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

July 4, 2014

Abstract
We study optimal price setting in a model with customer markets. Customers face search frictions preventing them from freely 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. We find that customer markets substantially reduce the pass-through of cost shocks to prices and introduce. The inertia in the response of the customer base to price variation causes the response of demand to be higher in the long-run than in the short-run.

JEL classification: E30, E12, L16
Keywords: customer markets, price setting, real rigidities, product market frictions

*Corresponding Author: luigi.paciello@eief.it. Previous drafts of this paper circulated under the title “Price Setting with Customer Retention”. We benefited from comments on earlier drafts at the Minnesota Workshop in Macroeconomic Theory, 2nd Rome Junior Macroeconomics conference, 2nd Annual UTDT conference in Advances in Economics, 10th Philadelphia Search and Matching conference, ESSET 2013, MBF-Bicocca conference, MacCI Mannheim, the Collegio Carlo Alberto Pricing Workshop, Goods Markets Paris Conference 2014, the Macroeconomy and Policy, and seminars at Bank of France, Bank of Spain, Columbia University, Federal Reserve Bank of Richmond, the Ohio State University, University of Pennsylvania, Macro Faculty Lunch at Stanford, and University of Tor Vergata. We thank Fernando Alvarez, Lukasz Drozd, Huberto Ennis, Mike Golosov, Hugo Hopenhayn, Eric Hurst, Pat Kehoe, Pete Klenow, Francesco Lippi, Kiminori Matsuyama, Guido Menzio, Dale Mortensen, Ezra Oberfield, Facundo Piguillem, Valerie Ramey, Leena Rudanko, Stephanie Schmitt-Grohe, and Martin Uribe. Luigi Paciello thanks Stanford University for hospitality. The views expressed in this article are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.
1 Introduction

A primary goal of business strategy is to maintain and grow the customer base of a firm, by which we mean the set of customers currently buying from it. Pricing is an important instrument to achieve this, since customers pay close attention to economic convenience when choosing where to shop. In this paper we study the price setting problem of firms competing for customers and its implications for the dynamics of demand. The presence of customer markets creates a link among firms’ pricing decisions, providing a natural foundation for strategic complementarities in price setting (Phelps and Winter (1970)). The stickiness in the customer base creates a link among current and future firms’ demand (Foster et al. (2012), Gourio and Rudanko (2014b)). As firms’ pricing and demand formation are prominent determinants of dynamics in modern macroeconomic models (Bai et al. (2012) and Kaplan and Menzio (2013)), it is important to shed light on their underpinnings.

We develop a tractable model where the customer base of a firm responds sluggishly to prices because of search frictions that prevent customers from freely moving across different suppliers of the same good. In this framework the customer base of a firm is a persistent determinant of the firm’s future demand. As a result, the firm has an incentive to invest in its customer base. We focus on the implications of the customer base dynamics for firms pricing decisions. Our model features two separate margins through which the price affects the level of demand. First, there is an extensive margin associated to the effect of a change in price on the number of customers buying from a firm we just described. In addition, there is an intensive margin which relates to the effect of a change in price on the quantity demanded by each customer from a given supplier. Only the latter is present in workhorse macroeconomic models arising from the possibility of substituting expenditure across suppliers of different varieties of goods (Dixit and Stiglitz (1977)). This class of demand models is nested in our framework when search costs go to infinity shutting down the extensive margin. We complement our modeling effort with an empirical analysis that relies on novel micro data documenting pricing and customer base evolution for a large retail firm. We use the data to both provide descriptive evidence of the relevance of pricing to shape customer base dynamics and to estimate the key parameters of our model. Finally, we use the estimated model to study the effect of an exchange rate shock and show that customer markets have important implications for the propagation of such shocks.

There are two types of agents in our model, firms and customers. Firms produce an homogeneous good with a linear production function in its single variable input. Firms are heterogenous in two dimensions: they have different idiosyncratic productivity and different customer base. Productivity follows an exogenous Markov process, whereas changes in the
customer base arise endogenously and depend on the search decisions of customers. There is no commitment technology and firms post their price each period after observing the realization of their productivity draw. Customers derive utility from the consumption of the homogenous good and start each period matched to the same firm from which they bought in the previous period. Customers have perfect information on the state of the economy as well as on the characteristics of their current match and must pay a cost to draw a price offer from another randomly assigned firm. Upon searching, the customer observes the characteristics of the new firm, compare them to those of her old supplier, and decides where to buy. The combination of the search decision with the (subsequent) decision of where to buy determines the extensive margin of demand. After customers have made their search and matching decisions, each customer decides her purchased quantity of the homogenous good according to a demand function decreasing in the price. This downward sloping function drives the intensive margin of demand. Firms cannot price discriminate and therefore face a trade-off between charging a higher price and extract more surplus from high search costs customers, versus posting a lower price to extract a lower surplus but from a larger mass of customers. Since the customer base is sticky, changes in customer base have persistent effects on demand making the price setting problem of the firm a dynamic one. We characterize an equilibrium with i.i.d. productivity and show that it features both price dispersion and customer dynamics. The equilibrium prices are decreasing in productivity and more productive firms enjoy higher growth in their customer base.

