Price Dynamics with Customer Markets*

Luigi Paciello

Einaudi Institute for Economics and Finance and CEPR

Andrea Pozzi

Einaudi Institute for Economics and Finance and CEPR

Nicholas Trachter

Federal Reserve Bank of Richmond

May 9, 2014

Abstract

We study optimal price setting in a model with customer markets. Customers face search frictions preventing them from costessly moving across firms. The stickiness in the customer base implies that firms consider customers as an asset and results in reduced markups, more markedly so for less productive firms. We exploit novel micro data on purchases from a panel of households from a large U.S. retailer to quantify the model. The predicted distribution of prices displays substantial excess kurtosis, consistent with recent empirical evidence. Furthermore, we provide evidence that this class of models can play an important role in shaping aggregate dynamics. We show that, coupled with nominal rigidities, customer markets substantially amplifies the real effects of nominal shocks.

JEL classification: E30, E12, L16

Keywords: customer markets, price setting, real rigidities, product market frictions

---

1   Introduction

There is ample consensus that dynamics in the firms' customer base -i.e. the set of customers currently purchasing from a firm- are important determinants of performance, and that firms act to influence its evolution (Foster et al. (2012)). The macroeconomic literature has suggested customer markets as an important determinant of firms' optimal pricing policy and a natural example of real price rigidity (Blanchard (2009)). In this paper we study a model of price setting with competition for customers and a sticky customer base. We characterize the equilibrium of the model and estimate it using novel micro data. Using the estimated model, we study the implications of customer markets for the pricing policy of firms. Our model features dispersion in the price of homogenous goods and a shape of the distribution of prices that is consistent with recent empirical evidence presented in Kaplan and Menzio (2014). Finally, we embed our model of customer markets in an otherwise standard macro model of nominal rigidities to study the propagation of nominal shocks. We find that the model substantially amplifies the persistence of the response of output and prices to the nominal shock, thus magnifying the size of its real effects.

We build on the seminal work on customer markets by Phelps and Winter (1970). In our model, customer dynamics are the result of customers hunting for lower prices. Incentives to hunt derive from price dispersion of otherwise homogeneous good supplied by a large mass of firms characterized by heterogeneous productivity. Search frictions along the lines of Burdett and Coles (1997) introduce stickiness in customer dynamics.[1] Each customer is matched to a particular firm at any point in time and draws a new search cost every period. She has perfect information on the state of the economy as well as on the characteristics of her supplier, and every period decides whether to search for a new supplier; if she does so, she incurs in a cost and is randomly matched to a new firm. After observing the characteristics of the new match, the customer decides if she wants to join the new firm or stay with the old one. Finally the customer allocates her income between the good sold by the supplier she is matched with and another good which is supplied in a centralized market and produced by a perfectly competitive sector without customer markets. The two goods are substitutes, giving rise to a downward sloping demand.

Each firm posts a price common to all customers without commitment every period, before search decisions are taken. The current price affects both the current and future demand of the firm. There are two channels through which the price affects demand. The first channel is static and stems from the standard downward sloping demand of each customer. The second channel is dynamic and concerns the effect of prices on customer dynamics. Inertia

---

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

in the customer base leads firms to consider customers as an asset. The firm faces a trade-off between maximizing static profits per customer and expanding its customer bas. A decrease in the price reduces current profits per-customer but persistently increase the future profits through better retention and acquisition of customers.

We solve for the equilibrium prices in the sector with customer markets. As a result of competition for customers, firms optimal markup over marginal cost is lower than would otherwise be if the customer base were inelastic to prices. When productivity is persistent, markups increase with firm productivity as more productive firms are associated to lower expected prices in the future, and thus customers are less willing to leave their supplier. The relatively less elastic customer base allows the firm to extract more surplus from their current customers.

We complement our modeling effort with an empirical analysis. We exploit scanner data from a major U.S. supermarket chain documenting purchases for a large sample of households between 2004 and 2006. We focus on regular shoppers at the chain and study the extent to which the occurrence of exits from the customer base is affected by variation in the price of the (household specific) basket of consumption. Household level scanner data are particularly well suited to study customer base dynamics. First, we observe a wealth of details on all the shopping trips each household makes to the chain (list of goods purchased, prices, quantities, etc...). More importantly, we can also infer the occurrence of exit from the customer base which we proxy by prolonged spells without purchasing at the chain.

Estimating a linear probability model, we show that customer base dynamics are affected by variation in the price: a one percent change in the price of the customer's typical basket of goods raises her likelihood of leaving the retailer by 0.2 percentage points. We control for demographic characteristics of the household as well as for other variables influencing the propensity of households to hunt for alternative suppliers (distance from the store, distance from competitor stores, prices of competitors, etc..).

The data contains variation useful to identify the key parameters of the model. We target the price elasticity of the customer base obtained in the exercise described above to estimate the size of search costs in our model. Intuitively, the larger the search costs, the less elastic the customer base is to a given variation in prices. We use data on store level prices to infer the volatility and persistence of the idiosyncratic productivity process, exploiting the relationship between equilibrium prices and productivity.

We quantitatively assess the relevance of customer markets for price dynamics by comparing it to an identical economy where however the customer base is inelastic. The estimated model delivers a leptokurtic distribution of prices featuring substantial mass of prices close to the mean, and at the same time fat tails (kurtosis 8.2). The distribution of prices in the

inelastic customer base model is also leptokurtic but with a much smaller kurtosis (2.9). The presence of customer markets reduces dispersion of prices, as less productive firms charge lower price than they would otherwise in an economy with inelastic customer base in order to retain their customers, with the consequence that a large mass of prices concentrates around a lower mean. At the same time, a non-negligible fraction of firms set prices relatively far away from this mean. Part of them (very low productivity firms) charge a higher price because, given the persistence of their status, they find much costly to retain customers and settle for a price that is lower than the static profit maximization but still substantially above the mean. Another set of firms (very high productivity firms) charge a price substantially below the mean simply because they enjoy a productivity advantage which makes such lower price optimal. The relatively high kurtosis of the distribution of prices is consistent with the recent evidence by Kaplan and Menzio (2014). As a consequence of the smaller dispersion in prices, dispersion in markups substantially increases in our model, as the estimated model delivers a strong positive relationship between markup and idiosyncratic productivity (Petrin and Warzynski (2012)).

Customer markets have been suggested in the literature as a natural source of real rigidities (Rotemberg and Woodford (1991), Blanchard (2009)), which are an important determinant in the magnification of the real effects of nominal shocks (Chari et al. (2000), Gopinath and Itskhoki (2011)). This paper develops and estimates a microfounded model of customer markets, delivering a natural laboratory to assess the relevance of this type of real rigidity for the propagation of nominal shocks. In order to address this question we introduce our model of customer markets in a standard macro framework with nominal rigidities. We consider an unexpected shock that permanently increases aggregate nominal spending and compare the propagation with and without the presence of customer markets. We find that customer markets substantially magnify the real effects of nominal shocks: the cumulated impulse response of output is four times larger with customer markets than without; persistence measured by the half-life of output response also increases by a factor of three. This shows that customer markets can substantially amplify the real effects of nominal shocks, hinting towards a larger role for these shocks in explaining business cycle fluctuations.

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

Analyzing the implications of build-up of a customer base for pricing and markup, we tie into a growing body of literature using models where the market share of the firm is sluggish to study a number of issues such as pricing-to-market (Alessandria (2009), Drozd and Nosal (2012)), firm investment (Gourio and Rudanko (2014)), firm dynamics (Luttmer (2006), Dinlersoz and Yorukoglu (2012)), and advertising (Hall (2012)). We instead focus on the influence of customer base concerns on firm price setting as in Bils (1989), Burdett and Coles (1997), Menzio (2007) and Kleshchelski and Vincent (2009). We differ from them in the specifics of the modeling approach and because we quantify the model using empirical evidence directly documenting the comovement of customers and prices. Moreover, none of these papers uses customer markets to study the implications for the distribution of prices nor the propagation of aggregate shocks.

Our model delivers real price rigidities as in models of kinked demand (Kimball (1995)), or in models of imperfect competition where the demand elasticity depends on the market share (Atkeson and Burstein (2008)). A distinctive characteristic of our model is that it introduces a dynamic element in the pricing decision, due to the stickiness of the customer base.

We add to the literature using scanner data to document empirical regularities in pricing and shopping behavior. A series of contributions (Aguiar and Hurst (2007), Coibion et al. (2012), and Kaplan and Menzio (2013, 2014)) integrates store and customer scanner data to show that intensity of search for lower prices depends on income and opportunity cost of time. We instead focus on documenting how the decision to search is triggered by prices.

The rest of the paper is organized as follows. In Section 2 we lay out the model and in Section 3 we characterize the equilibrium. Section 4 presents the data and descriptive evidence of the relationship between customer dynamics and prices. In Section 5 we discuss identification and estimation of the model, and use it to quantify the implications of the model for price and customer dynamics. In Section 6 we perform a policy experiment and document the role of customer markets for the propagation of nominal shocks. Section 7 concludes.

2 The model

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

Customers. We use the index i to denote a customer. Let $d(p)$ and $v(p)$ denote the static demand and customer surplus functions respectively which only depend on the current price

p. We assume that: (i) $d(p)$ is continuously differentiable with $d'(p) < 0$, $\lim_{p\to\infty} d(p) = 0$ and $\varepsilon_d(p) \equiv -\partial \ln d(p)/\partial \ln p \geq 1$; and (ii) $v(p)$ is continuously differentiable with $v'(p) < 0$, $v''(p) \leq 0$ and $\lim_{p\to\infty} v'(p) = -\infty$, $\lim_{p\to 0^+} v'(p) = 0$. Assumption (i) states that the demand function is decreasing in prices and it approaches zero as the price grows, while assumption (ii) states that the surplus of the customer is decreasing and concave in the price. In Appendix C we show that these properties are satisfied in models with CRRA utility functions and CES demand. Each customer starts a period matched to a particular firm, whose characteristics she observes perfectly. Customers are characterized by a random search cost ψ measured in units of customer surplus. The search cost is drawn each period from the same distribution with density $g(\psi)$ on positive support, and associated cumulative distribution function denoted by $G(\psi)$. We restrict our attention to density functions that are continuous on all the support. Upon payment of the search cost the customer draws a random price quote from another firm, with the probability of drawing a particular firm being proportional to its customer base.[2] The customer can decide to accept the offer and exit the customer customer base of her original firm, or decline and stay matched to the old firm. We assume no recall in the sense that once the customer exits the customer base of the firm she cannot go back to it unless she randomly draws it when searching. The customer can search at most once per period.

Firms. We use the index j to denote a firm. The production technology is linear in the unique production input, ℓ, and depends on the firm specific productivity z^j. That is, $y^j = z^j \ell^j$. We let the constant $w > 0$ denote the marginal cost of the input ℓ, p denote the price of the good, and $\pi(p, z) \equiv d(p)(p - w/z)$ denote the profit per customer. We assume that $\pi(p, z)$ is single-peaked. We assume that productivity z is distributed according to a conditional cumulative distribution function $F(z'|z)$ with bounded support $[\underline{z}, \bar{z}]$. We also assume that $F(z'|z_h)$ first order stochastically dominates $F(z'|z_l)$ for any $z_h > z_l$. The only choice firms make is to to set prices.

Timing of events. A firm starts a period matched to the set of customers she had retained at the end of the previous period (m_{t-1}^j). The timing of events is the following: (i) productivity shocks are realized for all firms and each firm j posts a price p_t^j without commitment, (ii) each customer draws her search cost ψ_t^i and observes the price p_t^j as well as the relevant state of the firm she is matched with (i.e. z_t^j and m_{t-1}^j), (iii) each customer decides whether to search for a new firm or remain matched to her current one, (iv) if the customer decides to search, she pays the search cost and draws a new supplier j' with probability $m_{t-1}^{j'}/\Gamma$. The

---

[2]This captures the idea that larger firms attract more customers (Rob and Fishman (2005)).

customer perfectly observes not only the price but also productivity and customer base of the prospective match and decides whether to exit the customer base of the current supplier to join that of the new match or to stay with the current match. Finally, (v) customer surplus $v(p_t^j)$ and profits $\pi(p_t^j, z_t^j)$ are realized.

Equilibrium. A firm and its customers play an anonymous sequential game. We look for a stationary Markov Perfect equilibrium where strategies are a function of the current state. There are no aggregate shocks. Although the relevant state for the pricing decision of the firm could in principle include both the stock of customers and the idiosyncratic productivity, we conjecture and show the existence of an equilibrium where optimal prices only depend on productivity, and we denote by $\mathcal{P}(z)$ the equilibrium pricing strategy of the firm. The relevant state for the search decision of a customer includes the expectations about the path of current and future prices of the firm she is matched to, as well as the idiosyncratic search cost. Given the Markovian equilibrium we study, the current realization of idiosyncratic productivity is a perfect statistic about the distribution of future prices. As a result, the search strategy of the customer depends on the current price and productivity of the firm she is matched to, and on her own search cost. We denote the search decision as $s(p, z, \psi) \in \{0, 1\}$, where $s = 1$ means that the customer decides to search. Conditional on searching, the exit decision depends on the continuation value associated to the firm the customer starts matched to (the outside option), which is fully characterized by posted price and productivity, as well as on productivity of the firm she has drawn upon the search, z', which fully characterizes the continuation value associated to the new firm. We denote the exit decision as $e(p, z, z') \in \{0, 1\}$, where $e = 1$ means that the customer decides to exit the customer base of her original firm.

2.1 The problem of the customer

Consider a customer buying goods from firm j, and let $V(p, z, \psi)$ denote the value function for her of being matched to firm j -which has current productivity z and posted price p- and that has drawn a search cost equal to ψ. We have that this value function solves the following problem,

$$V(p_t^j, z_t^j, \psi_t^i) = \max\left\{ \bar{V}(p_t^j, z_t^j), \, \tilde{V}(p_t^j, z_t^j) - \psi_t^i \right\} , \tag{1}$$

where $\bar{V}(p, z)$ is the customer's value if she does not search, and $\tilde{V}(p, z) - \psi$ is the value if she does search. Given the pricing function $\mathcal{P}(\cdot)$ mapping future productivity into prices in

the Markov equilibrium, the value in the case of not searching is given by

$$\bar{V}(p_t^j, z_t^j) = v(p_t^j) + \beta \int_0^\infty \int_{\underline{z}}^{\bar{z}} V(\mathcal{P}(z'), z', \psi') \, dF(z'|z_t^j) \, dG(\psi') \,. \tag{2}$$

The value when searching is given by

$$\tilde{V}(p_t^j, z_t^j) = \int \max \left\{ \bar{V}(p_t^j, z_t^j), \, x \right\} dH(x) \,,$$

where the customer takes expectations over all possible draws of potential new firms, each of them providing a value $\bar{V}'$ to the customer is she decides to join the new firm, and where $H(\cdot)$ is the equilibrium cumulative distribution of continuation values from which the firm draws a new potential match when searching. For instance, $H(\bar{V}(p_t^j, z_t^j))$ is the probability of drawing a potential match offering a continuation value smaller or equal than the current match.

