive 10%	34.6	0.0
Most productive 25%	34.6	0.0
Least productive 25%	10.0	7.0
Least productive 10%	10.0	7.0

Notes: The pass-through is computed as the log-change in the price of each producer relatively to the log-change in production cost. Then we take an average across the different producers in a given set of firms, using the weight each producer has into aggregate consumption, and report in % changes. This table reports the unconditional averages, as well as averages conditional on the productivity percentiles of firms, for firms $\in \tilde{J}$ and for firms $\notin \tilde{J}$.

the mass of firms $j \in \tilde{J}$ is larger. The reason for i) is that a decrease in production costs puts the firm in a better position with respect to the ability to retain customers, and thus the firm is more willing to pass-through a decrease in production cost than an increase in production cost. The reason for ii) is that the extensive component of demand of more productive firms is relatively less elastic to changes in production cost and prices. Thus, everything else being equal, these firms would pass-through a larger fraction of the change in production cost. As a result also the feedback effect on the remaining firms will be larger, so inducing also these firms to pass-through relatively more. The reason for iii) is simply that the more aggregate the common shock is, the higher the incentives to pass-through the shock given that a larger fraction of firms is directly affected by such shock, and thus pass-through more.

Table 5 displays the average pass-through to a persistent common increase in production cost, keeping everything else being equal with respect to the baseline specification. The half-life of the common shock 12 months, implying $T = 24$. We obtain the following results: i) the pass-through at t_0 is smaller in the case of the persistent shock than in the baseline case of a one period shock; ii) the size of the pass-through increases as more time elapse since the shock hit.

Table 4: The avg. % pass-through to a change in production cost: other experiments

Change wrt. baseline	Firms $\in \tilde{J}$	Firms $\notin \tilde{J}$
Decrease in cost: $Z_{t_0} = 1.1$	16.1	5.4
Avg. productivity of firms $\in \tilde{J}$ larger	10.0	7.0
Fraction of firms $j \in \tilde{J}$: $1/20$	34.6	0.0
Fraction of firms $j \in \tilde{J}$: $1/3$	34.6	0.0

Notes: The pass-through is computed as the log-change in the price of each producer relatively to the log-change in production cost. Then we take an average across the different producers in a given set of firms, using the weight each producer has into aggregate consumption, and report in % changes. This table reports the unconditional averages for firms $\in \tilde{J}$ and for firms $\notin \tilde{J}$. The first row refers to the experiment where firms $\in \tilde{J}$ experience a 10% decrease in production costs (instead of a 10% increase); the second and third rows refers to the cases where the mass of firms $\in \tilde{J}$ amount to $1/20$ and $1/3$ of total respectively (instead of $1/4$ of total); the fourth row refers to the case where firms $\in \tilde{J}$ have the same idiosyncratic productivity distribution of firms $\notin \tilde{J}$ but shifted so that the former have 10% higher mean productivity.

Table 5: The avg. % pass-through to a change in production cost: persistent shock

Horizon	Firms $\in \tilde{J}$	Firms $\notin \tilde{J}$
Impact: t_0	16.1	5.4
Quarter: $t_0 + 3$	34.6	0.0
Year: $t_0 + 12$	34.6	0.0

Notes: The pass-through is computed as the log-change in the price of each producer relatively to the log-change in production cost. Then we take an average across the different producers in a given set of firms, using the weight each producer has into aggregate consumption, and report in % changes. This table reports the unconditional averages for firms $\in \tilde{J}$ and for firms $\notin \tilde{J}$. This table measures the pass-through at different horizons from the realization of the common shock at t_0 .

7 Concluding remarks

To be completed

References

- 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.
- Coibion, O., Gorodnichenko, Y. and Hong, G. H. (2012). The cyclicalities 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. (2011). Equilibrium wage and employment dynamics in a model of wage posting without commitment, *NBER Working Papers 17284*, National Bureau of Economic Research, Inc.
- 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.

- Goldberg, P. and Hellerstein, R. (2009). How rigid are producer prices?, *Working Papers 1184*, Princeton University, Department of Economics, Center for Economic Policy Studies.
- Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the u.s. automobile industry, *Econometrica* **63**(4): 891–951.
- Goldberg, P. K. and Hellerstein, R. (2013). A structural approach to identifying the sources of local currency price stability, *The Review of Economic Studies* **80**(1): 175–210.
- Gourio, F. and Rudanko, L. (2011). Customer capital, *NBER Working Papers 17191*, National Bureau of Economic Research, Inc.
- Hale, G. and Hobijn, B. (2011). The u.s. content of made in china, *FRBSF Economic Letter* (Aug 8).
- Holmes, T. (2011). The diffusion of wal-mart and economies of density, *Econometrica* **79**: 253–302.
- 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.
- 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 Zerom, D. (2010). Accounting for incomplete pass-through, *Review of Economic Studies* **77**(3): 1192–1230.
- Petrin, A. and Warzynski, F. (2012). The impact of research and development on quality, productivity and welfare, *mimeo*.
- Pistaferri, L. (2003). Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply, *Journal of Labor Economics* **21**(3): 729–754.