The quantitative predictions of our model will depend on its parameters. It is crucial that we are able to measure them in order to assess the relevance of customer markets for pricing. Therefore, we estimate our model taking advantage of scanner data from a major U.S. retailer 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 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 indeed affected by variation in the price: a 1% change in the price of the customer’s typical basket of goods would raise the firm yearly customer turnover from 14% to 21%. This suggests that firms can indeed influence customer dynamics through variation in prices. We use this result together with our model to explore the implications of this margin for optimal price setting.
We use the price elasticity of the customer base, jointly with moments from the distribution of prices posted by the chain, to identify the key objects of the model: the distribution of search costs and the properties of the productivity process. We assess the relevance of customer markets for price dynamics by comparing it to a counterfactual economy where only the intensive margin is present. This is an interesting benchmark, as the pricing problem of the firm in such economy is similar to the one of standard macro models where competition comes only from the downward sloping demand of each customer, and the customer base is constant. We design the experiment so that the counterfactual economy is observationally equivalent with respect to average demand elasticity, price persistence and dispersion. We find that customer markets have several implications for firms pricing and document that they are non-trivial in magnitude. First, our model features equilibrium customer dynamics: the yearly average turnover is 9%. Yearly turnover in the data is 14%, implying that customer dynamics triggered by prices can explain a large fraction of overall turnover. More productive firms gain customers because they charge lower prices which both reduces customer attrition and increases the inflow of new customers. The reverse hold for firms with low productivity.

A second set of results relates to the pricing decision. We find that competition for customers reduces the average pass-through of idiosyncratic cost shocks to prices to about a quarter of the corresponding figure in counterfactual economy with no extensive margin. The pass-through of cost shocks is incomplete in our model because both the extensive and the intensive margin elasticities of demand increase with the price, so that optimal markups decrease with an increase in production costs. The pass-through of cost shocks is substantially lower in our model than in the counterfactual economy because a change in the price has persistent effects on demand, amplifying the effect on optimal prices of the extensive margin elasticity. The low pass-through predicted by the model is consistent with what we measure from our data, and similar evidence is found for firm level pass-through of exchange rate shocks in Burstein and Jaimovich (2012). In addition, we show that the scale of the search friction affects the shape of the distribution of prices. As search costs decrease, the price distribution features higher kurtosis. The kurtosis of the distribution of prices is the result of the cross-sectional variation in pass-through of cost shocks, which display an inverse-U shape with respect to firm productivity. Highly productive firms are attractive to customers and can therefore pass-through a higher fraction of cost changes than firms with average productivity. Firms with very low productivity also pass-through more than average productivity ones as their profit margins are already low and they have low chance to retain customers anyway, mitigating their incentive to contain prices. We will use the mapping from the distribution of search costs to the shape of the distribution of prices in our data to identify the scale of
Finally, we use our model to explore the effects of a productivity shock that affects a subset of the firms. This is far from an abstract setup: it captures the main features of an exchange rate shock (which affects firms’ costs differently according to whether or not they buy inputs on the international market) or a change in the state sale tax (which penalizes and benefits competing firms on the basis or where their headquarters are established). Our model represents an interesting setting to study this type of shocks. In our economy, shocks that asymmetrically hit firms affect price dispersion. This in turn influences the incentives of customers to search generating persistent dynamics in demand. This margin is absent in models who do not include customer markets. We simulate a persistent real exchange rate appreciation which makes foreign firms more productive relatively to domestic ones. We find a low pass-through of the shock to the price of both imported and domestically produced goods, with the price of imported goods declining more than the price of domestic goods. Aggregate demand increases by little but the composition shifts significantly in favor of imports. The customer relocation has persistent effects, implying that the long-run effect of the exchange rate shock on the ratio of demand for imported and domestic goods (Armington elasticity) is significantly larger than the short-run one. This is consistent with empirical evidence (Ruhl (2008)) and it is in stark contrast with the predictions of the counterfactual model without customer markets.

Related Literature. Our paper relates to the seminal work by Phelps and Winter (1970) who study the pricing problem of the firm facing customer retention concerns. In their paper, the response of the firm’s customer base to a change in the firm’s price is modeled with an ad hoc function. We instead endogenize customer dynamics in response to firms’ pricing as the outcome of customers’ optimal search decisions. Alessandria (2004) and Menzio (2007) also study the firm price setting problem in models where search frictions prevent customers from freely moving to the lowest price supplier. In Alessandria (2004) the firm’s pricing affects the exit rate, but not the arrival rate, of customers. The author shows that his model can generate large and persistent deviations from the low of one price, consistent with the empirical evidence on international prices. In Menzio (2007) the firm’s price path affects the arrival rate, but not the exit rate, of customers. The author studies how firms optimal pricing depends on the ability of firms to commit to a price path and on the information available to customers. In these papers customers are homogeneous with respect to the search friction and, as a result, optimal pricing is such that no endogenous customer dynamics occur in equilibrium. Kleshchelski and Vincent (2009) use idiosyncratic switching costs,  

1See also Kaplan and Menzio (2014) for detailed evidence on the distribution of retail prices.
modeled as shocks to preferences of different brands, to obtain a model where customer
dynamics occur in equilibrium and firms pricing affect the exit rate of customers. Differently
from us, they focus on a symmetric equilibrium that features no price dispersion due to the
absence of heterogeneity in firms productivity, and use perturbation methods around this
equilibrium to study the pass-through of cost shocks to prices. Apart from the specifics of
the model, we differ from these studies because we characterize an economy that features both
endogenous customer dynamics and a non-degenerate distribution of prices in equilibrium.
This is interesting not only from a theoretical perspective but also because, combined with
observable statistics from the cross-sectional variation in prices and customer dynamics, it
allows us to quantify the key structural parameters of the model.