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

Lemma 1 *The customer matched to a firm with productivity z_t^j charging price p_t^j: i) searches if she draws a search cost ψ_t smaller than a threshold, i.e. $\psi_t \leq \hat{\psi}(p_t^j, z_t^j)$, where $\hat{\psi}(p, z) = \int_{\bar{V}(p,z)}^\infty \left(x - \bar{V}(p, z)\right) dH(x) \geq 0$; ii) conditional on searching, exits if she draws a new firm promising a continuation value $\bar{V}'$ larger than the current match, i.e. $\bar{V}' \geq \bar{V}(p_t^j, z_t^j)$.*

The proof of the lemma is in Appendix A.1. Given that search is costly, not all customers currently matched to a given firm exercise the search option; only those with a low search cost ψ do so. Notice the threshold $\hat{\psi}(p, z)$ depends on both the price of the firm, p, and its productivity, z. The dependence on the price is straightforward, following from its effect on the surplus $v(p)$ that the customer can attain in the current period. The intuition behind the dependence on the firm's productivity is that, as searching is a costly activity, the decision of which firm to patronize is a dynamic one, and involves comparing the value of remaining in the customer base of the current firm with the value of searching. Because of the Markovian structure of prices, customer's expectation about future prices are completely determined by the firm's current productivity. firm.

The next lemma discusses some useful properties of the continuation value function $\bar{V}(p, z)$.

Lemma 2 *The value function $\bar{V}(p, z)$ (the threshold $\hat{\psi}(p, z)$) is strictly decreasing (increasing) in p. If $\hat{V}(z) \equiv \bar{V}(\mathcal{P}(z), z)$ is increasing in z, the value function $\bar{V}(p, z)$ (the threshold $\hat{\psi}(p, z)$) is increasing (decreasing) in z.*

The proof of Lemma 2 is in Appendix A.2. An important implication of the lemma is that, not only customers are more likely to search and exit from firms charging higher prices, but also that they are more likely to do so from firms with lower productivity. This follows from the dependence of the expected future path of prices on the firm's current productivity as, under the assumption that $\hat{V}(z)$ is increasing in z, firms with lower productivity offer low continuation value to customers.

2.2 The problem of the firm

In this section we describe the pricing problem of the firm. We start by discussing the dynamics of the customer base as a function of price and productivity, given the optimal search and exit strategy of the customers. Then we move to the setup and characterization of the firm pricing strategy.

The customer base of a generic firm j at period t (m_t^j) is the mass of customers buying from firm j in period t. It evolves as follows

$$m_t^j = m_{t-1}^j - \underbrace{m_{t-1}^j \, G\left(\hat{\psi}(p_t^j, z_t^j)\right)\left(1 - H(\bar{V}(p_t^j, z_t^j))\right)}_{\text{customers outflow}} + \underbrace{\frac{m_{t-1}^j}{\Gamma} Q\left(\bar{V}(p_t^j, z_t^j)\right)}_{\text{customers inflow}}, \qquad (3)$$

where m_{t-1}^j is the mass of old customers, $G(\hat{\psi}(p_t^j, z_t^j))$ is the fraction of old customers searching, a fraction $1 - H(\bar{V}(p_t^j, z_t^j))$ of which actually finds a better match and exits the customer base of firm j. The ratio m_{t-1}^j/Γ is the probability that searching customers in the whole economy draw firm j as a potential match. The function $Q(\bar{V}(p_t^j, z_t^j))$ denotes the equilibrium mass of searching customers currently matched to a firm with continuation value smaller than $\bar{V}(p_t^j, z_t^j)$. Therefore, the product of the two amounts to the mass of new customers entering the customer base of firm j. We can express the dynamics in the customer base as $m_t^j = m_{t-1}^j \Delta(p_t^j, z_t^j)$, where the function $\Delta(\cdot)$ denotes the growth of the customer base and is given by

$$\Delta(p, z) \equiv 1 - G\left(\hat{\psi}(p, z)\right)\left(1 - H(\bar{V}(p, z))\right) + \tfrac{1}{\Gamma} Q\left(\bar{V}(p, z)\right). \qquad (4)$$

The assumption that the probability that a firm is proposed to a searching customer as her new potential match is proportional to its customer base, coupled with linear production technology, implies that the growth of a firm is independent of its size. This result is known as Gibrat's Law, and is consistent with existing empirical evidence on the distribution of firms' size (see Luttmer (2010)). The next lemma discusses the properties of the customer base growth with respect to prices and productivity.

Lemma 3 *Let $\bar{p}(z)$ solve $\bar{V}(\bar{p}(z), z) = \max_z \{\bar{V}(\mathcal{P}(z), z)\}$; $\Delta(p, z)$ is strictly decreasing in p for all $p > \bar{p}(z)$, and constant for all $p \leq \bar{p}(z)$. If $\hat{V}(z) \equiv \bar{V}(\mathcal{P}(z), z)$ is increasing in z, then $\Delta(p, z)$ is increasing in z.*

The proof of Lemma 3 follows directly from Lemma 2. The growth of the customer base is decreasing in the current price because a higher price reduces the current surplus and therefore the value of staying matched to the firm. When the price is low enough that no firm in the economy offers a higher value to the customer, the customer base is maximized ant a further decrease in the price has no impact on the customer growth. If $\hat{V}(z)$ is increasing in z, the growth of the customer base increases with firm productivity, as a larger z is associated to higher continuation value which increases the value of staying matched to the firm.

We next discuss the pricing problem of the firm. The firm pricing problem in recursive form solves

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

subject to equation (3), where $\tilde{W}(z_t^j, m_{t-1}^j)$ denotes the firm value at the optimal price and $\pi(p, z_t^j) = d(p)(p - w/z_t^j)$ is profits per customer. We study equilibria where the pricing decision of the firm only depends on productivity. Thus, we conjecture that in this equilibrium the value function for a firm is homogeneous of degree one in m, i.e., $\tilde{W}(z, m) = m\, \tilde{W}(z, 1) \equiv m\, W(z)$, where $W(z)$ solves

$$W(z) = \max_p \; \Delta(p, z) \underbrace{\left(\pi(p, z) + \beta \int_{\underline{z}}^{\bar{z}} W(z') dF(z' \,|\, z) \right)}_{\text{present discounted value of a customer} \; \equiv \; \Pi(p, z)} \; , \tag{5}$$

where we used equation (3) and we dropped time and firm indexes to ease the notation. We assume that the discount rate β is low enough so that the maximization operator in equation (5) is a contraction, so that by the contraction mapping theorem we can conclude that our conjecture about homogeneity of $\tilde{W}(z, m)$ is verified.

We can express the objective of the firm maximization problem as the product of two terms. The first term is the growth in the customer base, $\Delta(p, z)$, which according to Lemma 3 is decreasing in the price for all $p > \bar{p}(z)$ and is maximized at any price $p \leq \bar{p}(z)$. The second term is the expected present discounted value to the firm of each customer, which we denote by $\Pi(p, z)$. The function $\Pi(p, z)$ is maximized at the static profit maximizing price,

$$p^*(z) = \frac{\varepsilon_d(p)}{\varepsilon_d(p) - 1} \frac{w}{z} \; . \tag{6}$$

9

It follows that setting a price above the static profit maximizing price is never optimal. Moreover, if $\bar{p}(z) \leq p^*(z)$, the optimal price will not be below $\bar{p}(z)$, because in that region profit per customer decrease in the price but the customer base is unaffected, so that $\hat{p}(z) \in [\bar{p}(z), p^*(z)]$. If instead $\bar{p}(z) \geq p^*(z)$, then the optimal price is the static profit maximizing price, $\hat{p}(z) = p^*(z)$, as at this price both the customer base and the profit per customer are maximized. The following proposition collects these results and provides necessary conditions under which the optimal price can be characterized as the solution to the first order condition and uses it to discuss the properties of optimal markups.

Proposition 1 *Let $\bar{p}(z)$ solve $\bar{V}(\bar{p}(z), z) = \max_{z \in [\underline{z}, \bar{z}]} \{\bar{V}(\mathcal{P}(z), z)\}$, and let $p^*(z)$ be the price that maximizes the static profit in equation (6). Denote by $\hat{p}(z)$ a price that solves the firm problem in equation (5). If $G(\cdot)$ is differentiable for all $\psi \in [0, \infty)$, then $\hat{p}(z) \in [\bar{p}(z), p^*(z))$ if $\bar{p}(z) < p^*(z)$, and $\hat{p}(z) = p^*(z)$ otherwise. Moreover, if the distributions $Q(\cdot)$ and $H(\cdot)$ are differentiable on all their supports, $\hat{p}(z)$ must solve the following first order condition,*

$$\frac{\partial \Pi(p, z)}{\partial p} \frac{p}{\Pi(p, z)} = \underbrace{\frac{p}{\Delta(p, z)} \frac{\partial \Delta(p, z)}{\partial p}}_{customer\ growth\ elasticity\ \equiv\ \varepsilon_m(p,z)} \geq 0 \ , \tag{7}$$

for each z.

A proof of the proposition can be found in Appendix A.3. The first order condition is not in general sufficient for an optimum to the firm problem as the firm objective, and in particular $\Delta(p, z)$, is not in general a concave function of p. The first order condition is however sufficient if the customer growth elasticity, $\varepsilon_m(p, z)$, is non-decreasing in p for all $p \leq p^*(z)$, i.e. $\Delta(p, z)$ is a concave function of p in the relevant region of prices. This is a useful property to know because it will be satisfied at the parameter estimates in our empirical exercise.

The first order condition illustrates the trade-off the firm faces when setting the price in a region where customer retention is a concern: if $\bar{p}(z) < p^*(z)$ the optimal price balances the marginal benefit of an increase in price (more profit per customer) with the cost (decrease in the customer base). From equation (7) we can obtain an expression for the optimal markup,

$$\mu(p, z) \equiv \frac{p}{w/z} = \frac{\varepsilon_d(p)}{\varepsilon_d(p) - 1 + \varepsilon_m(p, z)\, x(p, z)} \ . \tag{8}$$

The terms $\varepsilon_d(p)$ and $\varepsilon_m(p, z)$ represent the price elasticities of quantity purchased (per-customer) and of customer growth, respectively. An increase in price reduces total current demand both because it reduces quantity per customer (*intesinve margin effect*) and because it reduces the number of customers (*extesinve margin effect*). Moreover, the optimal markup

10

solves a dynamic problem as a loss in customers has persistent consequences for future demand due to the inertia in the customer base. This dynamic effect is captured by the term $x(p, z) \equiv \Pi(p, z)/(d(p) \, p) \geq 0$ which measures the firm present discounted value of a customer scaled by the current revenues. It follows that active customer markets are associated to strictly lower markup than the one that maximizes static profit, the lower, the larger the product $\varepsilon_m(p, z) \, x(p, z)$.

To clarify the importance of the dynamic effect on optimal markups, consider the following thought experiment. Define the overall demand elasticity of an economy as the sum of its quantity elasticity and its customer growth elasticity: $\varepsilon_q(p, z) \equiv \varepsilon_d(p) + \varepsilon_m(p, z)$. Take two firms characterized by the same productivity z and the same overall demand elasticity, but by different combinations of $\varepsilon_d(\cdot)$ and $\varepsilon_m(\cdot)$. In particular, one firm has lower quantity elasticity but higher customer growth elasticity than the other. Then the optimal markup for the former is strictly lower than that for the latter.[3]

3   Equilibrium

In this section we define an equilibrium, discuss its existence, and characterize its general properties. We start by defining the type of equilibrium we study.

Definition 1 *Let $\hat{V}(z) \equiv \bar{V}(\mathcal{P}(z), z)$ and $p^*(z)$ be given by equation (6). We study stationary Markovian equilibria where $\hat{V}(z)$ is non-decreasing in z, and for all $z \in [\underline{z}, \bar{z}]$ the firm pricing strategy lies in the compact set $[p^*(\bar{z}), p^*(\underline{z})]$. A stationary equilibrium is*

(i) a search and an exit strategy that solve the customer problem for given equilibrium pricing strategy $\mathcal{P}(z)$, as defined in Lemma 1;

(ii) a firm pricing strategy $\hat{p}(z)$ that solves the firm's problem in equation (5), given customers' strategies and equilibrium pricing policy $\mathcal{P}(z)$, and is such that $\hat{p}(z) = \mathcal{P}(z)$ for each z;

(iii) two distributions over the continuation values to the customers, $H(x)$ and $Q(x)$, that solve $H(x) = K(\hat{z}(x))$ and $Q(x) = \Gamma \int_{\underline{z}}^{\hat{z}(x)} G(\hat{\psi}(\hat{p}(z), z)) \, dK(z)$ for each $x \in [\hat{V}(\underline{z}), \hat{V}(\bar{z})]$, where $\hat{z}(x) = \max\{z \in [\underline{z}, \bar{z}] : \hat{V}(z) \leq x\}$, and $K(z)$ solves

$$K(z) = \int_{\underline{z}}^{z} \int_{\underline{z}}^{\bar{z}} \Delta(\hat{p}(x), x) \, dF(s|x) \, dK(x) \, ds \ , \tag{9}$$

for each $z \in [\underline{z}, \bar{z}]$ with boundary condition $\int_{\underline{z}}^{\bar{z}} dK(x) = 1$.

---

[3]More details are available in Appendix A.4.

The requirement that the continuation value to customers is non-decreasing in productivity implies that customers rank of firms coincides with their productivity. This is a natural outcome as more productive firms are better positioned to offer lower prices and therefore higher values to customers. The requirement that $\hat{p}(z) \in [p^*(\bar{z}), p^*(\underline{z})]$ excludes those equilibria where firms cut prices below the static profit maximizing price of the most productive firm. Notice that Proposition 1 implies that $\hat{p}(z) \le p^*(z)$ for all $z \in [\underline{z}, \bar{z}]$, so that we are effectively restricting only the lower bound.

The next proposition states the existence of an equilibrium and characterizes its properties.