- Pollard, P. S. and Coughlin, C. C. (2004). Size matters: asymmetric exchange rate pass-through at the industry level, *Working Papers 2003-029*, Federal Reserve Bank of St. Louis.
URL: <http://ideas.repec.org/p/fip/fedlwp/2003-029.html>
- Ravn, M., Schmitt-Grohe, S. and Uribe, M. (2006). Deep habits, *Review of Economic Studies* **73**(1): 195–218.
- Shi, S. (2011). Customer relationship and sales, *Working Papers tecipa-429*, University of Toronto, Department of Economics.
- Sweezy, P. M. (1939). Demand under conditions of oligopoly, *Journal of Political Economy* **47**: 568.

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 $\theta > (\theta - 1)\frac{1}{N}$. 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 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

First we prove that if $\hat{\mathcal{P}}(z) \leq \bar{\mathcal{P}}(z)$ then $\mathcal{P}(z) = \tilde{\mathcal{P}}(z) = \{\hat{\mathcal{P}}(z)\}$ follows. By definition of $\hat{\mathcal{P}}(z)$ and $\bar{\mathcal{P}}(z)$, and given that firm's value is increasing in both profits per customer and customer base, $\hat{\mathcal{P}}(z) \leq \bar{\mathcal{P}}(z)$ implies $\mathcal{P}(z) = \hat{\mathcal{P}}(z)$. Given that for all $p \leq \bar{\mathcal{P}}(z)$, $\frac{\partial \Delta(p, z)}{\partial p} = 0$, the first order condition in equation (9) reduces to $\frac{\partial \pi(p, z)}{\partial p} = 0$, which is satisfied at $p = \hat{\mathcal{P}}(z)$. Therefore, $\hat{\mathcal{P}}(z) \in \tilde{\mathcal{P}}(z)$. Finally, given that $\pi(p, z)$ is strictly concave in p , attaining its maximum at $\hat{\mathcal{P}}(z)$, and $\Delta(p, z)$ non-increasing in p , we have that $\tilde{\mathcal{P}}(z) = \{\hat{\mathcal{P}}(z)\}$. Next, we show that the converse is true. Assume by contradiction that $\hat{\mathcal{P}}(z) > \bar{\mathcal{P}}(z)$ and $\mathcal{P}(z) = \tilde{\mathcal{P}}(z) = \{\hat{\mathcal{P}}(z)\}$. Because $\frac{\partial \Delta(p, z)}{\partial p} < 0$ for any $p > \bar{\mathcal{P}}(z)$, and given the definition of $\hat{\mathcal{P}}(z)$, we have that $\hat{\mathcal{P}}(z)$ cannot satisfy equation (9), so $\hat{\mathcal{P}}(z) \notin \tilde{\mathcal{P}}(z)$, which constitutes a contradiction of $\tilde{\mathcal{P}}(z) = \{\hat{\mathcal{P}}(z)\}$. Given that $\Delta(p, z)$ is strictly decreasing in p for $p > \bar{\mathcal{P}}(z)$, and given that $\frac{\partial \pi(p, z)}{\partial p} = 0$ at $p = \hat{\mathcal{P}}(z)$ and continuous, the value function of the firm is strictly decreasing at $p = \hat{\mathcal{P}}(z)$. As a result there exists another value of p that gives higher value to the firm, contradicting $\hat{\mathcal{P}}(z) \in \mathcal{P}(z)$.

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

We next show that if the elasticity of the customer base is continuous, then $\mathcal{P}(z) \subset (\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$ and $\mathcal{P}(z) \subset \tilde{\mathcal{P}}(z)$. Start by noting that the requirement for the elasticity

being continuous is that $g(0) = 0$. First, for any $p < \bar{\mathcal{P}}(z)$, the firm's value function is strictly increasing in p because (i) the profit function is strictly increasing in p given that $\bar{\mathcal{P}}(z) < \hat{\mathcal{P}}(z)$, (ii) the customer base is constant. Moreover, because $\Delta(p, z)$ is continuously differentiable with $\lim_{p \uparrow \bar{\mathcal{P}}(z)} \frac{\partial \Delta(p, z)}{\partial p} = 0$, the value function is also strictly increasing in p at $p = \bar{\mathcal{P}}(z)$. As a result, $\mathcal{P}(z) \subset (\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$. Finally, given the value function is strictly increasing in p at $p = \bar{\mathcal{P}}(z)$, and strictly decreasing at $p = \hat{\mathcal{P}}(z)$, and because the value function is continuously differentiable in p , then there is at least a maximum in $(\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$, and such maximum has to satisfy the first order condition, so that $\mathcal{P}(z) \subset \tilde{\mathcal{P}}(z)$.