Another approach within the same literature allows for a dedicated technology firms can
use to attract new customers, such as advertising, in combination with the pricing instru-
ment, and explores the connection between product market frictions and firms behavior.\(^2\) In
particular, Gourio and Rudanko (2014b) explore the relationship between the firm’s effort to
capture customers and its performance, focusing on firms investment and Tobin’s q. Using
cross-industry variation in selling expenses, they find that customer markets have non-trivial
implications for the relationship between investment and Tobin’s q. Dinlersoz and Yorukoglu
(2012) study a setting where firms exert advertising effort to disseminate information to
uninformed customers who instead sustain no search cost. They study how the presence
of customer markets impact on industry dynamics, focusing on firms entry and exit, and
provide descriptive evidence in support of the implications of their model.\(^3\) Drozd and Nosal
(2012) show that a standard international real business cycle model where however producers
have to exert effort to find new customers and increase sales can explain several patterns in
the dynamics of international prices and trade, and in particular a large long-run Armington
elasticity. Shi (2011) studies a model where firms cannot price discriminate across customers
and use sales to attract new customers. We differ from these papers because we focus on the
implications of customer retention, rather than acquisition, concerns for the pricing decision
of the firm.

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)). Real rigidities are mechanisms that dampen price
responses of firms because of factors such as strategic complementarities in price setting.

\(^2\)See Hall (2012) and Gourio and Rudanko (2014a) for studies of the business cycle properties of advertising
expenditure.

\(^3\)Burdett and Coles (1997) use search frictions to study a model where young firms charge lower prices to
increase the arrival probability of new customers. This approach is also used in the context of labor markets. See
With respect to this literature, a distinctive characteristic of our model is that the pricing problem is dynamic, as the current price affects both current and future demand. This has important consequences both for price and demand dynamics. First, everything else being equal, the persistent loss in customers that follows an increase in the price induces firms to charge relatively lower markups and pass-through shocks to costs less. Second, after a change in the firm price, the demand responds sluggishly so that the long-run response is larger than the short-run one.

On the empirical side, a large literature has used 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 (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 variation in prices.

Finally, a stream of studies analyzes the implications of product market frictions for business cycle fluctuations (Petrosky-Nadeau and Wasmer (2011), Bai et al. (2012) and Kaplan and Menzio (2013)). In these papers, aggregate shocks cause the cost-opportunity of searching to vary, having consequences for markup and demand dynamics over the business cycle. While we abstain from this type of analysis, our quantified model could be extended to allow for the search opportunity-cost to vary with the aggregate state and be used as a laboratory for these studies.

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 exchange rate type of shocks. Section 7 concludes.

2 The model

The economy is populated by a measure one of firms producing an homogeneous good, and a measure $\Gamma$ 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
\[ \varepsilon_d(p) \equiv -\partial \log d(p)/\partial \log p \geq 1; \text{ and (ii) } v(p) \text{ is continuously differentiable with } v'(p) < 0, \]
\[ v''(p) \leq 0 \text{ and } \lim_{p \to \infty} v'(p) = -\infty, \lim_{p \to 0^+} v'(p) = 0. \]
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 \( \psi \geq 0 \) measured in units of customer surplus. The search cost is drawn each period from the same distribution with density \( g(\psi) \) continuous on the support, and associated cumulative distribution function denoted by \( G(\psi) \) which we assume differentiable. 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. 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, \( \ell \), 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 \( \ell \), \( p \) 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 \([z, \bar{z}]\). 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 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 \( \psi_t^i \) 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 customer perfectly observes not only the price but also productivity and customer base of the

---

4This captures the idea that larger firms attract more customers (Rob and Fishman (2005)).
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 \( P(z) \) 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 \( \psi \). We have that this value function solves the
following problem,

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

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 \( P(\cdot) \) mapping future productivity into prices in
the Markov equilibrium, the value in the case of not searching is given by

\[ \bar{V}(p^t_j, z^t_j) = v(p^t_j) + \beta \int_0^\infty \int_z^{\bar{z}} \bar{V}(P(z'), z', \psi') \, dF(z' | z^t_j) \, dG(\psi') . \]  

(2)

The value when searching is given by

\[ \tilde{V}(p^t_j, z^t_j) = \int \max \{ \bar{V}(p^t_j, z^t_j), x \} \, dH(x) , \]

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 \( \psi_t \) smaller than a threshold, i.e. \( \psi_t \leq \hat{\psi}(p^t_j, z^t_j) \), where \( \hat{\psi}(p, z) = \int_{\bar{V}(p, z)}^{\infty} (x - \bar{V}(p, z)) \, dH(x) \geq 0 \); 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 \( \psi \) 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.

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 \( \bar{V}(z) \equiv \bar{V}(P(z), z) \) 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 $\tilde{V}(z)$ 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 $t$ ($m_{jt}$) is the mass of customers buying from firm $j$ in period $t$. It evolves as follows

$$m_{jt} = m_{j,t-1} - m_{j,t-1} G(\hat{\psi}(p_{jt}, z_{jt})) \left(1 - H(\tilde{V}(p_{jt}, z_{jt}))\right) + \frac{m_{j,t-1}}{\Gamma} Q(\tilde{V}(p_{jt}, z_{jt})),$$

where $m_{j,t-1}$ is the mass of old customers, $G(\hat{\psi}(p_{jt}, z_{jt}))$ is the fraction of old customers searching, a fraction $1 - H(\tilde{V}(p_{jt}, z_{jt}))$ of which actually finds a better match and exits the customer base of firm $j$. The ratio $m_{j,t-1}/\Gamma$ is the probability that searching customers in the whole economy draw firm $j$ as a potential match. The function $Q(\tilde{V}(p_{jt}, z_{jt}))$ denotes the equilibrium mass of searching customers currently matched to a firm with continuation value smaller than $\tilde{V}(p_{jt}, z_{jt})$. 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 $m_{jt} = m_{j,t-1} \Delta(p_{jt}, z_{jt})$, where the function $\Delta(\cdot)$ denotes the growth of the customer base and is given by

$$\Delta(p, z) = 1 - G(\hat{\psi}(p, z)) \left(1 - H(\tilde{V}(p, z))\right) + \frac{1}{\Gamma} Q(\tilde{V}(p, z)).$$