Proposition 2 *Let productivity be i.i.d. with $F(z'|z_1) = F(z'|z_2)$ continuous and differentiable for any z' and any pair $(z_1, z_2) \in [\underline{z}, \bar{z}]$, and let $G(\psi)$ be differentiable for all $\psi \in [0, \infty)$, with $G(\cdot)$ differentiable and not degenerate at $\psi = 0$. There exists an equilibrium as described in Definition 1 where $\hat{p}(z)$ satisfies equation (7), and*

    (i) $\hat{p}(z)$ is strictly decreasing in z, with $\hat{p}(\bar{z}) = p^*(\bar{z})$ and $\hat{p}(\bar{z}) < \hat{p}(z) < p^*(z)$ for $z < \bar{z}$, implying that $\hat{V}(z)$ is strictly increasing;

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

The proposition highlights the main properties of the equilibrium we study. The equilibrium is characterized by price dispersion: more productive firms charge lower prices and, therefore, offer higher continuation value to customers and grow faster. As shown in Proposition 1, the presence of customer markets reduces markups for each productivity level relatively to the case where firms maximize static profits, i.e. $\hat{p}(z) < p^*(z)$ for all $z < \bar{z}$. In equilibrium, there is a positive mass of lower productivity firms that have a shrinking customer base, and a positive mass of higher productivity firms that are expanding their customer base.

Notice that differentiability of the distribution of productivity F is needed to ensure that $H(\cdot)$ and $Q(\cdot)$ are almost everywhere differentiable so that equation (7) is a necessary condition for optimal prices. However, equation (7) is not necessary for the existence of an equilibrium as described in Definition 1. Even when F is not differentiable and the first order condition cannot be used to characterize the equilibrium, an equilibrium with the properties of Proposition 2 exists where $\hat{p}(z)$ and $\hat{\psi}(\hat{p}(z), z)$ are monotonic in z but not necessarily strictly monotonic for all z. Monotonicity of optimal prices follows from an application of Topkis theorem. In order to apply the theorem to the firm problem in equation (5) we need to establish increasing differences of the firm objective $\Delta(p, z) \Pi(p, z)$ in $(p, -z)$. Under

the standard assumptions we stated on $\pi(p, z)$ it is easy to show that $\Pi(p, z)$ satisfies this property. The customer base growth does not in general verifies the increasing difference property. However, under the assumption of i.i.d. productivity $\Delta(p, z)$ is independent of z which, together with Lemma 3, is sufficient to obtain the result. More details on the proof of the proposition can be found in Appendix A.5. Finally, while the results of Proposition 2 refer to the case of i.i.d. productivity shocks, numerical results in Section 5.2 show the properties of Proposition 2 extend to the case of a persistent productivity process.

The next remark shows that our model nests two limiting cases that have been extensively studied in the literature. First, if we let the search cost diverge to infinity, i.e. $G(\psi) = 0$ for all $\psi < \infty$, we obtain a model where customer base concerns are not present. Because the customer base is unresponsive to prices, the firm problem reduces to the standard price setting problem under monopolistic competition. In the second limiting case, we explore the equilibrium under the assumption that firms share the same constant level of productivity, i.e. $F(z'|z)$ is degenerate at some productivity level $z_0 \in [\underline{z}, \bar{z}]$. Under the type of Markovian equilibrium we study, firms have to charge the same price when they have the same productivity. In this case the equilibrium price must be the price that maximizes static profits, a result reminiscent of Diamond (1971).

Remark 1 *Two limiting cases of the equilibrium stated in Definition 1:*

(1) *Suppose that $G(\psi) = 0$ for any arbitrarily large level of ψ. Then, in equilibrium: (i) the optimal price maximizes static profits, $\hat{p}(z) = p^*(z)$ for all z, (ii) equilibrium markups $\mu(\hat{p}(z), z)$ are increasing in productivity, and (iii) there is no search in equilibrium. Furthermore, the equilibrium is unique.*

(2) *Suppose that $G(\cdot)$ is differentiable and not degenerate at $\psi = 0$. Let the productivity distribution be degenerate at some $z = z_0$. Then, there is a unique Markovian equilibrium where each firm charges the price that maximizes static monopoly profits, i.e. $\hat{p}(z_0) = p^*(z_0)$.*

A proof can be found in Appendix A.6.

4   Data

We complement the theoretical analysis with an empirical investigation that relies on cashier register data from a large US supermarket chain. The empirical analysis has two purposes. First, we document that changes in the price posted by the firm influence customers'decision to exit the customer base and measure the size of this effect. Second, we use the data to

estimate our model and quantify the importance of customer markets in shaping firm price setting.

4.1 Data sources and variable construction

The data include purchases by households who carry a loyalty card of the chain. For every trip made at the chain by a panel of households between June 2004 and June 2006, we have information on the 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. This data are particularly suitable to our focus as we are interested in the behavior of regular customers, who typically carry a loyalty card. To be conservative, we keep in our data only households who shop at the chain at least twice a month, so to remove occasional shoppers. Within this sample, the average number of shopping trips at the chain is 157 shopping trips over the two years; if those trips were uniformly distributed that would imply visiting a store of the chain six times per month.

In the theoretical model we studied the behavior of customers buying from firms producing a single homogeneous good. Our application documents the exit decisions of customers from supermarket stores.[4] In this context, customers buy bundles of goods and therefore we assume that their behavior depends on the price of the basket of goods they typically buy at the supermarket.[5] While the multiproduct nature of the problem may have implications for the pricing decision of the firm, we abstract from this issue and only focus on the resulting price index of the customer basket which we use to measure the comovement between the customer's decision to exit the customer base and the price of her typical basket of goods posted at the chain. To do so we need to construct two key variables: (i) an indicator signaling when the household is exiting the chain's customer base, and (ii) the price of the household basket. Below we briefly describe the procedure followed to obtain them, the details are left to Appendix B.

We consider every customer shopping at the retailer in a given week as belonging to the chain's customer base in that week. We assume that a household has exited the customer base when she has not shopped at the chain for eight or more consecutive weeks and we date the exit event to the last time the customer visited the chain. The eight-weeks window is a conservative choice since households in our sample shop much more frequently than

---

[4]The choice of focusing on the customer base of the store rather than that of one of the branded product it sells is data driven. With data from a single chain we cannot track the evolution of the customer base of a single brand. In fact, if we observed customers stopping to buy a particular brand we can only infer she is not buying it at our chain, but we cannot exclude she is not buying it elsewhere.

[5]Note that since customers baskets are in large majority composed of package goods, which are standardized products, the assumption that the basket is a homogenous good is not unwarranted.

that. Regular customers are unlikely to experience a eight-weeks spell without shopping for reasons other than having switched to another chain (e.g. consuming their inventory). In fact, the average number of days elapsed between consecutive trips is close to four and the 99th percentile is 24 days, roughly half the length of the absence we require before inferring that a household is buying its grocery at a competing chain.

We construct the price of the basket of grocery goods usually purchased by the households in a fashion similar to Dubois and Jodar Rosell (2010). We identify the goods belonging to a household's basket using scanner data on items the household purchased over the two years in the sample. The price of its basket in a particular week is then computed as the average of the weekly prices of the goods included in the basket, weighted by their expenditure share in the household budget. Since households differ in their choice of grocery products and in the weight such goods have in their budget, the price of the basket is household specific. We face the common problem that household scanner data only contain information on prices and quantities of UPCs when they are actually purchased. Therefore we complement our data with store level data on weekly revenues and quantities sold.[6] This data allows us to to construct weekly prices of 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.

4.2   Evidence on 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 chain in a particular week. Our regressor of interest is the logarithm of the price of the basket of grocery goods usually purchased by the households at the chain ($p^{retailer}$).

In Table 1, we report results of regressions of the following form,

$$Exit_{it} = b_0 + b_1 p_{it}^{retailer} + X_i' b_2 + \varepsilon_{it} \ . \tag{10}$$

In the regression we include year-week fixed effects to account for time-varying drivers of the decision of exiting the customer base common across households and we control for observable characteristics, such as age, income, and education, through inclusion of household's demographics matched from Census 2000. We add the number of competing grocery retailers in the zipcode, as well as the distance (in miles) from the closest store of the chain and that from the closest store of the competition to account for the fact that households living closer to outlets of the chain and far away from alternative options will be less likely to leave

---

[6]The retailer changes the price of the UPCs at most once per week, hence we only need to construct weekly prices to capture the entire time variation.

the customer base of the chain. Finally, we include as regressors the logarithm of the price of the basket in the first week in the sample and the standard deviation of price changes for each household over the sample period. These are meant to control for differences in the composition of the basket across shoppers. For example, some customers may purchase product categories more prone to promotions than others and experience more intense price fluctuations as a result.

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

Exiting: Missing at least 8 consecutive weeks

	(1)	(2)
$\log(P_i^{retailer})$	0.20***	0.23***
	(0.074)	(0.073)
$\log(P_i^{competitors})$		0.001
		(0.001)
Observations	71,049	52,670

Notes: An observation is a household-week pair. The sample only includes households who prominently shop at stores for which we have complete price data for all the UPCs they purchase. We exclude from the sample the top and bottom 1 % in the distribution of the number of trips over the two years. Demographic controls rely on a subsample of households for which information on the block-group of residence was provided and include as regressors ethnicity, family status, age, income, education, and time spent commuting (all matched from Census 2000) as well as distance from the closest outlet of the supermarket chain and distance from the closest competing supermarket (provided by the retailer). The logarithm of the price of the household basket in the first week in the sample and the standard deviation of changes in the log-price of the household basket over the sample period are included as a controls in all specifications. Week-year fixed effects are also always included. Standard errors are in parenthesis. ***: Significant at 1% **: Significant at 5% *: Significant at 10%.

Table 1 reports the results of a regression whose dependent variable is an indicator that takes value one if in that week the customer decides to leave the retailer, and zero otherwise. Column (1), documents that a 1% increase in the weekly price of the customer-specific grocery basket is associated with 0.2 percentage points increase in the probability that the customer leaves the chain to patronize a rival firm.[7] The coefficient on the price of the basket is identified by *UPC-chain* specific shocks as those triggered, for example, by the expiration of a contract between the chain and a manufacturer of a UPC. Furthermore, we also exploit variation in our data from *UPC-store* specific shocks: within the chain, the price of a same good moves differently in different stores. This can be due, for instance, to variation in

---

[7]Notice that since productivity does not enter the equation as a separate regressor, the coefficient b_1 conflates two different effects of the price on the customer's exit decision when interpreted through the lens of our model. The first is static and stems from the impact of the price on the contemporaneous utility of the customer. The second is dynamic and depends on information the price contains about future productivity and continuation value of the customer.

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).

Endogeneity of prices to the exit decision of a customer is unlikely in this setting. First, our approach differs from the standard discrete choice studies of demand. We are not modeling the household's choice of the preferred retailer among a set of potential alternatives, but rather the decision of whether to leave a given retailer or to stick with it. It implies that the usual concern of price endogeneity driven by unobserved store characteristics is not relevant in this context. Moreover, the customer reacts to his own specific basket price, whereas the firm sets UPC prices common to all customers. Even if the retailer were to observe variables predicting the exit from the customer base of specific households, it is unlikely that it can react with targeted prices for them. In fact the basket of different households will partially overlap making it impossible to fine tune the basket price faced by some households without affecting the price of others.

The results in column (1) do not control for the pricing behavior of the competitors. This may raise concerns on the precision of our estimate of the elasticity to price. Absent information on the level of prices at competing stores, we cannot tell whether shifts in the price of the basket at the chain are idiosyncratic or due to shocks common to all the other retailers in the market. Only shock idiosyncratic to the chain should be expected to affect the probability of leaving the chain. Aggregate cost shocks do not change the relative price and, therefore, should not trigger exit from the customer base. Furthermore, our retailer is a major player in the markets included in our sample and it is reasonable to assume that the competition takes its prices into consideration when deciding on their own. This possibly introduces correlation between price variations at the chain and price variations at the alternative outlets the customer may visit. Disregarding the prices of the competitors may therefore lead to biases in the magnitude and even the sign of the own-price elasticity. In column (2) we address both of these concerns by directly controlling for the prices posted by competitors of the chain using the IRI Marketing data set. This source includes weekly UPC's prices for 30 major product categories for a representative sample of chain stores across 64 markets in the US.[8] Using this data, we can compute the price of each UPC in the Metropolitan Statistical Area of residence of a customer by averaging the price posted for the item by all the chains sampled by IRI. Then, we construct the average market price of the basket bought by the customer in the same fashion described for the price of the basket at

---

[8]A detailed description of the data can be found in Bronnenberg et al. (2008). All estimates and analyses in this paper based on Information Resources Inc. data are by the authors and not by Information Resources Inc.

our chain.[9] Even after controlling for the general level of prices, the coefficient on the price of the basket at the chain stays negative and significant and nearly unaffected in magnitude, suggesting that most of the variation in the index comes from chain, or UPC-chain, specific shocks.

5 Estimation and quantitative analysis

In this section we discuss the procedure we follow to calibrate and solve the model. We need to choose the discount factor β and the nominal wage w, as well as four functions: the demand function, $d(p)$, the surplus function $v(p)$, the distribution of search costs $G(\psi)$, and the conditional distribution of productivity $F(z'|z)$ for all $z \in [\underline{z}, \bar{z}]$. We next we discuss the parametrization of the model in detail.

Discount factor. We assume that a period in the model corresponds to a week to mirror the frequency of our data. We fix the firm discount rate is $\beta = 0.995$. In the set of parameters that we consider, this level of β ensures that the max-operator in equation (5) is a contraction.[10]

Customer demand and surplus functions. We assume that customers derive utility from consumption according to the function $\log(c)$, where c is a composite of two types of goods defined as $c = \left(d^{\frac{\theta-1}{\theta}} + n^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}$, with $\theta > 1$.[11] One (that we label d) is the good supplied by the type of firms described in Section 2.2; the other good that we label (n) acts as a numeraire and is sold in a centralized market. The sole purpose of good n is to microfound a downward sloping demand $d(p)$. The parameter θ is chosen so that the implied average markup in absence of customer retention concern (Monopolistic economy) is about 10%, a value in the range of those used in the macro literature. The customer budget constraint is given by $p\,d + n = I$, where I is the agent's nominal income which we normalize to one.

Firms productivity process. We assume that the productivity follows a simple process of the following form: $\log(z_t^j) = \log(z_{t-1}^j)$ with probability ρ, and $\log(z_t^j) = \sigma \epsilon_t^j$ with probability

---

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

[10]One can think of the effective discount rate faced by the firm as the product of the usual time preference discount factor and a rescaling element which takes into account the time horizon of the decision maker, as for instance the average tenure of CEOs in the retail food industry reported in Henderson et al. (2006). This could also be modeled by a lower value of β.

[11]In Appendix C we show that moving from these assumptions we can derive a demand function ($d(p)$) and a customer surplus function ($v(p)$) consistent with the assumptions made in Section 2.