We now show that, when $g(0) = 0$ and the elasticity of the customer base is non-decreasing in $\bar{\mathcal{P}}(z)$, then $\tilde{\mathcal{P}}(z)$ is a singleton. The right-hand-side term of [equation \(9\)](#) is strictly decreasing, strictly positive for $p < \hat{\mathcal{P}}(z)$, and strictly negative otherwise. Next we study the properties of the left-hand-side term of [equation \(9\)](#). In particular, by continuity of $g(\cdot)$ we have that $-\partial \Delta(p, z) / \partial p$ is equal to zero at $p = \bar{\mathcal{P}}(z)$, and strictly positive for $p > \bar{\mathcal{P}}(z)$. Thus, at least in a neighborhood of $p = \bar{\mathcal{P}}(z)$, $-\partial \Delta(p, z) / \partial p$ is increasing. Moreover, $\Delta(p, z)$ is non-increasing in p . Under the assumption that $-\partial \Delta(p, z) / \partial p$ is also non-decreasing in p for $p \in (\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$, there will be only one level of $p \in (\bar{\mathcal{P}}(z), \hat{\mathcal{P}}(z))$ at which [equation \(9\)](#) is equal to zero.

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

Part (i) of the proposition follows immediately from [Proposition 1](#).

We now prove part (ii) of the proposition. It is useful to recall that $\hat{V} = \int_{\underline{z}}^{\bar{z}} \bar{V}(\mathcal{P}^*(z), z) \mu(z) dz$ and that a customer exits the customer base of the firm if $\hat{V} - \psi > \bar{V}(\mathcal{P}^*(z), z)$.

We first show that in equilibrium there is at least one productivity level where firms experience customers leaving. To prove this notice that the only way no customer would exit the customer base of any firm is if $\bar{V}(\mathcal{P}^*(z), z) = \hat{V}$ for every z . However, under this restriction, $\bar{\mathcal{P}}(z) = \bar{\mathcal{P}}$ for every z . Then, there exists some productivity level at which $\hat{\mathcal{P}}(z) > \bar{\mathcal{P}}$. Then, [Proposition 1](#) implies that at this productivity level the optimal price is strictly above $\bar{\mathcal{P}}$. As a result, there are some productivity levels at which $\bar{V}(\mathcal{P}^*(z), z) < \hat{V}$. Then, for firms with these productivity levels some customers exit and $\mathcal{P}^*(z) > \bar{\mathcal{P}}(z)$. Furthermore, using again [Proposition 1](#), we get that these firms optimal price is below $\hat{\mathcal{P}}(z)$, so that \mathcal{Z}_2 is non-empty. A similar argument implies that \mathcal{Z}_1 is non-empty. This follows because, if some firms' optimal price is such that $\bar{V}(\mathcal{P}^*(z), z) < \hat{V}$, then for another set of firms it has to be the case that $\bar{V}(\mathcal{P}^*(z), z) > \hat{V}$. Because no one exits the customer base of these firms, this implies that $\hat{\mathcal{P}}(z) < \bar{\mathcal{P}}(z)$ so that $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$.

A.4 Proof of Proposition 4

The outline of the proof is the following. We start by showing that the firm pricing decision $\mathcal{P}(z)$ is continuous in z and $\bar{\mathcal{P}}(\cdot)$. Then, we show that, under the assumptions of the proposition, $\bar{\mathcal{P}}(\cdot)$ is continuous in z and firm's optimal pricing $\mathcal{P}(z)$. As a consequence, we combine all these results to obtain a continuous mapping from prices $\mathcal{P}(z)$ into itself, and apply the Brouwer's fixed-point theorem to show that there exists a solution $\mathcal{P}^*(z)$.

Let $\mathcal{T}_1(z, \bar{\mathcal{P}}(z))$ denote the mapping from productivity z and the threshold price $\bar{\mathcal{P}}(z)$ to optimal pricing $\mathcal{P}(z)$.