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 and 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

$$\tilde{W}(z^t_j, m^t_j - 1) = \max_p m^t_j \pi(p, z^t_j) + \beta \int_{z}^{\bar{z}} W(z', m^t_j) \, dF(z' | z_t),$$

subject to equation (3), where $\tilde{W}(z^t_j, m^t_j - 1)$ 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., $\bar{W}(z, m) = m \bar{W}(z, 1) \equiv m W(z)$, where $W(z)$ solves

$$W(z) = \max_p \Delta(p, z) \left( \pi(p, z) + \beta \int_{z}^{\bar{z}} W(z') dF(z' | z_t) \right) ,$$

where we used equation (3) and we dropped time and firm indexes to ease the notation. We assume that the discount rate $\beta$ 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 $\bar{W}(z, m)$ 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,

$$p^*(z) = \frac{\varepsilon_d(p)}{\varepsilon_d(p) - 1} \frac{w}{z} .$$
It follows that setting a price above the static profit maximizing price is never optimal. Moreover, if \( \hat{p}(z) \leq p^*(z) \), the optimal price will not be below \( \hat{p}(z) \), because in that region profit per customer decrease in the price but the customer base is unaffected, so that \( \hat{p}(z) \in [\hat{p}(z), p^*(z)] \). If instead \( \hat{p}(z) \geq p^*(z) \), 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.

**Proposition 1** Let \( \hat{p}(z) \) solve \( \hat{V}(\hat{p}(z), z) = \max_{x \in [\bar{z}, \bar{z}]} \{ \hat{V}(P(z), z) \} \), 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). We have \( \hat{p}(z) \in [\hat{p}(z), p^*(z)] \) if \( \hat{p}(z) < p^*(z) \), and \( \hat{p}(z) = p^*(z) \) otherwise.

A proof of the proposition can be found in Appendix A.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 \( \hat{V}(z) \equiv \hat{V}(P(z), z) \) 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 \( z \in [\bar{z}, \bar{z}] \) the firm pricing strategy \( \hat{p}(z) \in [p^*(\bar{z}), p^*(\bar{z})] \) solves the first order condition to the firm problem in equation (5) given by

\[
\frac{\partial \Pi(p, z)}{\partial p} \frac{p}{\Pi(p, z)} = -p \frac{\partial \Delta(p, z)}{\partial p} \geq 0 . \tag{7}
\]

A stationary equilibrium is then

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

(ii) a firm pricing strategy \( \hat{p}(z) \) that solves equation (7) for each \( z \), given customers’ strategies and equilibrium pricing policy \( P(z) \), and is such that \( \hat{p}(z) = P(z) \) 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 \( Q(x) = \Gamma \int_{\hat{z}}^x G(\hat{\psi}(\hat{p}(z), z)) dK(z) \) for each \( x \in [\hat{V}(\hat{z}), \hat{V}(\hat{z})] \), where \( \hat{z}(x) = \max\{z \in [\bar{z}, \bar{z}] : \hat{V}(z) \leq x\} \), and \( K(z) \) solves

\[
K(z) = \int_{\bar{z}}^z \int_{\bar{z}}^{\hat{z}} \Delta(\hat{p}(x), x) dF(s|x) dK(x) ds , \tag{8}
\]

for each \( z \in [\bar{z}, \bar{z}] \) with boundary condition \( \int_{\bar{z}}^{\bar{z}} dK(x) = 1 \).
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 \( \hat{p}(z) \in [p^*(z), p^*(\bar{z})] \) excludes those equilibria where firms cut prices below the static profit maximizing price of the most productive firm. Notice that Proposition 1 implies that \( \hat{p}(z) \leq p^*(z) \) for all \( z \in [\bar{z}, \bar{z}] \), so that we are effectively restricting only the lower bound. The first order condition in equation (7) 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).

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 \( (z_1, z_2) \in [\bar{z}, \bar{z}] \), 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, implying that the optimal markups are given by

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

where \( \varepsilon_d(p) \equiv \partial \log(d(p)) / \partial \log(p) \), \( \varepsilon_m(p, z) \equiv \partial \log(\Delta(p, z)) / \partial \log(p) \), and \( p = \hat{p}(z) \) for each \( z \).

(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 \( \Delta(\hat{p}(z), z) < 1 \).

We first comment on the assumptions behind the results of the Proposition. We 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. 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 verify 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. 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. More details on the proof of the proposition can be found in Appendix A.5.

We now comment on the properties of the equilibrium highlighted in the Proposition. The equilibrium is characterized by price dispersion: more productive firms charge lower prices and, therefore, offer higher continuation value to customers and grow faster. 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. As shown in equation (9), optimal markups depend on three distinct terms: $\varepsilon_d(p)$, $\varepsilon_m(p, z)$ and $x(p, z) \equiv \Pi(p, z)/(d(p) p)$. The terms $\varepsilon_d(p)$ and $\varepsilon_m(p, z)$ represent the price elasticities of quantity purchased (per-customer) and of customer growth, respectively. An increase in price reduces total current demand both because it reduces quantity per customer (intensive margin effect) and because it reduces the number of customers (extensive margin effect). Moreover, the optimal markup solves a dynamic problem as a loss in customers has persistent consequences for future demand due to the inertia in the customer base. This dynamic effect is captured by the term $x(p, z)$ which measures the firm present discounted value of a customer scaled by the current revenues. It follows that active customer markets are associated 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, and to highlight the different properties from a model of static demand, 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 $\varepsilon_q(\cdot)$, 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.\(^5\) Intuitively, a loss in demand associated to a loss in customers is a persistent loss and, therefore, has larger impact on firm value, inducing it to charge lower markups with respect to a firm that operates in a market where the customer base is less elastic.