$1 - \rho$; where ϵ is i.i.d. and distributed according to a standardized normal. In the numerical solution of the model we approximate the normal distribution on a finite grid, using the procedure described in Tauchen (1986). Finally, we set the nominal wage equal to the price of the numeraire good, so that $w = 1$. This is equivalent to assume that the numeraire good n is produced by a competitive representative firm with linear production function and unitary labor productivity.

Search cost distribution. We assume that the search cost is drawn from a Gamma distribution with shape parameter ζ, and scale parameter λ. The Gamma distribution is appealing because it is flexible and fits the assumptions we made over the G function in the specification of the model. In particular, we focus on parameter values of $\zeta > 1$, so that the distribution of search costs is differentiable at $\psi = 0$; in our estimates, this restriction is not binding.

5.1 Identification and estimation

We aim to estimate the persistence and volatility of the productivity process (ρ and σ), and the scale and shape parameters of the search cost distribution (λ and ζ). Below we discuss the intuition behind the source of identification for each one of them.

We estimate persistence and volatility by matching the autocorrelation and the volatility of log-prices predicted by the model to the posted price measured using the store-level prices provided by the IRI data. We construct the posted price of a store in a particular week is the revenue weighted average of the prices of all the UPC in stock at the store. The data implies an autocorrelation of log-prices equal to 0.7, and a volatility of 0.06.[12]

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 4. We identify the scale parameter λ by matching the average effect of log-prices on the exit probability predicted by the model in equilibrium to its counterpart in the data measured by the parameter b_1 in equation (10). The model predicts that the ex-ante probability (before drawing the search cost) of exiting the customer base of a firm charging p and with productivity z is $\tilde{G}(p, z) \equiv G(\hat{\psi}(p, z))(1 - H(\bar{V}(p, z)))$. In the region of parameters we study, the marginal effect of prices on the probability of exiting, - i.e. $E[\partial \tilde{G}(p, z)/\partial \log(p) \mid_{p=\hat{p}(z)}]$- is decreasing in the scale of the search cost, creating a mapping between the mean of the search

---

[12]These statistics are obtained from fitting an AR(1) process to the time series of prices separately for each stores for which we have store level price data, $\log(p_t^s) = \rho^s \log(p_{t-1}^s) + \sigma^s \epsilon_t^s$. This step delivers 126 estimates of the persistence parameters ρ^s and of the volatility of the residuals σ^s. We then take the median across estimates for each store.

cost and the average elasticity.[13] We target the coefficient estimated in the specification of column (2) in Table 1, i.e. $b_1 = 0.2$.

The parameter ζ measures the inverse of the coefficient of variation of the search cost distribution. In the model, higher dispersion of search costs (i.e. lower ζ) implies more mass on the tails of the distribution of search costs. The latter is associated to larger variation in the sensitivity of exit probability to price. In the data, we measure this variation by fitting a spline to equation (10), allowing the price marginal effect on the probability of exit to vary for different terciles of price levels. We find that the difference in the estimate dispersion of b_1 is 0.04, with higher prices commanding a higher value of b_1 as predicted by the model. The parameter ζ is estimated by matching this number to an equivalent statistic generated by the model.[14]

We define $\Omega \equiv [\zeta \ \lambda \ \rho \ \sigma]'$ 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. Each iteration n of the estimation procedure unfolds according to the following steps:

1. Pick values for the parameters ρ_n, σ_n, λ_n, ζ_n,

2. Solve the model and obtain the vector $v(\Omega_n)$,

3. Evaluate the objective function $(v_d - v(\Omega_n))'\,(v_d - v(\Omega_n))$.

We select as estimates the parameter values that minimize the objective function.

Implementing step 2 requires solving a fixed point problem in equilibrium prices $\mathcal{P}(z)$ for all $z \in [\underline{z}, \bar{z}]$. In particular, given our definition of equilibrium and the results of Proposition 2, we look for equilibria where $\mathcal{P}(z) \in [p^*(\bar{z}), p^*(\underline{z})]$ for each z, and $\mathcal{P}(z)$ is strictly decreasing in z. In principle, our model could have multiple equilibria. However, numerically we always converge to the same equilibrium. In Appendix D we provide more details on the numerical solution of the model.

The estimates from this procedure are summarized in Table 2.

---

[13]Notice that $\partial \tilde{G}(p,z)/\partial p = -v'(p)\,[G'(\hat{\psi}(p,z))(1 - H(\bar{V}(p,z)))^2 + G(\hat{\psi}(p,z))H'(\bar{V}(p,z))]$. The parameter λ directly affects G' and G. In particular, for given equilibrium prices, an increase in λ is associated to a decrease in $G(\cdot)$ for all ψ, and to a decrease in $G'(\cdot)$ for all ψ small enough given $\zeta > 1$. Finally, notice that we are targeting the persistence and volatility of the empirical price distribution, which are indeed fixed in our analysis. This implies that $\partial \tilde{G}(p,z)/\partial p$ is decreasing in λ for values of $\hat{\psi}$ small enough.

[14]In the model, we construct the equivalent statistic as follows: if $z_{(1)}$ is the first and $z_{(2)}$ the second tercile of the estimated productivity distribution in the model, we compute $E[\partial \tilde{G}(p,z)/\partial \log(p)\,|_{p=\hat{p}(z),\, z \geq z_{(2)}}] - E[\partial \tilde{G}(p,z)/\partial \log(p)\,|_{p=\hat{p}(z),\, z \leq z_{(1)}}]$.

Table 2: Parameter estimates

|	Value	Target
Persistence of productivity innovations, ρ	0.7	Log-price autocorrelation: 0.7
Volatility of productivity innovations, σ	0.09	Log-price dispersion: 0.06
Distribution of cost, $g(\psi) \sim Gamma(\zeta, \lambda)$		
Shape parameter, ζ	2	Dispersion in marginal effect: 0.04
Scale parameter, λ	0.3	Average marginal effect: 0.2

5.2 Quantitative analysis of the model

We use the estimates obtained in the previous section to illustrate the quantitative implications of the model on several objects of interest. We begin by reporting on the relationship between customer base growth and pricing behavior with idiosyncratic productivity. We then analyze how the presence of customer markets affects prices and markups dispersion.

The idea that firms are endowed with a set of customers that they try to preserve is at the core of our model. In Figure 1 we display the annualized net growth rate of the customer base as a function of the production cost (i.e. the inverse of productivity z). Production cost influences dynamic in the customer base of a firm as it determines the price a firm can charge in the current period and signals its future prices. This type of relationship has been observed by Foster et al. (2012) who show that modeling the accumulation of the stock of demand idiosyncratic to a firm as a function of its price history helps explaining differences in growth between incumbent firms and new entrants.

Firms with low cost experience positive net growth of their customers base; whereas high cost firms are net loosers of customers. Net customer base growth declines in cost at an increasing pace: for firms in the right tail of the cost distribution it is very costly to experience even a marginal increase in their production cost. However, these instances are rare as the tails of the cost distribution are thin.

Figure 2 displays the equilibrium markups in our model (henceforth "Baseline economy") as a function of a firm's cost. It also relates them to the benchmark of an "Inelastic customer base economy". The latter is obtained by letting search costs diverge to infinity so that customers will never want to search for a new firm and will be tied to the firm they are initially matched with. To make the comparison meaningfull, we fix θ so that the resulting average total elasticity of demand (i.e., $\int_{\underline{z}}^{\bar{z}} \varepsilon_q(\hat{\mathcal{P}}(z), z) f(z) dz$) is the same as in our Baseline

Figure 1: Firm growth and productivity
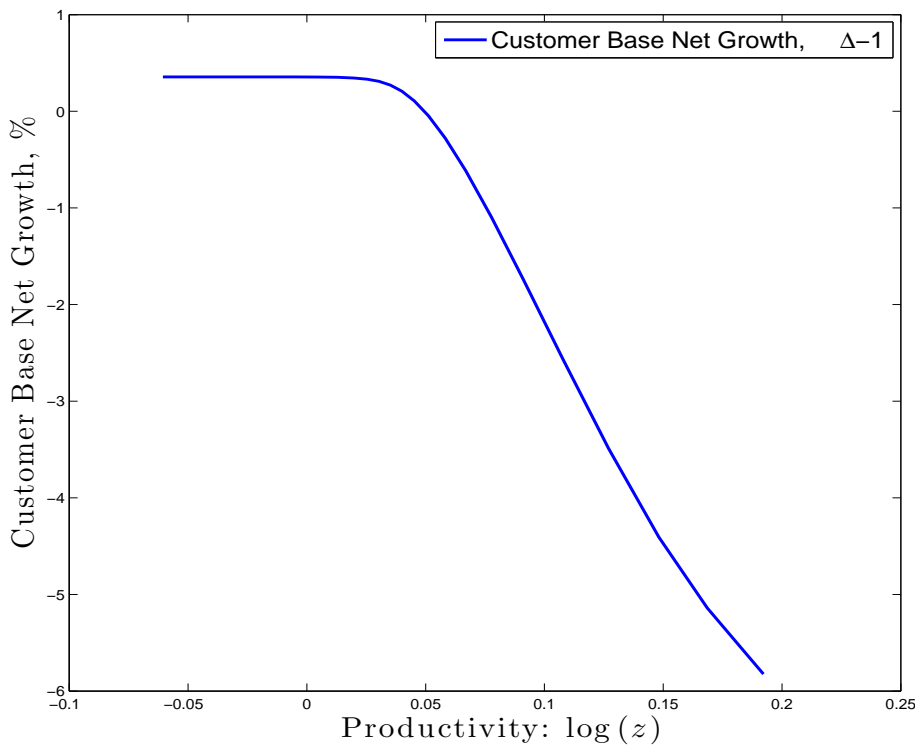
economy. It is worth noting that this alternative model is analogous to the standard model of monopolistic competition widely used in the macroeconomics literature.

Markups are strictly decreasing in production cost in both economies. In fact, in both models the intensive elasticity of demand, $\varepsilon_d(p)$, is increasing in p, giving rise to a negative co-movement between markups and production cost.[15] Since equilibrium prices are monotonically increasing in production cost, it follows that firms with higher production cost face higher elasticity of demand, so that optimal markups are decreasing in production cost.

However, the presence of customer base concerns causes markups to decrease more steeply in the Baseline economy. With a positive extensive margin elasticity, increases in price lead to the loss of customers on top of contraction in the quantity sold to retained customers. This results in an extra incentive to compress prices, which is stronger the higher the extensive margin elasticity faced by the firm and causes the average markup to be lower in the Baseline

---

[15]With CES preferences the demand of good i depends on the relative price p_i/P. With a finite number of goods in the basket of the customer, an increase in p_i, also increases the price of the basket, P, thus reducing the overall increase in p_i/P and effect on demand. The effect on P is larger, the higher the weight of good i in the basket, that is the lower the price p_i and the higher its demand. Therefore, the elasticity of demand $\varepsilon_d(p)$ increases in p.

than in the inelastic customer base economy (15% vs. 21%). In fact, highly productive firms deliver such high a value to their customers that they will never want to leave them; as a result those firms can act as if they were in the inelastic customer base economy. Less productive firms, instead face an actual risk of losing customers if they decide to charge higher prices; therefore the have to set markups lower than those they would choose in absence of an elastic customer base. This implies that our model delivers markups that are pro-cyclical with respect to productivity shocks; firms gaining production efficiency can set higher markups. The average elasticity of markups to productivity shocks generate by our model is 0.74, which is consistent with the empirical evidence provided by Petrin and Warzynski (2012) using firm level data. In the inelastic customer base economy, the average elasticity of markups to productivity is only 0.21.

Figure 2: Markup and productivity
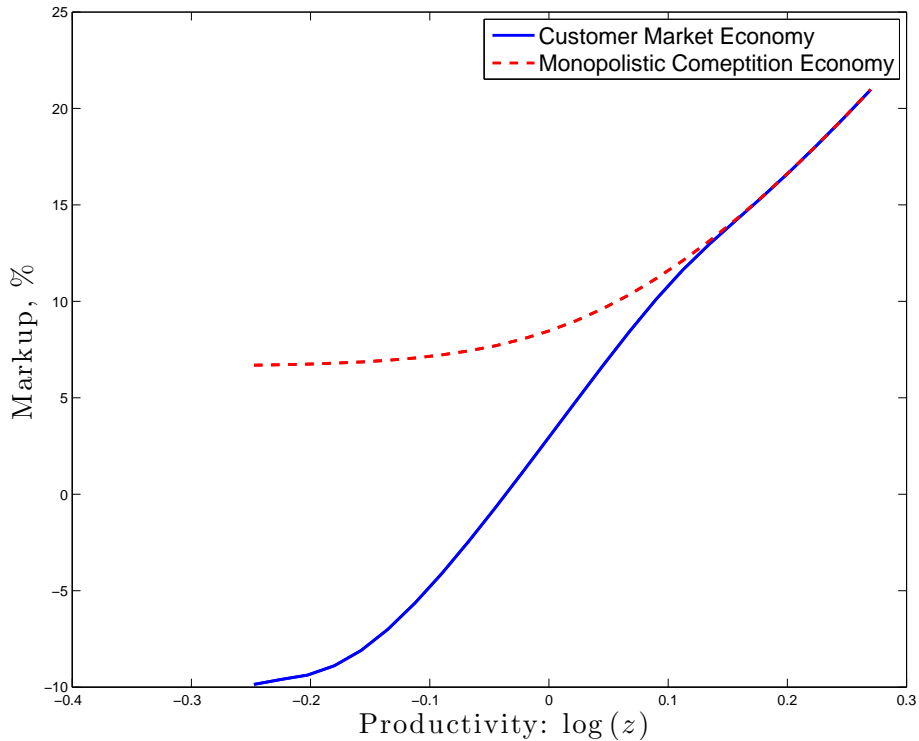
A direct consequence of the presence of a sticky customer base is a substantial increase in the dispersion of markups and a reduction in that of prices with respect to the inelastic customer base economy, as illustrated in Figure 3. The standard deviation of markups rises from 1.5% to 5.6%; the standard deviation of prices shrink from 6% to 2%. The two features are obviously related. The presence of dynamic concerns leads the firms to reduce their