Claim 1 *If the assumptions of Lemma 2 are satisfied the mapping $\mathcal{T}_1(\cdot, \cdot)$ is unique and continuous in both arguments.*

The proof follows immediately from Proposition 1, equation (3) and Lemma 2. Because the first order condition is continuous in p and $\bar{\mathcal{P}}(z)$, $\mathcal{T}_1(\cdot, \cdot)$ is continuous in both arguments.

Let $\mathcal{T}_2(z, \mathcal{P})$ denote the mapping from productivity z and the optimal price function $\mathcal{P}(\cdot)$ to threshold pricing $\bar{\mathcal{P}}(z)$.

Claim 2 *The mapping $\mathcal{T}_2(\cdot, \cdot)$ is unique and continuous in both arguments.*

The proof follows immediately from equations (1)-(2).

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

Claim 3 *$\mathcal{T}(\cdot, \cdot)$ 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 1, we have that $\mathcal{P}^*(z)$ lies in a compact set. Moreover, we know that $\mathcal{P}(z) \in [W/\bar{z}, \hat{\mathcal{P}}(\underline{z})]$, so that the operator $\mathcal{T}(z, \mathcal{P})$ maps prices \mathcal{P} from the set $[W/\bar{z}, \hat{\mathcal{P}}(\underline{z})]$ to a subset of it. By applying the Brouwer fixed-point theorem, we obtain existence of a solution $\mathcal{P}^*(z) = \mathcal{T}(z, \mathcal{P}^*)$ for all z .

A.5 Proof of Remark 2

We first provide a proof of the first part of the corollary. Under the assumptions of the corollary every firm producing good 1 has the same productivity z at every point in time. As a result, as previously discussed, every firm chooses the same price $\mathcal{P}^*(z)$. We prove the statement of the corollary 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.

Conjecture that firms choose $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$. It is immediate to see that $\bar{\mathcal{P}}(z) = \hat{\mathcal{P}}(z)$. Because $\mathcal{P}^*(z) = \bar{\mathcal{P}}(z)$, a direct application of [Proposition 1](#) validates the conjecture.

We now show by contradiction that $\mathcal{P}^*(z) = \hat{\mathcal{P}}(z)$ is the unique price schedule in equilibrium. There are two cases: one were firms price above $\hat{\mathcal{P}}(z)$ and one were they price below. For the first case, conjecture that $\mathcal{P}^*(z) > \hat{\mathcal{P}}(z)$. Again, it is immediate to see that $\bar{\mathcal{P}}(z) > \hat{\mathcal{P}}(z)$. However, by [Proposition 1](#), given that $\hat{\mathcal{P}}(z) < \bar{\mathcal{P}}(z)$ firms should deviate and price at $\hat{\mathcal{P}}(z)$. Therefore, $\mathcal{P}^*(z) > \hat{\mathcal{P}}(z)$ cannot happen in equilibrium. For the second case, conjecture that $\mathcal{P}^*(z) < \hat{\mathcal{P}}(z)$. Now we have that $\bar{\mathcal{P}}(z) < \hat{\mathcal{P}}(z)$. However, by [Proposition 1](#), given that $\hat{\mathcal{P}}(z) > \bar{\mathcal{P}}(z)$ a firm's optimal price should satisfy $\bar{\mathcal{P}}(z) < \mathcal{P}^*(z) < \hat{\mathcal{P}}(z)$. Therefore, $\mathcal{P}^*(z) < \hat{\mathcal{P}}(z)$ cannot happen in equilibrium.

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

A.6 Proof of [Proposition 5](#)

It proves useful to define the following object,

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

So that the value for a remaining customer is $\bar{V}(p, z) = v(p) + \beta\tilde{V}(z)$. Notice that, because $\mathcal{P}^*(z)$ is decreasing in z , $v(p)$ decreasing in p , and the process for the productivity shock displaying persistence, we have that $\bar{V}(\mathcal{P}^*(z), z)$ is strictly increasing in z . It is then immediate to get that $\bar{\mathcal{P}}(z)$ is strictly decreasing in z . Furthermore, because the price that maximizes static profits $\hat{\mathcal{P}}(z)$ is strictly decreasing in z , implies the existence of the unique threshold \hat{z} where firms with productivity above \hat{z} charge $\hat{\mathcal{P}}(z)$ and those below the threshold charge a price below the price that maximizes their static profits. Finally, [Proposition 3](#) provides that \hat{z} is interior.

A.7 Proof of [Proposition 6](#)

It proves useful to define the following object,

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

So that the value for a remaining customer is $\bar{V}(p, z) = u_0(p) + \beta\tilde{V}(z)$.