\(^5\)More details are available in Appendix A.4.
The next remark explores two interesting limiting cases of our model.

**Remark 1** Let search costs be defined as \( n \psi \), where \( n > 0 \). That is, let the value function in equation (1) be
\[
V(p, z, \psi) = \max \left\{ \bar{V}(p, z), \tilde{V}(p, z) - n\psi \right\}.
\]

Two limiting cases of the equilibrium stated in Definition 1:

1. Let \( n \) diverge to infinity. Then, in equilibrium: (i) the optimal price maximizes static profits, \( \hat{p}(z) = p^*(z) \) for all \( z \in [\bar{z}, \bar{z}] \), (ii) equilibrium prices \( \hat{p}(z) \) are decreasing in productivity and 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. Let \( \pi(p^*(\bar{z}), \bar{z}) > 0 \) and let the assumptions of Proposition 2 be satisfied. Then, \( \max \{\hat{p}(z)\} = \hat{p}(\bar{z}) \) approaches \( \hat{p}(\bar{z}) = p^*(\bar{z}) \) as \( n \to 0 \). As a result, in the limit, there is no price dispersion in equilibrium and customers do not search.

A proof can be found in Appendix A.6. These limiting cases showcase the effect of search frictions on price dispersion. The first limiting case explores the resulting equilibrium when we let search costs diverge to infinity; we do this by introducing a scale parameter \( n \) and letting \( n \) to diverge.\(^6\) This case is interesting because the model reduces to one 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 widely explored in the macroeconomics literature: the firm sets the price \( p \) taking into account only its impact on static demand \( d(p) \). Not surprisingly, the equilibrium is unique, optimal prices maximize static profits, i.e. \( \hat{p}(z) = p^*(z) \) for all \( z \in [\bar{z}, \bar{z}] \), there is price dispersion, and there is no search in equilibrium. The second limiting case explores the resulting equilibrium when search costs become arbitrarily small, something that we do by letting \( n \) become arbitrarily small. We restrict attention to the model that satisfies the assumptions of Proposition 2, so that the first order condition presented in equation (7) is necessary for optimality.\(^7\) In this case, as the scale of search costs becomes arbitrarily small, equilibrium prices approach the lowest price in the economy, \( p^*(\bar{z}) \). As a result, there is no price dispersion and customers do not search; this result is reminiscent of Diamond (1971).

\(^6\)This is analogous to assume that \( G(\psi) = 0 \) for all \( \psi < \infty \).

\(^7\)The assumption \( \pi(p^*(\bar{z}), \bar{z}) > 0 \) is purely technical and it ensures that the first derivative of the profit function is bounded in the relevant range.
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 supermarket chain shared with us scanner data detailing purchases by a panel of households carrying a loyalty card of the chain. The chain operates over a thousand stores across ten states and the data reflect this geographical dispersion. For every trip made at the chain by those customers 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.

The population of shoppers visiting a retail store in a given time period can be divided into regular and occasional customers. Regular customers shop routinely, at least for some period of time, at the same store. Occasional customers are instead agents who typically shop elsewhere and visited the chain for convenience, for instance because they happened to be in the vicinities. For the sake of our analysis, it makes sense to restrict attention to regular customers: including occasional customers in the customer base of a firm may lead to misleading results. Since such agents make few purchases and far between, we would label their disappearances as exits from the customer base; whereas their presence at the chain depended on incidental factors. Luckily, our data are ideal to separate regular and occasional customers as they only include information on households carrying a loyalty card of the chain. The willingness to sign up for the loyalty card signals some form of commitment of these households to the chain which random shoppers are quite unlikely to bother with. The customers in our sample make an average of 150 shopping trips at the chain over the two years; if those trips were uniformly distributed that would imply visiting a store of the chain six times per month. The average expenditure per trip is 69 dollars for the average household. There is a great deal of variation (the 10th percentile is 29 dollars; the 90th is 118 dollars) explained, among other things, by income and family size of the different households.

\(^8\)The chain is able to associate the loyalty cards belonging to different members of a same family to a single household identifying number, which is the unit of observation in our data. Therefore, in the analysis we use the terms “customer” and “household” interchangeably.
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.\(^9\) 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.\(^{10}\) Although we acknowledge that the multiproduct nature of the problem may have implications for the pricing decision of the firm, we abstract from this issue in our analysis 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 given the shopping frequency of households in our sample.\(^{11}\) Regular customers are unlikely to experience an eight-weeks spell without shopping for reasons other than having switched to another chain (e.g. consuming their inventory). In fact, on average four days elapse between consecutive trips and the 99th percentile of this statistic is 28 days, half the length of the absence we require before inferring that a household is buying its 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. In a particular week \(t\), the price paid by customer \(i\), shopping at store \(j\) and for its basket, represented by the collection of UPC’s in \(K_i\), is

\[
p_{ijt} = \sum_{k \in K_i} \omega_{ik} p_{kjt}, \quad \omega_{ik} = \frac{\sum_t E_{ikt}}{\sum_{k \in K_i} \sum_t E_{ikt}},
\]

\(^9\)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 that they are not buying it at our chain, but we cannot exclude they are buying it elsewhere.

\(^{10}\)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.