23

Figure 3: The distributions of markups and prices
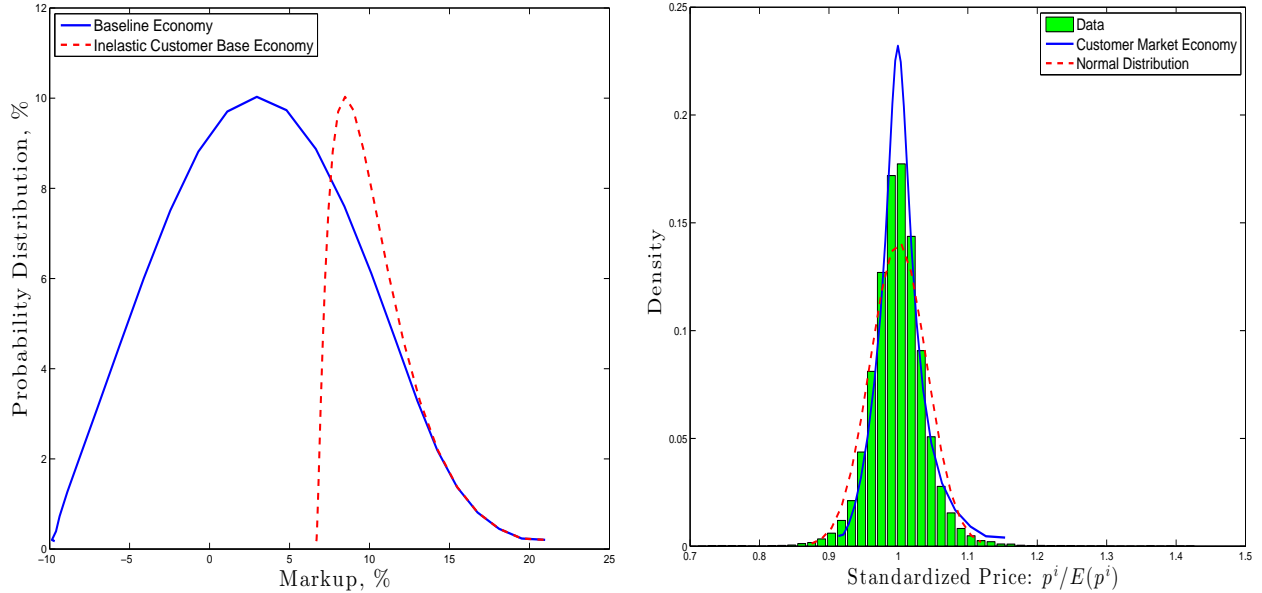
markups exactly to avoid changing their price. Therefore, changes in cost process result contractions and expansions of markups and contribute much less to the volatility of the posted prices.

It is interesting to notice that the distributions of markups in the baseline economy has positive, though tiny, mass on negative markups. As we have seen, firms with very high production cost can experience severe customer losses. Therefore, they are willing make temporary negative profits in order to keep customers around and charge higher markups when the production cost mean reverts.[16] In the inelastic customer base economy negative markups cannot occur, as a firm can always decide not to produce without any real impact on future demand.

On top of providing insights when compared to that arising from a model without customer concerns, the distribution of log-prices in the baseline economy displays some interesting features of its own. The price distribution generate by the model is leptokurtic. In particular the kurtosis of the price distribution in the baseline economy is much higher than that arising in the inelastic customer base scenario (8.2 vs 2.9) and matches closely the statistic documented by Kaplan and Menzio (2014) for the price distribution of homogeneous packaged goods. The presence of customer markets reduces dispersion of prices, as less

---

[16]Note that despite the temporary negative profits, the value of these firms is strictly positive.

productive firms charge lower price than they would otherwise in an economy with inelastic customer base in order to retain their customers, with the consequence that a large mass of prices concentrates around a lower mean. At the same time, a non-negligible fraction of firms set prices relatively far away from this mean. Part of them (very low productivity firms) charge a higher price because, given the persistence of their status, they find much costly to retain customers and settle for a price that is lower than the static profit maximization but still substantially above the mean. Another set of firms (very high productivity firms) charge a price substantially below the mean simply because they enjoy a productivity advantage which makes such lower price optimal.

6   Customer markets and propagation of nominal shocks

The macroeconomic literature has highlighted the role of real rigidities in increasing the persistence of inflation and output response to aggregate nominal shocks (see, for instance, Chari et al. (2000), Klenow and Willis (2006), or Gopinath and Itskhoki (2011)). A number of studies, such as Rotemberg and Woodford (1991), or Blanchard (2009), have pointed to customer markets as a natural source of real rigidities, not least due to the observation that firms do behave consistently with the idea that they perceive customers as assets. This calls for an assessment of the effect that this type of real rigidity can have on the propagation of nominal shocks and we are uniquely well positioned to provide it. First, we extend the early work of Phelps and Winter (1970) by allowing for the dynamics of customers to be endogenously determined, and therefore to respond to shocks. This is important because customer dynamics affects both the extensive demand elasticity and the relative value of a customer in different ways, something that could not be simply picked up using a model that posts a reduced-form equation for this margin. Furthermore, we exploit micro data to discipline our exercise and ground our quantification of the impact of customer dynamics on markups dynamics.

A nominal shock is represented in our economy as an unexpected, permanent innovation in the price of the numeraire good q. We consider the economy in its steady state at $t = t_0$ and hit it with an unforeseen increase in q so that q_{t_0} jumps from a steady state value of 1 to the new value of $1 + \delta$. We study the transition of the economy to the new steady state.

In order to consider the general equilibrium effects of the nominal shock on wages, income and stochastic discount factors, we need to implement a series of extensions to our model. In particular, we introduce perfectly competitive labor markets through a representative household, as well as allow for dynamics in the aggregate state. These extensions are standard and details are given in Appendix E.

In our standard environment nominal shocks would not have any effect on real variables (consumption, labor, search decisions, etc.). It is well known that, in absence of nominal rigidities, firms would completely pass-through the increase in nominal marginal cost, with no effect on their demand and customer base.[17] Therefore, we need a simple and tractable way to introduce nominal rigidities in our model if we want to use it to study the propagation of nominal shocks. To do so, we assume that firms are not perfectly informed about the realization of the aggregate shocks. At each point in time, each firm has a probability $\alpha \in (0, 1]$ to become aware that the nominal shock has occurred. We set $\alpha = 0.1$, implying that on average it takes roughly a quarter for a firm to realize that the shock has realized.[18] This information friction causes that, even though all firms are allowed to adjust the price in every period, a fraction of them will behave as the aggregate shock did not occur, meaning that their price will not respond to it. We can interpret the friction in the spirit of the rational inattentiveness literature that developed after the work of Mankiw and Reis (2002), where firms review infrequently the aggregate state. For simplicity, and as it is typically assumed in this literature, we assume that customers are perfectly informed.[19]

Figure 4 plots the response of aggregate output to a nominal shocks of size $\delta = 5\%$ in our baseline economy with customer markets as well as in an alternative economy with inelastic customer base. The response of output is sizable, and is magnified by the presence customer markets: the cumulated output response (i.e. the area under the impulse response) is 3.75 times larger in our baseline economy than in the economy with inelastic customer base. The half life of the output response is 33 weeks in our baseline model, against 10 weeks in the alternative economy.

In order to understand the causes behind the magnification effect induced by customer markets, we study the response of markups for firms active in the locally produced good d, which is where customer markets matter. In the competitive sector (good n) markups are constant and equal to zero. Figure 5 shows that markups decline more on impact and recover much more slowly in our baseline economy than in the alternative scenario where customer markets are absent. The lower markups stimulate demand resulting in a boost for aggregate output and employment.

There are in turn two reasons explaining why markups are persistently below their steady

---

[17]Perfectly competitive labor markets, together with a perfectly competitive sector producing good n imply that the equilibrium wage is $w_t = q_t$, and thus moves one for one with the numeraire.

[18]Cross-country evidence from survey data places the frequency of information acquisition by firms between 2 and 4 times per year. See Fabiani et al. (2007) for a review.

[19]An alternative specification could be one where firms adjust prices infrequency with a Calvo type lottery. While it would not alter the qualitative conclusions of the experiment of this section, that environment would however change the steady state pricing problem of the firm substantially, as it would also affect the response of prices to idiosyncratic shocks.

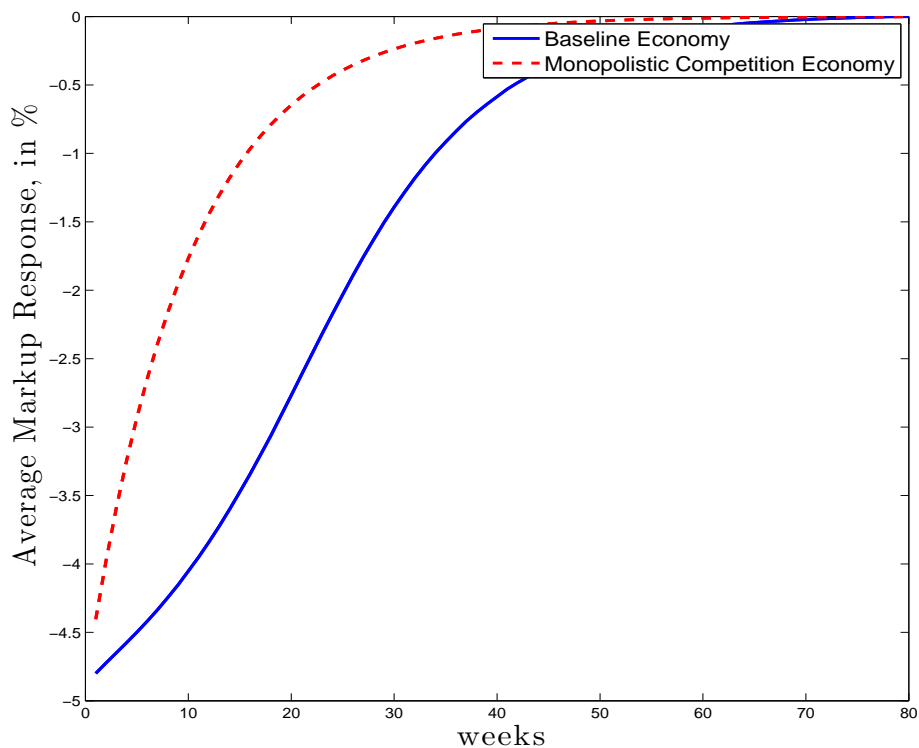Figure 4: The response of aggregate output to a 5% nominal shock
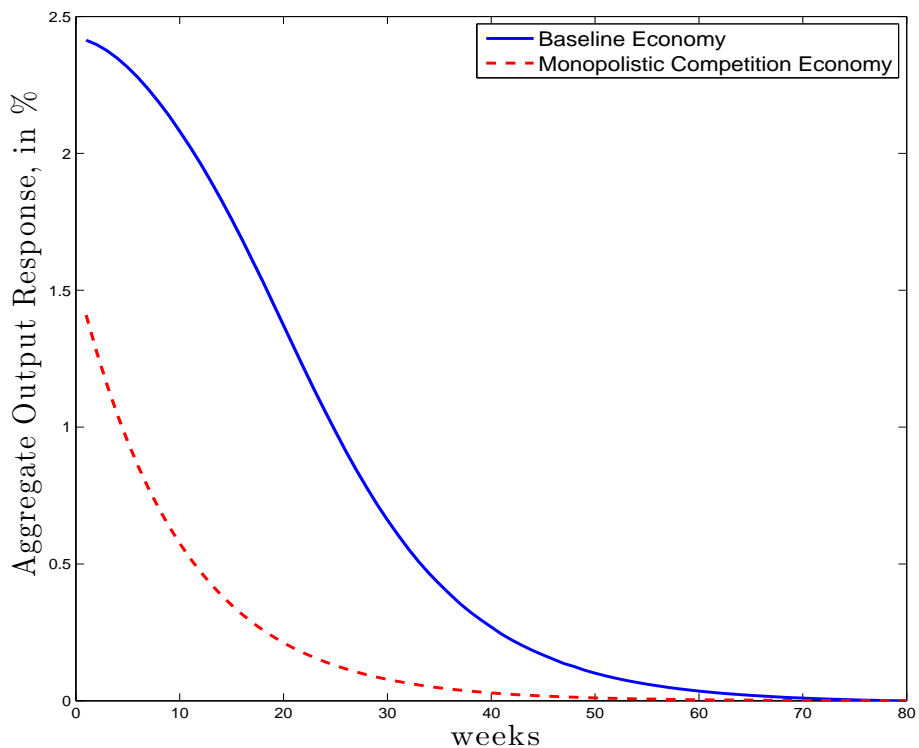
The vertical axis refers to % deviations from the steady state of aggregate output after a 5% permanent increase in q. The blue solid line refers to our baseline model with at parameter values estimated in Section 5. The red dashed line refers to the model without customer markets, i.e. the same parameter estimates but $\lambda \to \infty$.

state value in our baseline economy. The first one is mechanical: it takes time for firms to become informed about the realization of the aggregate shock. Firms unaware of the shock realization cannot react to it. A second reason arises from competition for customers. Customer markets introduce a strong element of strategic complementarity in price setting. Firms that learned about the nominal shock should react raising their price; however, they know that a fraction of their competitors will not do so because they are not aware of this event. The result is that, on average, even informed firms will have to respond only partially in order to avoid losing customers.

In the economy with inelastic customer base only the first effect operates. The optimal markup of each firm is unresponsive to the aggregate shock, as in standard CES economies firms fully pass-through the increase in nominal marginal cost. The only reason average markup does not immediately adjust is that some firms have not yet learned about the nominal shocks and, therefore, have not raised their markup by a factor of δ. In contrast, in

27

Figure 5: The response of average markup to a 5% nominal shock

The vertical axis refers to % deviations from the steady state of average markup in the sector facing customer retention concerns (good d), after a 5% permanent increase in q. The blue solid line refers to our baseline model with at parameter values estimated in Section 5. The red dashed line refers to the model without customer markets, i.e. the same average total elasticity, but with $\lambda \to \infty$.

our economy the optimal markup of each firm is not constant in response to the aggregate shock. As a consequence of competition for customers, firms that have the chance to respond to the aggregate shock are on average worsening their position with respect to the ones that do not (which keep their price at sub-optimally low level). Therefore they will not want to fully pass-through the increase in nominal marginal cost, amplifying the persistence of the response to the shock.