1) *Productivity shocks are i.i.d., $\rho = 0$. Points (i)-(ii).* We start by showing that $\bar{\mathcal{P}}(z) = \bar{\mathcal{P}}$. Under the assumptions of the proposition we have that $\mu(z'|z) = \mu(z'|y)$ for all z, y . As

a result, we get that $\tilde{V}(z) = \hat{V}$. Because the continuation value of being matched with any firm equals \tilde{V} , using [equation \(2\)](#) we get that $\bar{\mathcal{P}}(z) = \bar{\mathcal{P}}$. Notice that this result immediately implies that $\varepsilon_m(p, z) = \varepsilon_m(p)$.

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

We now prove that the pricing schedule $\mathcal{P}^*(z)$ is monotonic. We do this in three steps:

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

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

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 $\mathcal{P}^*(z_l) \leq \mathcal{P}^*(z_h)$. We rewrite $\tilde{V}(z)$ as follows,

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

from where it is clear that if $\mathcal{P}^*(z_l) \leq \mathcal{P}^*(z_h)$ then, using [equation \(2\)](#), it is necessary that $\bar{\mathcal{P}}(z_l) \geq \bar{\mathcal{P}}(z_h)$. We now show that when $\bar{\mathcal{P}}(z_l) \geq \bar{\mathcal{P}}(z_h)$, it is the case that $\mathcal{P}^*(z_l) > \mathcal{P}^*(z_h)$, which constitutes a contradiction.

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

$$\begin{aligned} H_1(p, z) &\equiv \frac{\delta + 1 - G(\bar{\psi}(p, z))}{1 - \rho \hat{\beta}(1 - \tau) [\delta + 1 - G(\bar{\psi}(p, z))]} \pi(p, z) , \\ H_2(p, z) &\equiv [\delta + 1 - G(\bar{\psi}(p, z))] \hat{\beta}(1 - \rho) \mathcal{F} , \end{aligned}$$

where \mathcal{F} is constant in productivity z . Notice that, using [equation \(3\)](#), $\arg \max_p H_1(p, z_H) < \arg \max_p H_1(p, z_L)$ if $\bar{\mathcal{P}}(z_l) \geq \bar{\mathcal{P}}(z_h)$. Also, notice that $H_2(p, z)$ is constant for any $p \leq \bar{\mathcal{P}}(z)$ and then strictly decreasing at a rate which is independent of z . As a result, it follows immediately that, as long as $\bar{\mathcal{P}}(z_l) \geq \bar{\mathcal{P}}(z_h)$, 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 $\mathcal{P}^*(z_H) < \mathcal{P}^*(z_L)$, which constitutes a contradiction. Then, $\mathcal{P}^*(z)$ decreases monotonically with z .

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

Point (iii). Monotonicity of $\mathcal{P}^*(z)$ and $\bar{\mathcal{P}}(z)$, together with [Lemma 2](#) and [Lemma 3](#) imply that $\partial \bar{\psi}(p, z) / \partial z \leq 0$ so that $\varepsilon_m(\mathcal{P}^*(z), z)$ decreases with z . The process in [equation \(13\)](#) implies $\alpha(p, z) = 1 + \beta \rho + \beta(1 - \rho) \bar{\Pi} / \pi(p, z)$, where $\bar{\Pi}$ is the average profits. Finally, $\pi(\mathcal{P}^*(z), z)$ increasing in z guarantees that $\alpha(p, z)$ decreases with z . Then by using [equation \(10\)](#) Point (iii) follows.

A.8 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

$$\max_{\ell} \int_{\underline{z}}^{\bar{z}} \left(v(\mathcal{P}^*(z)) - \int_0^{v(\bar{\mathcal{P}}(z)) - v(\mathcal{P}^*(z))} \psi g(\psi) d\psi \right) \frac{M(z)}{\Gamma} dz - \frac{\ell^{1+\phi}}{1+\phi} , \quad (15)$$

subject to

$$v(\mathcal{P}^*(z)) = \frac{\left(I (\mathcal{P}^*(z)^{1-\theta} + 1)^{-\frac{1}{1-\theta}} \right)^{1-\gamma}}{1-\gamma},$$

$$I = w \ell + \int_{\underline{z}}^{\bar{z}} \pi(\mathcal{P}^*(z), z) \frac{M(z)}{\Gamma} dz,$$

and [equation \(3\)](#). 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 \(2003\)](#)).

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 members 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 discounts 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 6 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 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

Table 6: Descriptive statistics on customer shopping behavior

	<i>Mean</i>	<i>Std.dev.</i>	<i>25th pctile</i>	<i>75th pctile</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		

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 7](#) reports descriptive statistics on the change in price of the basket.

Table 7: 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	