\(^{11}\)We experimented with 4 weeks and 12 weeks as alternative lengths of the period of absence required to infer the exit from the customer base. In both cases the results are qualitatively similar. However, in the 12 weeks case the number of exit events becomes so small that it makes it hard to detect significant effects.
where $p_{kjt}$ is the price of UPC $k$ at store $j$ in week $t$ and $E_{ikt}$ is the expenditure (in dollars) by customer $i$ in UPC $k$ in week $t$. Note that the price of the basket is household specific because households differ in their choice of grocery products ($K_i$) and in the weight such goods have in their budget ($\omega_{ik}$). 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 them with store level data on weekly revenues and quantities sold.\footnote{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.} This data allows us to back out weekly prices of each UPC in the sample by dividing total revenues by total quantity sold as in Eichenbaum et al. (2011).

### 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 aim is to capture the effect of the price ($\log(p_{retailer})$) posted by the chain for the basket of goods purchased by the customer on her decision to exit. To isolate this magnitude in a way consistent with the mechanism described in the theoretical model, we need to include a series of controls.

We are interested in the effect of price variation induced by cost shifts idiosyncratic to a firm. 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. To identify idiosyncratic price variations, we control for the prices posted by the 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.\footnote{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.} Thanks to this data, we can observe the weekly price of a specific UPC at every chainstore sampled by IRI in the Metropolitan Statistical Area of residence of a customer. We can, therefore, construct the price of the basket bought by the customer at each store in the MSA in the same fashion described for the price of the basket at our chain. Finally, we take the average of such prices across all stores to compute the average market price of the basket ($p_{competitors}$). To further control for sources of aggregate variation, we include in the regression year-week fixed effects to account for time-varying drivers of the decision.
of exiting the customer base common across households (e.g. disappearances due to travel during holiday season).

The coefficient on the retailer price of the basket is therefore 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 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).

Another set of regressors is meant to acknowledge that, unlike posited in the model, customers are heterogenous in more dimensions than their cost to search. We start by controlling for customer heterogeneity including observable characteristics (age, income, and education) matched from Census 2000. We also consider location as a potential driver of the decision to exit. We control for the size of the set of potential store choices by including the number of supermarket stores in the zipcode of residence of the customer and factor her convenience in shopping by calculating the distance in miles between her residence and the closest store of the chain and the closest alternative supermarket. To pick up the heterogeneity in the types of goods different customers include in their basket, we control for the price volatility of the customer-specific basket and for its price in the first week in the sample, as a scaling factor. Finally, we calculate customer tenure, defined as the number of consecutive weeks the customer has spent in the customer base of the chain, and include it in the regression to account for the fact that long term customers of the chain may be less willing to leave it *ceteris paribus*.

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

\[
\text{Exit}_{it} = b_0 + b_1 \log(p_{retailer}^{i}) + b_2 \log(p_{competitors}^{i}) + b_3 \text{tenure}_{it} + X_i' \beta + \epsilon_{it}. \tag{10}
\]

In the regression we include

The retailer price in equations (10) would be endogenous if the chain conditioned its price to unobserved (to the econometrician) variables which also influence the customer’s decision to leave. We deal with the price of the individual specific basket, which implies that the firm would need to acquire relevant information on each customer (market level variables would not suffice) and be willing to tweak the prices of their baskets in response to that.\footnote{Note that if blockbuster goods, i.e. goods which account for a significant share of many consumers’ baskets, exist the chain may face a combinatorial problem and be unable to tune the prices of the individual baskets at will. For instance, the chain may want to have the price of the basket rise for some customers and fall for others. However, if it raises the price of a blockbuster good, it may prove hard to undo this with “reasonable” price decreases in other goods and obtain lower basket prices for some customers.}
Table 1: Effect of price on the probability of exiting the customer base

<table>
<thead>
<tr>
<th>Exiting: Missing at least 8 consecutive weeks</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \log(p_{retailer}) )</td>
<td>0.14**</td>
<td>-0.01</td>
<td>0.16*</td>
<td>0.15**</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.030)</td>
<td>(0.089)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Walmart entry</td>
<td>0.019*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \log(p_{competitors}) )</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.002***</td>
<td>-0.003***</td>
<td>-0.004***</td>
<td>-0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>52,670</td>
<td>52,670</td>
<td>66,182</td>
<td>52,101</td>
</tr>
</tbody>
</table>

Notes: An observation is a household-week pair. The results reported are calculated through two-stages least squares where we use the logarithm of the cost of the basket (constructed based on the replacement cost provided for each UPC by the retailer) as instrument for the logarithm of the retailer price. We trim from the sample households in the top and bottom 1% in the distribution of the number of trips over the two years. Demographic controls rely on a random 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%.

Nevertheless, the retailer provided along with the store price data a measure of replacement cost which represents a natural instrument so that we can estimate equations (10) through instrumental variables and be protected against potential endogeneity of price. We use replacement cost measure to construct an individual specific cost of the basket, following the same procedure used to obtain the price of the basket. The cost of the basket is then used as an instrument for the price of the basket in all the specifications.

The results are reported in Table 1. The main specification in column (1) shows that the basket price posted by the retailer significantly impacts the probability of leaving. The effect is also quantitatively important. The average probability of exiting the customer base (0.3% weekly) implies a yearly turnover of 14%; if the retailer’s prices were 1% higher, its yearly turnover would jump to 21%. The coefficient on the competitors’ price, which we

---

15 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.
would expect to enter with a negative sign, is not significant. This may be due to the fact that the IRI data only allow us to match a customer with the stores in her Metropolitan Statistical Area of residence. The set of stores the customer considers as alternative to the chain are probably located in a much smaller area and this introduces measurement error. The negative coefficient on tenure confirms the intuition that the longer the relationship between a firm and a customer the less likely they are to be interrupted. Among the several individual characteristics we control for (not reported for brevity) it is worth mentioning that distance from stores of the chain and distance from the closest competing store enter with the expected sign. Customers living in proximity of a store of the chain are less likely to leave it and those leaving closer to competitors’ stores are more inclined to do so.