Finally, notice that the response of output (average markup) to the nominal shock in our baseline economy with customer markets displays substantial persistence. This occurs because the strength of the strategic complementarity effect increases in the first few weeks after the nominal shock, thus pushing towards lower markups. Figure 6 shows the source of the transitory strengthening of the strategic complementarity after the nominal shock: the equilibrium mass of customers searching for a new match increases on impact and keep growing until peaking after about a quarter, before reverting to steady state. A larger mass

Figure 6: Impulse response of searching customers
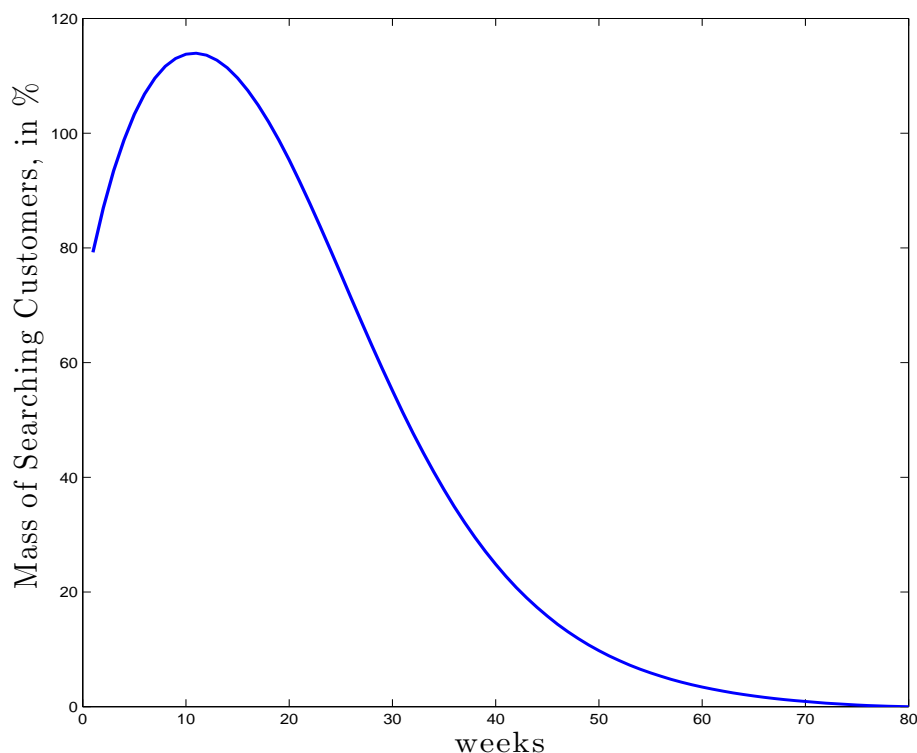
The figure plots the response of the mass of customers searching for a new firm in % deviation from the steady state.

of searching customers derives from the larger benefit from search due to the presence of a fraction of firms unaware of the shock, which are keeping prices lower than steady state. The hump-shaped response of the mass of searching customers is due to two competing forces. First notice that the benefit from search is on average higher for customers matched to aware than unaware firms. As time goes by two things happen. On the one side, more firms become aware of the shock and react to it, so that the pool of unaware firms reduces and the probability of drawing an unaware firm goes down. This reduces on average the benefit from search of a customer matched to aware firms. For given mass of customers matched to aware firms, this effect pushes towards a reduction in the mass of searching customers. On the other side, the mass of customers matched to aware firms increase mechanically because a fraction α of unaware firms become aware every period. Given that customers of aware firms search on average more than customers of unaware firms, this effect pushes towards an increase in the total mass of customers searching.

7   Conclusions

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

We use scanner data on households' purchases at a U.S. supermarket chain to provide direct evidence that customers do respond to variation in the price of their consumption basket. We also exploit the data to estimate the key parameters of the model and provide a quantification of the effect of customer retention concerns on firm pricing. We use the estimated model to gauge the role of customer markets in the propagation of nominal shocks, showing that they can greatly amplifying the real effects of nominal shocks.

References

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

Alessandria, G. (2009). Consumer search, price dispersion, and international relative price fluctuations, *International Economic Review* **50**(3): 803–829.

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

Bils, M. (1989). Pricing in a customer market, *Quarterly Journal of Economics* **104**(4): 699–718.

Blanchard, O. (2009). The state of macro, *Annual Review of Economics* **1**(1): 209–228.

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

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

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment, *International Economic Review* **39**(2): 257–73.

Chari, V. V., Kehoe, P. J. and McGrattan, E. R. (2000). Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?, *Econometrica* **68**(5): 1151–1180.

Coibion, O., Gorodnichenko, Y. and Hong, G. H. (2012). The cyclicality of sales, regular and effective prices: Business cycle and policy implications, *IMF Working Papers 12/207*, International Monetary Fund.

Coles, M. G. (2001). Equilibrium wage dispersion, firm size and growth, *Review of Economic Dynamics* **4**(1): 159–187.

Coles, M. G. and Mortensen, D. T. (2013). Equilibrium labor turnover, firm growth and unemployment, *Working paper*.

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

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

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

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.

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

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.

Gopinath, G. and Itskhoki, O. (2011). In search of real rigidities, *NBER Macroeconomics Annual 2010, Volume 25*, NBER Chapters, National Bureau of Economic Research, Inc, pp. 261–309.

Gourio, F. and Rudanko, L. (2014). Customer capital1, *The Review of Economic Studies* .

Hall, R. E. (2012). The cyclical response of advertising refutes counter-cyclical profit margins in favor of product-market frictions, *Working Paper 18370*, National Bureau of Economic Research.

Henderson, A. D., Miller, D. and Hambrick, D. C. (2006). How quickly do CEOs become obsolete? Industry dynamism, CEO tenure, and company performance, *Strategic Management Journal* **27**: 447–460.

Kaplan, G. and Menzio, G. (2013). Shopping externalities and self-fulfilling unemployment fluctuations, *NBER Working Papers 18777*, National Bureau of Economic Research, Inc.

Kaplan, G. and Menzio, G. (2014). The morphology of price dispersion, *Working paper*, University of Pennsylvania.

Kimball, M. S. (1995). The Quantitative Analytics of the Basic Neomonetarist Model, *Journal of Money, Credit and Banking* **27**(4): 1241–77.

Klenow, P. J. and Willis, J. L. (2006). Real rigidities and nominal price changes, *Research working paper*, Federal Reserve Bank of Kansas City.

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

Luttmer, E. G. (2006). Consumer search and firm growth, *Technical report*.

Luttmer, E. G. (2010). Models of growth and firm heterogeneity, *Annual Review of Economics* **2**(1): 547–576.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve, *The Quarterly Journal of Economics* **117**(4): 1295–1328.

Menzio, G. (2007). A search theory of rigid prices, *PIER Working Paper Archive 07-031*, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

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

Petrin, A. and Warzynski, F. (2012). The impact of research and development on quality, productivity and welfare, *mimeo*.

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

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

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

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

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

A   Proofs

A.1   Proof of Lemma 1

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

A.2   Proof of Lemma 2

The proof of Lemma 2 follows from the assumption of $v(p)$ being strictly decreasing in p so that $\bar{V}(p, z)$ is decreasing in p; the threshold $\hat{z}(p, z)$ is increasing in z because of the assumptions that $\hat{V}(z)$ is increasing in z and the productivity process assumed to exhibit persistence, so that $\bar{V}(p, z)$ increases with z. Moreover, the assumption that $\hat{V}(z)$ is increasing in z also implies that $\bar{p}(z)$ is increasing in z. Notice that $\frac{\partial \hat{\psi}(p,z)}{\partial p} = -v'(p) + \frac{\partial \tilde{V}(p,z)}{\partial p} \geq 0$ by the definition of $\tilde{V}(p, z)$ and $v(p)$ being decreasing in p. Also, we have that

$$\frac{\partial \hat{\psi}(p, z)}{\partial z} = -\frac{\partial \bar{V}(p, z)}{\partial z} \left(1 - H(V(p, z))\right) \leq 0 \, ,$$

as $\bar{V}(p, z)$ is increasing in z if $\hat{V}(z)$ is increasing in z and the productivity process exhibits persistence.

A.3   Proof of Proposition 1

Let $\bar{p}(z)$ be the level of price at which no customer searches. Then $\bar{p}(z)$ satisfies $\bar{V}(\bar{p}(z), z) = \max_{z \in [\underline{z}, \bar{z}]} \{\bar{V}(\mathcal{P}(z), z)\}$. First, given the definition of $p^*(z)$ and the fact that $\Delta(p, z)$ is strictly decreasing in p for all $p > \bar{p}(z)$, and constant otherwise, it immediately follows that $\hat{p}(z) \in [\bar{p}(z), p^*(z)]$ if $\bar{p}(z) < p^*(z)$, and $\hat{p}(z) = p^*(z)$ otherwise. Next, we show that $\hat{p}(z) < p^*(z)$ if $\bar{p}(z) < p^*(z)$. The results follows because at $p = \hat{p}(z)$ $W(p, z)$ is strictly decreasing in p as by

definition of $\bar{p}(z)$, and assumptions about G, $\Delta(p, z)$ is strictly decreasing in p at $p = \hat{p}(z)$, so that $p = \hat{p}(z)$ cannot be a maximum.

We next prove necessity of equation (7) for an optimum. The latter follows immediately from the assumption of G and H differentiable, as $\bar{V}(p, z)$ is continuously differentiable in p given $v(p)$ is so, and $Q(x) = \Gamma \int_s^x G\left(\int_s^\infty (u - x)dH(u)\right) dH(u)$ is continuously differentiable in p. Thus, $\Delta(p, z)$ is continuously differentiable in p. Finally, $\Pi(p, z)$ is continuously differentiable in p because $d(p)$ has been assumed to be continuously differentiable. Thus, the objective of the firm problem is continuously differentiable in p, and the first order condition is necessary and sufficient.

Finally, we provide an expression for the extensive margin elasticity:

$$\varepsilon_m(p, z) = -\frac{v'(p)\,p}{\Delta(p, z)} \left[G'(\hat{\psi}(p, z)) \left(1 - H(\bar{V}(p, z))\right)^2 + 2\,G(\hat{\psi}(p, z))\,H'(\bar{V}(p, z)) \right] .$$

If $\varepsilon_m(p, z)$ is non-decreasing in p then the first order condition is also sufficient for an optimum. The term outside the square brackets is indeed increasing in p. The terms inside the square brackets are all increasing in p, but G' and H' about which may or may not be increasing in p. Thus a sufficient condition for $\varepsilon_m(p, z)$ to be non-decreasing in p for some range of p is that $G'' > 0$ and $H'' < 0$ at that range of p.

A.4 Proof of the thought experiment in Section 2.2

We show that $\mu(p, z)$ is increasing in $\varepsilon_q(p, z)$. Notice that equation (8) can be rewritten as

$$\mu(p, z) = \frac{\varepsilon_q(p, z) + \varepsilon_m(p, z)\tilde{x}(p, z)}{\varepsilon_q(p, z) - 1 + \varepsilon_m(p, z)\tilde{x}(p, z)},$$

where $\tilde{x}(p, z) \equiv \Pi(p, z)/\pi(p, z)$. From the equation above we obtain

$$\frac{\partial \mu(p, z)}{\partial \varepsilon_m(p, z)} = \frac{\tilde{x}(p, z)}{\varepsilon_q(p, z) - 1 + \varepsilon_m(p, z)\tilde{x}(p, z)}(1 - \mu(p, z)).$$

A direct implication of nonnegative prices is that $\varepsilon_q(p, z) - 1 + \varepsilon_m(p, z)\tilde{x}(p, z) \geq 0$, so that $\text{sign}\left[\partial \mu(p, z)/\partial \varepsilon_m(p, z)\right] = \text{sign}[(\tilde{x}(p, z))(1 - \mu(p, z))]$. There are two cases two consider. The first one is when $\pi(p, z) > 0$, which occurs if and only if $\mu(p, z) > 1$. It implies $\tilde{x}(p, z) > 0$ and, therefore, $\partial \mu(p, z)/\partial \varepsilon_m(p, z) < 0$. The second case is when $\pi(p, z) < 0$, which occurs if and only if $\mu(p, z) < 1$. It implies $\tilde{x}(p, z) < 0$ and, therefore, $\tilde{x}(p, z) < 0$. As a result, $\partial \mu(p, z)/\partial \varepsilon_m(p, z) < 0$.

A.5 Proof of Proposition 2

Monotonicity of prices. We first show that optimal prices $\hat{p}(z)$ are non-increasing in z. Given, that productivity is i.i.d. and we look for equilibria where $\hat{p}(z) \geq p^*(\bar{z})$, we have that $\bar{p}(z) = p^*(\bar{z})$ for each z. From Proposition 1 we know that, for a given z, the optimal price $\hat{p}(z)$ belongs to the set $[p^*(\bar{z}), \ p^*(z)]$. Over this set, the objective function of the firm,

$$W(p, z) = \Delta(p, z)\left(\pi(p, z) + \beta \text{ constant}\right), \tag{11}$$

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

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

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

Existence of equilibrium. Next we prove existence of an equilibrium. The fixed point problem is a mapping from candidate function of equilibrium prices, $\mathcal{P}(z)$, to firm's optimal pricing strategy, $\hat{p}(z)$, where an equilibrium is one where $\hat{p}(z) = \mathcal{P}(z)$ for each z. Given that $W(p, z)$ is continuously differentiable in p, the operator that maps $\mathcal{P}(\cdot)$ into $\hat{p}(\cdot)$ is given by the first order condition in equation (7). Moreover, notice that $W(p, z)$ in equation (11) is continuous in (p, z). By the theorem of maximum $\hat{p}(z)$ is upper hemi-continuos and $W(\hat{p}(z), z)$ is continuos in z. Given that $\hat{p}(z)$ is non-increasing in z it follows that $\hat{p}(z)$ has a countably many discontinuity points. We proceed as follows. Let $\hat{\mathcal{P}}(z)$ be the set of prices that maximize the firm problem. Whenever a discontinuity arises at some $\tilde{z}$ (so that $\hat{\mathcal{P}}(\tilde{z})$ is not a singleton) we modify the optimal pricing rule of the firm and consider the convex hull of the $\hat{\mathcal{P}}(\tilde{z})$ as the set of possible prices chosen by the firm with productivity $\tilde{z}$. The constructed mapping from z to $\hat{\mathcal{P}}(z)$ is then upper-hemicontinous, compact and convex valued. We then apply Kakutani fixed point theorem to this operator and obtain a fixed point. Finally, notice that since the convexification procedure described above has to be applied only a countable number of times, the convexified prices have measure zero with respect to the density of the z. Hence,

they do not affect the fixed point.

Necessity of the first order condition. We show that Q and H are almost everywhere differentiable, so that Proposition 1 implies that equation (7) is necessary for an optimum. We guess that $\hat{p}(z)$ is strictly decreasing and almost everywhere differentiable. It immediately follows that $\hat{V}(z)$ is strictly increasing in z and almost everywhere differentiable. Then, given the assumption that F is differentiable, we have that K is differentiable. From $H(x) = K(\hat{V}^{-1}(x))$ it follows that H is also almost everywhere differentiable. Given that G and H are differentiable, so is Q. Then the first order condition in equation (7) is necessary for an optimum, which indeed implies that $\hat{p}(z)$ is strictly increasing and differentiable in z in any neighbor of the first order condition. Finally, given that $\hat{p}(z)$ has a countably many discontinuity points, it has countably many points where it is not differentiable, and the first order condition does not apply at those points, but applies everywhere else. These points have measure zero with respect to the density of the z and therefore $\hat{p}(z)$ is almost everywhere differentiable.