In columns (2)-(4) we assess the robustness of these findings. We start by replacing the price of the individual basket with a price index for the store basket, defined as the average of the price of all the UPC’s sold by a store, weighted for their sales. This price is, by construction, equal for all the customer shopping in the same store. Column (2) shows that this results in price coefficient negative and not significant. We take this as evidence that the customer-specific basket price used in our main specification is a meaningful object, whose significance stems from being able to capture the set of prices each customer cares about and not just reflecting some aggregate trend in pricing. The fact that the competitor price comes up as not significant may raise suspicion that the variable is too noisy to control for the effect of competition. This is important as we are interested in price movements driven by idiosyncratic firm shocks. To assess whether this is the case, we experiment in column (3) with an alternative way to control for the effect of competition: we exploit episodes of entry by Walmart, a major retailer with which our chain is in direct competition. We use data from Holmes (2011) to identify the date of entry by a Walmart supercenter -i.e. a store selling grocery on top of general discount goods- in a zipcode where our retailer also operates a supermarket. The resulting event study allows us to measure the effect of the retailer price on the probability of exit controlling for the most relevant change in the competitive environment. The estimated coefficient falls in the same ballpark as that estimated in the main specification, which reassures on the effectiveness of the IRI price in measuring the competitors’ behavior. In column (4), we change the assumption on the imputation of the date of exit; rather than assuming that the customer left on the occasion of her last trip to the store we posit that the exit occurred in the first week of her absence.¹⁶ Even in this case, the main result stays unaffected. Finally, we performed a placebo test running the main specification 1,000 randomly assigning exits from the customer base but keeping their total

¹⁶This alternative assumption matches more closely our model where the customer leaves after having seen the prices of her current supplier and decided not to buy there.
number the same as in the actual data. This exercise is meant to assess the likelihood that the significance of our result is only due to a lucky occurrence. We find that only in 2.8% of the cases the simulation yields and price coefficient positive and significant at 5%.

5 Estimation and quantitative analysis

In this section we discuss the procedure we follow to estimate the model. We need to choose the discount factor $\beta$ 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 $F(z'|z)$. Below 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 $\beta$ ensures that the max-operator in equation (5) is a contraction.\footnote{This level of $\beta$ reflects that the effective discount rate faced by the firm is 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).}

Customer demand and surplus functions. We assume that customers have logarithmic utility in consumption, $c$. Consumption 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 $\theta > 1$.\footnote{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.} The first good (that we label $d$) is supplied by firms facing product market frictions as described in Section 2.2; the other good ($n$) acts as a numeraire and is sold in a frictionless centralized market. The sole purpose of good $n$ is to microfound a downward sloping demand $d(p)$ and, therefore, allowing for an intensive margin of demand. The parameter $\theta$ is chosen so that the implied average intensive margin elasticity of demand ($\varepsilon_d(p)$) is 7, a value in the range of those used in the macro literature. The customer budget constraint is given by $pd + n = I$, where $I$ is the agent’s nominal income which we normalize to one.\footnote{In Appendix D we show that $I$ can be derived based on a model of the labor markets.}

Firms productivity process. We assume that the productivity evolves according to a process of the following form:
\[
\log(z_j^t) = \begin{cases} 
\log(z_{j-1}^t) & \text{with probability } \rho, \\
z_{\text{new}} \sim N(0, \sigma) & \text{with probability } 1 - \rho
\end{cases}
\]

When solving the model numerically, we approximate the normal distribution on a finite grid, using the procedure described in Tauchen (1986). Finally, we normalize the nominal wage equal to the price of the numeraire good, so that \( w = 1^{20} \).

**Search cost distribution.** We assume that the search cost is drawn from a Gamma distribution with shape parameter \( \zeta \), and scale parameter \( \lambda \). The Gamma is a flexible distribution and fits the assumptions we made over the \( G \) function in the specification of the model. In particular, for \( \zeta > 1 \), we obtain that the distribution of search costs is differentiable at \( \psi = 0^{21} \).

### 5.1 Identification and estimation

We want to estimate the persistence and volatility of the productivity process (\( \rho \) and \( \sigma \)), and the scale and shape parameters of the search cost distribution (\( \lambda \) and \( \zeta \)). These are the key parameters of the model: the productivity process influences the variability of prices, which is necessary for customers to obtain any benefit from search. The parameters of the \( G \) distribution, on the other hand, directly determine how costly it is to search. These parameters are chosen by matching moments from our data with the analogue computed from the numerical solution of the model. Below we explain how the choice of the moments contributes to identifying particular parameters. This argument is provided only for the sake of intuition; given the nonlinearity of the model, all the moments contribute to the identification of all the parameters.

The model specifies how the parameters of the productivity process impact on the autocorrelation and volatility of firm prices. We therefore estimate persistence and volatility of productivity by matching the autocorrelation and the volatility of the logarithm of firm prices to those measured in the data for the logarithm of ratio between the customer’s basket price at the retailer and the average price of the competitors (\( \text{p}^{\text{ratio}} \equiv \text{p}_{\text{retailer}} / \text{p}_{\text{competitors}} \) in Section 4).\(^{22}\) Normalizing the firm’s price by that of the competitors is necessary to purge

---

\(^{20}\) This is equivalent to assume that the numeraire good \( n \) is produced by a competitive representative firm with linear production function and unitary labor productivity. See Appendix D for details.

\(^{21}\) In our estimation procedure we do not impose any constraints on the values the parameter \( \zeta > 1 \) can take. Our unconstrained point estimates lies in the desired region.