Points (i)-(iv). We first prove part (i). We already proved that $\hat{p}(z)$ is non-increasing in z. The proof that $\hat{p}(z)$ is strictly decreasing in z for some region of z is by contradiction. Suppose that $\hat{p}(z)$ is everywhere constant in z at some $\tilde{p}$. Then $\bar{p}(z) = \tilde{p}$ for all z. If $\tilde{p} > p^*(\bar{z})$, then $\tilde{p}$ would not be optimal for firm with productivity $\bar{z}$ which would choose a lower price. If $\tilde{p} = p^*(\bar{z})$, then continuous differentiability of G together with $H = G = Q = 0$ at the conjectures constant equilibrium price, imply that the first order condition is locally necessary for an optimum, and a firm with productivity $z < \bar{z}$ would have an incentive to deviate according to equation (7), and set a strictly higher price than $\tilde{p}$. Finally, the result that $\hat{p}(z) < p^*(z)$ for all $z < \bar{z}$ and that $\hat{p}(\bar{z}) = p^*(\bar{z})$ follows from applying Proposition 1, and using that $\hat{p}(z) \geq \hat{p}(\bar{z})$ and $\bar{p}(z) = \hat{p}(\bar{z})$ for all z.

We next prove part (ii); monotonicity of prices, $v'(p) < 0$, i.i.d. productivity, and application of the contraction mapping theorem ensure that $\hat{V}(z) = \bar{V}(\hat{p}(z), z)$ is increasing in z, strictly so on some interval of productivities. Given $\hat{p}(z) \geq \hat{p}(\bar{z})$, then $\bar{V}(\hat{p}(z), z) \leq \bar{V}(\hat{p}(\bar{z}), \bar{z})$.

We now prove part (iii); $\hat{\psi}(p, z) \geq 0$ immediately follows its definition. Lemma 2 together with point (i) readily imply that $\hat{\psi}(p, z) > 0$ for some interval of z;

Finally, part (iv) follows immediately from application of Lemma 3 which ensures that $\Delta(p, z)$ is strictly decreasing in p, which coupled with point (i), and the fact that productivity is i.i.d., implies that $\Delta(\hat{p}(z), z)$ is increasing in z, strictly so for some range of productivity. Because of price dispersion, some customers are searching in the economy, and because $\hat{p}(\bar{z}) \leq \hat{p}(z)$ for all z (strictly so for some z) we have that $\Delta(\hat{p}(\bar{z}), \bar{z}) > 1$. Similarly, given prices are monotonic, $\hat{p}(\underline{z}) \geq \hat{p}(z)$ for all z (strictly so for some z) we have that $\Delta(\hat{p}(\underline{z}), z) < 1$ for some z.

A.6   Proof of Remark 1

Part (1) of the remark follows since, given that the search cost diverges, firms do not face customer base concerns and therefore find it optimal to charge the price that maximizes static profits. Formally, when $G(\psi) = 0$ for all $\psi \to \infty$, we have that $\bar{p}(z) \to \infty$, so that $p^*(z) \leq \bar{p}(z)$ for all z. Then, $\hat{p}(z) = p^*(z)$.

We next provide a proof of the part (2) of the remark. In a Markovian equilibrium where each firm has the some constant productivity z_0, every firm has to choose the same price. We prove that this price must be $p^*(z_0)$. We first show that $\hat{p}(z_0) = p^*(z_0)$ can be an equilibrium. Conjecture that firms choose $\hat{p}(z_0) = p^*(z_0)$, implying that in equilibrium $\bar{p}(z_0) = \hat{p}(z_0) = p^*(z_0)$. A direct application of Proposition 1 then indeed implies that firms would choose $\hat{p}(z_0) = p^*(z_0)$, validating the conjecture.

We next show that $\hat{p}(z_0) = p^*(z_0)$ is the unique equilibrium. Consider two cases. First, the case where $\hat{p}(z_0) > p^*(z_0)$. In equilibrium this would imply $\bar{p}(z_0) = \hat{p}(z_0) > p^*(z_0)$. However, Proposition 1 implies that firms should deviate and choose $\hat{p}(z_0) = p^*(z_0)$ if $\bar{p}(z_0) > p^*(z_0)$. Therefore, $\hat{p}(z_0) > p^*(z_0)$ cannot happen in equilibrium.

Next, conjecture that $\hat{p}(z_0) < p^*(z_0)$. In equilibrium this would imply $\bar{p}(z_0) = \hat{p}(z_0) < p^*(z_0)$. We want to show that $\hat{p}(z_0) = \bar{p}(z_0)$ cannot be optimal for the firm. From equation (4) we now that in equilibrium the change in customer base triggered by an arbitrarily small increase in the price is approximately zero because no customer searches ($\hat{\psi}(\bar{p}(z_0), z_0) = 0$) and $G'(0) = 0$ by assumption. Moreover, in this equilibrium $Q = 0$ and is insensitive to prices as no other customer leaves other firms in the economy, while H is bounded. On the other side, the increase in the price has a strictly positive effect on profits per customer as $\hat{p}(z_0) < p^*(z_0)$. Therefore, this deviation is optimal for the firm and $\hat{p}(z_0) < p^*(z_0)$ cannot be an equilibrium.

B   Data sources and variables construction

B.1   Data sources

The empirical evidence presented in Section 4 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 revenues and quantities 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 memebr of the family to a single household identifier. Purchases made without using the card are not recorded. However, the chain ensures that the loyalty card has a high penetration by keeping to a minimum the effort needed to register for one. Furthermore, nearly all promotional discount are tied to ownership of a loyalty card, which provides a strong incentive to use it. Another potential drawback of the data is that we only follow households when purchasing at stores of a single, albeit large, supermarket chain. Other data sources on the same industry, like the Nielsen Homescan database, rely on households themselves scanning the barcodes of the items purchased once they return home after a trip and can therefore track them shopping at a plurality of competing firms. On the other hand, cash register data contain significantly less measurement error than databases relying on home scanning (Einav et al. (2010)).

B.2 Variables construction

Exit from customer base. The dependent variable in the regression presented in equation (10) 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 just consuming its inventory. In fact, we observe households when they buy grocery at the chain but do not have any information on their shopping at competing grocers. To circumvent this problem, we focus on a subsample of households who shop frequently at the chain. For them we can plausibly assume that sudden long spells without trips represent instances in which the household has left the chain and is fulfilling grocery needs shopping at one of its competitors. Operationally, we select households who made at least 48 trips at the chain over the two years spanned in the sample, implying that they would shop on average twice per month at the chain. When such households do not visit any supermarket store of the chain over at least eight consecutive weeks, we assume that the customer is shopping elsewhere. The *Exit* dummy is constructed so that it takes value of one in correspondence to the last visit at the chain before a spell of eight or more weeks without shopping there. Table 3 summarizes shopping behavior for households in our sample. It is immediate to notice that a 8-weeks spell without purchase is unusual,

as customers tend 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.

Table 3: Descriptive statistics on customer shopping behavior

	Mean	Std.dev.	25th pctile	75th pctile
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		

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

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

where U^i is the set of all the UPC's (u) purchased by household i during the sample period, p_{ut} is the price of a given UPC u in week t and the w_{iu}'s are a set of household-UPC specific weights.

We choose to calculate the weights using the total expenditure in the UPC by the household over the two years in the sample. This can lead to some inaccuracy in identifying the goods the customer cares for at a given point in time. For example, if a customer bough 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 par-

ticular UPC may induce the customer to buy it just because of the temporary convenience; this would give the UPC a high weight in the month. The effect of promotion is instead smoothed when we compute weights using expenditure over the entire sample period. Table Table 4 reports descriptive statistics on the change in price of the basket.

Our definition of the household basket implies that it is possible that a household does not purchase every week each UPC included in its basket. This raises a practical problem since 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 we would be unable to construct the basket price for that week. The issue can be solved using the store level data which allow us to calculate unit value prices every week for every item in stock in a given store, whether or not that particular UPC was bought by one of the households in our data. Unit value prices are computed using data on revenues and quantities sold as

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

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

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

It is important to notice that the retail chain sets different prices for the same UPC in different geographic areas, called "price areas". The retailer supplied store level information for 270 stores, ensuring that we have data for at least one store for each price area. In order to use unit value prices calculated from store level data to compute the price of the basket of a specific household, we need to determine to which price area the store(s) at which she regularly shop belong. This information is not supplied by the retailer that kept the exact

41

definition of the price areas confidential. A possible solution is to infer in which price areas the store(s) visited by a household are located by comparing the prices contained in the household panel with those in the store data. In principle the household data should give information on enough UPC prices in a given week to identify the price area representative store whose pricing they are matching. However, even though two stores belonging in the same price area should have the same prices, they may not have the same unit value prices if the share of shoppers using the loyalty card differs in the two stores. Therefore, we choose to restrict our analysis to the set of customers shopping predominantly in one of the 270 stores for which the chain provided complete store level data. Since the 270 representative stores are not selected following any particular criterion, the resulting subsample of households should not subject to any type of selection. However, this choice is costly in terms of sample size: only 1,336 households shop at one of the 270 stores for which we have data; by far the biggest loss of observation imposed by our data cleaning procedures.

Table 4: <u>Descriptive statistics on basket price changes</u>

	Mean	Std.dev.
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	

C   Parametrization of customer demand and utility

In this section we propose a microfounded model that can give rise to the assumptions we made on the paper regarding customer's demand $d(p)$ and surplus $v(p)$. We also check whether the profit function is single-peaked, i.e. whether $p^*(z)$ exists and it is unique.

We consider a setup where customers derive utility from the consumption of $K > 1$ different varieties, the consumption of each variety described by c_k. The different varieties are aggregated to produce aggregate consumption C by using a CES aggregator, $C \equiv \left(\sum_{k=1}^{K} c_k^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}$, $\theta > 1$. Finally, customers derive utility from the consumption level C by $u(C) = C^{1-\gamma}/(1-\gamma)$, where $\gamma > 1$. Notice that, through the lens of our model, we can interpret $d(p_k) \equiv c_k(p_k)$ and $v(p_k) \equiv u(C(p_k))$, where p_k denotes the price of variety k.

Given that customers are not allowed to save, the problem they face is a static one. Each period the customer maximizes

$$\max_{\{c_k\}_{k=1}^K} u(C) \text{ subject to } \sum_{k=1}^K p_k c_k = I .$$

Operating with the first order conditions provide that the demand for variety k is given by $c_k(p_k) = \frac{I}{P} \left(\frac{p_k}{P}\right)^{-\theta}$, where $P \equiv \left(\sum_{k=1}^K p_k^{1-\theta}\right)^{\frac{1}{1-\theta}}$ is a price level that solves $P\,C = I$.

We start by evaluating the properties of demand, $c_k(p_k)$, continued by the evaluation of the indirect utility function $u(C(p_k))$. We then evaluate the properties of the demand elasticity, and we finish the section by evaluating the properties of firm's profits.

It proves useful to characterize the derivative of P with respect to p_k, which we label by the function $b(p_k)$. Notice that $b(p_k) \equiv \partial P/\partial p_k = \left(\frac{P}{p_k}\right)^\theta$, so that $b(p_k) > 0$, with $b'(p_k) = \theta \frac{b(p_k)}{P}\left(b(p_k) - b(p_k)^{\frac{1}{\theta}}\right)$. Moreover, in a symmetric equilibrium, where $p_k = p$ for all k, $b(p) = K^{\frac{\theta}{1-\theta}}$ and $b'(p) = \frac{\theta}{Kp}\left(K^{\frac{\theta}{1-\theta}} - K^{\frac{1}{1-\theta}}\right)$.

Demand, $c_k(p_k)$. This is analogous, through the lens of the model, to evaluate $d(p)$. It is immediate to see, from the expression for $c_k(p_k)$, that the demand for variety k converges to zero as its price diverges to infinity. That is, $\lim_{p_k\to\infty} c_k(p_k) = 0$. We now show that, in a symmetric equilibrium, when the number of varieties is large, the demand for variety k is decreasing and convex in its price. It is straightforward to compute the following derivatives,

$$\frac{\partial c_k(p_k)}{\partial p_k} = \frac{c_k(p_k)}{P}\left(b(p_k)(\theta - 1) - \theta b(p_k)^{\frac{1}{\theta}}\right) ,$$

$$\frac{\partial^2 c_k(p_k)}{\partial p_k^2} = \frac{1}{c_k(p_k)}\left(\frac{\partial c_k(p_k)}{\partial p_k}\right)^2 - \frac{\partial c_k(p_k)}{\partial p_k}\frac{b(p_k)}{P} + \frac{c_k(p_k)}{P}b'(p_k)\left(\theta - 1 - b(p_k)^{\frac{1-\theta}{\theta}}\right) ,$$

which, in a symmetric equilibrium, reduce to

$$\frac{\partial c_k(p)}{\partial p_k} = \frac{c_k(p)}{p}\left(\frac{\theta - 1}{K} - \theta\right) ,$$

$$\frac{\partial^2 c_k(p)}{\partial p_k^2} = \frac{1}{c_k(p)}\left(\frac{\partial c_k(p)}{\partial p_k}\right)^2 - \frac{\partial c_k(p)}{\partial p_k}\frac{1}{Kp} + \theta\frac{c_k(p)}{p^2}\left(\frac{1}{K} - 1\right)\left(\frac{\theta - 1}{K} - 1\right) ,$$

where we also used that, in a symmetric equilibrium, $P = K^{\frac{1}{1-\theta}}p$. Notice that, in the symmetric equilibrium, if K is large, we have that $\frac{\partial c_k(p)}{\partial p_k} < 0$ and $\frac{\partial^2 c_k(p)}{\partial p_k^2} > 0$, consistent with the demand function $d(p)$ being decreasing and convex in p. Moreover, because $c_k(p_k)$ and the price index P are twice continuously differentiable in prices and number of varieties K, the result also applies more generally away from the symmetric equilibrium.

43

Indirect utility function, $u(C(p_k))$. This is analogous, through the lens of the model, to evaluate $v(p)$. Using the definition of $c_k(p_k)$ together with $PC = I$, we can obtain the following expressions,

$$\frac{\partial C(p_k)}{\partial p_k} = -\frac{c_k(p_k)}{P} \quad , \quad \frac{\partial^2 C(p_k)}{\partial p_k^2} = -\left[\frac{\partial c_k(p_k)}{\partial p_k}\frac{1}{P} - \frac{c_k(p_k)}{P^2}b(p_k)\right] \quad .$$

Notice that, because $\frac{\partial C(p_k)}{\partial p_k} < 0$, we also have that $\frac{\partial u(C(p_k))}{\partial p_k} < 0$. The second derivative of the indirect utility function with respect to p_k can be written as

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

so that $\frac{\partial^2 u(C)}{\partial p_k^2} \leq 0$ if $\gamma - \theta\left(1 - b(p_k)^{\frac{1-\theta}{\theta}}\right) + 2 \geq 0$. For example, in the symmetric equilibrium, the required condition can be rewritten as $\gamma - \theta(1 - K) \geq -2$, which is satisfied, for example, for any $K \geq 2$.

Intensive margin demand elasticity, $\varepsilon_d(p_k)$. Using the definition of $c_k(p_k)$ and price index P provide that the intensive margin demand elasticity of variety k is given by

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

where $c_k(p_k)p_k = \left(\sum_{i=1}^K \left(\frac{p_i}{p_k}\right)^{1-\theta}\right)^{-1}$. Notice that, because $\theta > 1$ and $0 < c_k(p_k)p_k < I$, we have that $\varepsilon_d(p_k) > 1$. Also notice that, in a symmetric equilibrium, as K diverges to infinity we get that $\varepsilon_d(p_k) = \theta$, so that when there are infinite many varieties the demand elasticity is constant. Moreover, notice that

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

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

Profits, $\pi(p_k, z)$. 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 $\partial \pi(p_k, z)/\partial p_k = 0$. Second, that this solution is a maximum (i.e. that the profit function is strictly concave).

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

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

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

We now show that this unique price maximizes the firm's profits. To this end, we show that in a symmetric equilibrium, for large K, the profit function evaluated at this price is concave. Then, because all objects are well behaved with respect to K and prices (i.e. they are twice continuous differentiable), concavity also applies more generally away of the symmetric equilibrium.

The second derivative of the profit function with respect to p_k reads,

$$\frac{\partial^2 \pi(p_k, z)}{\partial p_k^2} = -\frac{c_k(p_k)}{p_k} \left[\varepsilon_d(p_k) \left(1 - \varepsilon_d(p_k) \left(1 - \frac{w/z}{p_k} \right) \right) + p_k \frac{\partial \varepsilon_d(p_k)}{\partial p_k} (1 - \frac{w/z}{p_k}) + \varepsilon_d(p_k) \frac{w/z}{p_k} \right] \ .$$

Notice that, in a symmetric equilibrium, c_k, p, $\varepsilon_d(p)$, and $\frac{\partial \varepsilon_d(p)}{\partial p_k}$ are continuous in K. We will use this fact to prove that for large K the profit function is concave at the price maximizing static profits. Notice that, in a symmetric equilibrium, when K diverges to infinity the second derivative reduces to

$$\lim_{K \to \infty} \frac{\partial^2 \pi(p, z)}{\partial p_k^2} = -\frac{c_k(p)}{p} \left[\theta \left(1 - \theta \left(1 - \frac{w/z}{p} \right) \right) + \theta \frac{w/z}{p} \right] \ ,$$

because $\lim_{K \to \infty} \varepsilon_d(p_k) = \theta$ and $\lim_{K \to \infty} \frac{\partial \varepsilon_d(p)}{\partial p_k} = 0$. Moreover, the markup $p/(w/z)$ 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,

$$\lim_{K \to \infty} \frac{\partial^2 \pi(p, z)}{\partial p_k^2} = -\frac{c_k(p)}{p} (\theta - 1) < 0 \ ,$$

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 $c_k(p)$, p, $\varepsilon_d(p_k)$, and $\frac{\partial \varepsilon_d(p)}{\partial p_k}$ are continuous in K, it is also the case that, in a symmetric equilibrium, $\frac{\partial^2 \pi(p,z)}{\partial p_k^2} < 0$ for large K. In the end, we concluded that if there is a large number of varieties, the profit function is concave, and $\partial \pi(p, z)/\partial p = 0$

45

characterizes its maximizer.

D   Numerical solution of the model

First, we set parameters. A first group of parameters is constant throughout the numerical exercises. These include β, w, q and I. We consider a grid of values for each of the other parameters, i.e. $\lambda, \zeta, \theta, \rho$ and σ.

We next describe how we solve for the equilibrium of the model for a given combination of parameters. We start by discretizing the AR(1) process for productivity to a Markov chain featuring $K = 25$ different productivity values. We then conjecture an equilibrium function $\mathcal{P}(z)$. Given our definition of equilibrium and the results of Proposition 2, we look for equilibria where $\mathcal{P}(z) \in [p^*(\bar{z}), p^*(\underline{z})]$ for each z, and $\mathcal{P}(z)$ is strictly decreasing in z. Our initial guess for $\mathcal{P}(z)$ is given by $p^*(z)$ for all z. We tried other guesses and we found that the algorithm converges to the same equilibrium. Given the guess for $\mathcal{P}(z)$, we can compute the continuation value of each customer as a function of the current price and productivity, i.e. $\bar{V}(p, z)$, and solve for the optimal search and exit thresholds as described in Lemma 1. Given $\mathcal{P}(z)$ and the customers' search and exit thresholds we can solve for the distributions of customers $Q(\cdot)$ and $H(\cdot)$ as defined in Definition 1. Notice that the latter also amounts to solve for a fixed point in the space of functions. Here, standard arguments for the existence of a solution to invariant distribution for Markov chains apply. Therefore, the assumption that $F(z'|z) > 0$ and $\Delta(\hat{p}(z), z) > 0$ ensure the existence of a unique $K(z)$ that solves equation (9). Finally, given $Q(\cdot)$, $H(\cdot)$, $\mathcal{P}(z)$ and $\bar{V}(p, z)$, we solve the firm problem and the obtain optimal firm prices given by the function $\hat{p}(z)$. We use $\hat{p}(z)$ to update our conjecture about equilibrium prices $\mathcal{P}(z)$, and iterate this procedure until convergence to a fixed point where $\mathcal{P}(z) = \hat{p}(z)$ for all $z \in [\underline{z}, \bar{z}]$.

E   The model with aggregate shocks

In this appendix we provide some details on the model we use to evaluate aggregate shocks. We start by describing the household preferences. We assume that the household is divided into a mass Γ of shoppers/customers, and a representative worker. The preferences of the household are given by

$$E_t \left[\int_0^\Gamma V_t(p_t^i, z_t^i, \psi_t^i) \, di - J_t \right] , \tag{12}$$

where $V_t(p_t^i, z_t^i, \psi_t^i)$ is defined as in equation (1) and is the value function that solves the customer problem in Section 2.1, while $J_t \equiv \phi \sum_{T=t}^\infty \beta^{T-t} \ell_T$ with $\phi > 0$ denotes the disutility

from the sequence of labor ℓ_T. The aggregate state for the household includes the distribution of prices, the distribution of customers over the different firms and the level of income, the wage, and their law of motion. Given that here we allow for aggregate shocks, we have to allow for the possibility that the aggregate state varies over time. Therefore, we index such dynamics in the aggregate state through the time subscript t for the value function.

The worker chooses the path of ℓ_t that maximize household preferences in equation (12). The search problem of each customer is as described in Section 2.1. As for the consumption decision, each customer allocates her income across consumption of the good sold in the local market, the demand of which we denote by d, and another supplied in a centralized market by a perfectly competitive firms, the demand of which we denote by k, to solve the following problem

$$v_t(p_t) = \max_{d,\, n} \; \frac{\left(d^{\frac{\theta-1}{\theta}} + n^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}(1-\gamma)}}{1-\gamma} \tag{13}$$

$$\text{s.t.} \quad p_t\, d + q_t\, n \leq I_t, \tag{14}$$

where $\theta > 1$ and $I_t \equiv (w_t \ell_t + D_t)/\Gamma$ is nominal income, which the customer takes as given. Nominal income depends on the household labor income $(w_t \ell_t)$ and dividends from firms ownership (D_t). The first order condition to the problem in equations (13)-(14) gives a standard downward sloping demand function for variety d, which we denote by the function

$$d_t(p_t) = I_t \frac{p_t^{-\theta}}{p_t^{1-\theta} + q_t^{1-\theta}} \; . \tag{15}$$

Without loss of generality we use the price q_t as the numeraire of the economy. From the first order conditions for the household problem, we obtain that the stochastic discount factor is given by $\beta\, \Lambda_{t+1}/\Lambda_t$, where $\Lambda_{t+s} = \int_0^\Gamma \left(c_{t+s}^i \right)^{-\gamma} /P_{t+s}^i \, di$ is the household marginal increase in utility with respect to nominal income, with c_{t+s}^i denoting customer i's consumption basket in period $t + s$, and $P_{t+s}^i = ((p_{t+s}^i)^{1-\theta} + (q_{t+s})^{1-\theta})^{\frac{1}{1-\theta}}$ the associated price.

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

There are two exogenous driving processes in our economy, namely aggregate productivity Z and the numeraire q. We consider an economy in steady state at period t_0 where expectations are such that $Z_t = 1$ and $q_t = 1$ for all $t \geq t_0$. Notice that in this case the economy coincides with the economy described in Section 2.

E.1 The response to aggregate productivity shock

We consider the economy in its steady state at $t = t_0$, i.e. $Z_{t_0} = q_{t_0} = 1$, in absence of any foreseen aggregate shock; we then hit the economy with a one time unforeseen shock to firms' productivity and study the convergence of the economy back to its initial steady state. The aggregate productivity shock takes the form of a one time unforeseen parallel shift to firms productivities so that aggregate productivity jumps from $Z_{t_0} = 1$ to $Z_{t_0} = 1.05$. After the aggregate shock hits at $t = t_0$, Z_t decays exponentially for all $t > t_0$, i.e. $\log(Z_t) = \rho \log(Z_{t-1})$. Notice that, as long as the shock lasts, the household experiences an increase in current income.

The left panel of Figure 7 evaluates the impulse response of average markups. Start by noting that the aggregate shock does not affect markups in the inelastic customer base economy, while it has an effect, albeit small, in the model with customer markets: Average markups are procyclical, with an initial response of almost 0.3%.

Figure 7: Aggregate productivity shock and markups
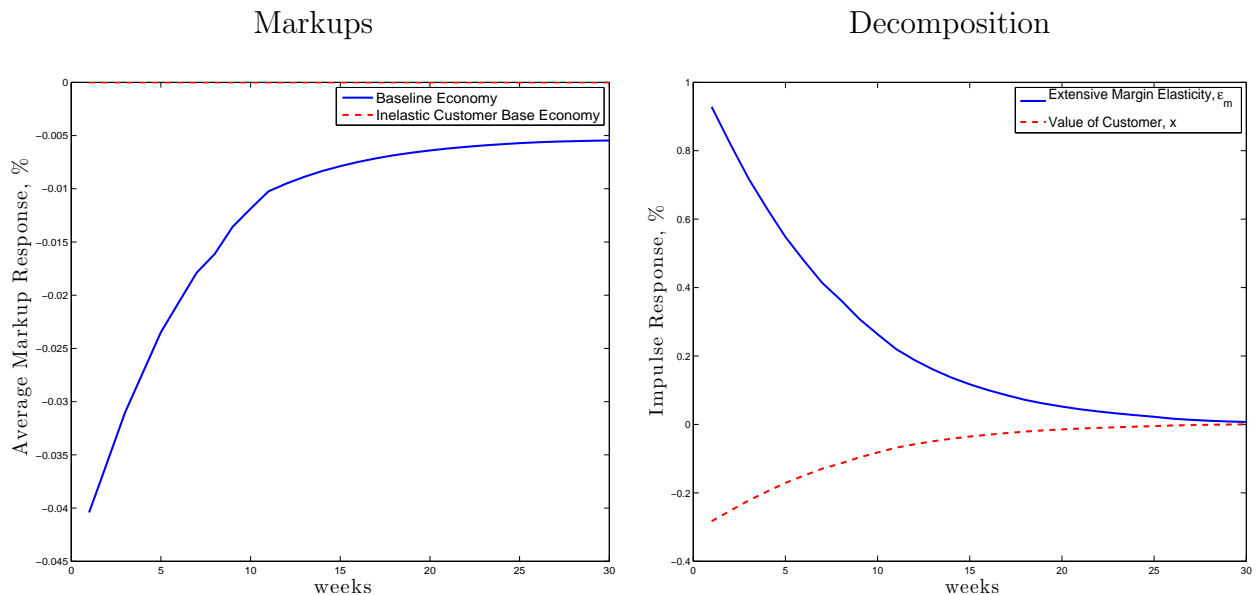
Values are presented as cross-firm averages in log-deviations from their steady state values.

48

Understanding the response of markups to aggregate shocks with customer markets, requires understanding how the extensive margin elasticity, $\varepsilon_m(\cdot)$, and the relative value of a customer, $x(\cdot)$, respond to the shock. This is important because, as shown in equation (8), equilibrium markups depend on the product of these two objects. With this in mind, the right panel of Figure 7 plots the impulse response of the average extensive margin elasticity and relative value of a customer. As the plot shows, both forces play in opposite directions. A fall in the relative value of a customer motivates firms to increase markups. However, an increase in the extensive margin elasticity motivates firms to decrease markups. We now turn into understanding how the aggregate productivity shock creates these two opposing effects. The shock implies that, on impact, household wealth increases, and then slowly returns to its steady state value. On the one hand, because consumption is a normal good, this translates into an initial increase in consumption, which also slowly fades aways with the shock. This affects the firms discount factor, $R_{t+1|t}$, as this one depends on the ratio of the customers marginal utilities derived from consumption. Because customers are risk averse, the interest rate falls if the consumption profile is decreasing. On impact, consumption increases, so that the interest rate increases, so that firms value less future income streams. As a result, the relative value of a customer, $x(\cdot)$, decreases; this motivates firms to charge higher markups. Later, as consumption falls, the interest rate decreases, and the value of a customer returns to its steady state value. On the other hand, the fact that the increase in wealth is transitory motivates customers to increase their interest on searching for cheaper stores. As a result, the extensive margin elasticity increases, on impact, slowly returning to its steady state value as the shock fades away. This effect motivates firms to decrease their markups. Overall, what we have is that there are two forces that go in opposite directions. The first force motivates firms to increase current markups: Because the relative value of a customer falls, firms want to disinvest on its customer base. The second force motivates firms to decrease current markups: Because of the transitory nature of the increase in household wealth, customers are more active on searching for cheaper stores, which makes the extensive margin of demand more elastic.