\(^{22}\) Autocorrelation and volatility of the innovation in the pricing process are obtained by estimating \( \log(p^\text{ratio})_t = k_0 + k_1 \log(p^\text{ratio})_{t-1} + \epsilon \) on the prices of the baskets of the customers included in the sample for the analysis in Section 4. The regression is estimated separately to remove the customer fixed effect and the median of the estimated \( k_1 \) (0.3) and the median of the standard deviation.
the data from aggregate effects and isolate the variation in price driven by the idiosyncratic component.

To identify the parameters of the search cost distribution we exploit two statistics. We first notice that in the Gamma distribution the shape parameter governs the hazard rate. This means that our parameter $\zeta$ influences the fraction of customers that draw a search cost low enough that they are allowed to search. The fraction of customers searching and that of customers exiting are obviously related. Therefore, we estimate $\zeta$ by matching the ex-ante (i.e., before drawing the search cost) probability of exit from the customer base delivered by the model to its counterpart in the data measured by the parameter $b_1$ in equation (10).\textsuperscript{23} Moreover, the shape of the distribution of prices is directly informative about the scale of search costs, $\lambda$: higher search costs are associated to higher kurtosis, as we illustrate in the next Section. Therefore, we pin down $\lambda$ by targeting the kurtosis of the (log-) ratio of prices $p_{\text{ratio}}$, which in our data is equal to 7.

Operationally, 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 $\Omega$, 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 $\rho_n$, $\sigma_n$, $\lambda_n$, $\zeta_n$,
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. In particular, given our definition of equilibrium and the results of Proposition 2, we look for equilibria where prices are in the interval $[p^*(\bar{z}), p^*(\bar{z})]$. In principle, our model could have multiple equilibria; however, numerically we always converge to the same equilibrium. In Appendix E we provide more details on the numerical solution of the model.

The estimation results are summarized in Table 2.

\textsuperscript{23}The equilibrium probability of exiting the customer base of a firm charging $p$ and with productivity $z$ is $G(\psi(p, z))(1 - H(V(p, z)))$ and the marginal effect of prices on the exit probability is proportional to $G'(\psi(p, z))/G(\psi(p, z))(1 - H(V(p, z)))^2$. The hazard rate $G'/G$ is decreasing in $\zeta$ for given equilibrium prices; this is exactly the case we are considering since we are targeting the persistence and volatility of the empirical price distribution, which is indeed fixed in our analysis.
Table 2: Parameter estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence of productivity process, $\rho$</td>
<td>0.3</td>
<td>Log-price autocorrelation: 0.3</td>
</tr>
<tr>
<td>Volatility of productivity innovations, $\sigma$</td>
<td>0.15</td>
<td>Log-price dispersion: 0.03</td>
</tr>
<tr>
<td>Shape parameter of search cost distribution, $\zeta$</td>
<td>2</td>
<td>Log-price kurtosis: 7</td>
</tr>
<tr>
<td>Scale parameter of search cost distribution, $\lambda$</td>
<td>0.125</td>
<td>Average marginal effect $b_1$: 0.14</td>
</tr>
</tbody>
</table>

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 predicted dynamics of the customer base growth, and then move on to discuss the implications for pricing. As shown in Remark 1, an advantage of our model is that it nests, when the scale of the search costs diverges to infinity, the standard case of an economy where demand elasticity comes only from an intensive margin associated to customers substituting across different varieties of goods (i.e. $\varepsilon_q = \varepsilon_d$). Therefore, to appreciate the importance of customer markets for equilibrium dynamics we contrasts the predictions of our model (henceforth “baseline economy”) with those of the limiting case with infinite scale of search costs, $\lambda \to \infty$, (henceforth “counterfactual economy”). To make the comparison meaningful, we fix $\theta$ in this counterfactual economy so that the resulting average total elasticity of demand is the same as in our baseline economy, and choose $\sigma$ and $\rho$ so that the counterfactual economy is observationally equivalent with respect to the volatility and autocorrelation of prices.

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 weekly net growth rate of the customer base as a function of productivity. This is the most striking difference between our model and the counterfactual economy in which there is no customer dynamics. In our setting, productivity 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. Firms with high productivity experience positive net growth of their customers base; whereas lower productivity firms are net losers of customers. It can be noticed that net customer base growth increases in productivity at a decreasing pace. This is dictated by the asymmetry between the retention and attraction margins in our model. When a firm experiences a change in productivity, the resulting price adjustment directly affects both the fraction of its own customers who decide to search and
Notes: The figure plots net customer base growth as a function of a firm’s idiosyncratic productivity, for the baseline and the counterfactual economy. The baseline economy is simulated using the parameter estimates in Table 2. The counterfactual economy’s parameters productivity process matches the same moments as in the baseline estimation (autocorrelation and volatility of the prices of customers’ baskets) but search is shut down ($\lambda \to \infty$). Therefore, in the counterfactual economy there is no extensive margin of demand. The parameter governing the intensive elasticity of demand is chosen for the counterfactual economy so that it matches the same overall elasticity of demand (intensive plus extensive margin) featured by the baseline economy.

their subset who will ultimately exit; which determine the outflow. Inflow in the customer base of the firm instead depends on the number of searching customers in the economy and on the fractions of these customers that, given that they have drawn the firm as potential new supplier, decide to enter its customer base. Idiosyncratic productivity influences, through the price, the latter but has no effect on the former. Hence, the penalizing effect of a productivity reduction is more severe than the gain from an equally sized productivity improvement. Allowing for an advertising technology would enable the firm to affect the mass of customers arriving and reinforce the link between the arrival rate of customers and firm productivity.

Quantitatively, the model predicts a yearly customer turnover of about 9%. This is lower than the 14% observed in our data. However, the latter is an unconditional figure whereas
our model only delivers an estimate of turnover due to price variation.

Figure 2: Prices and productivity

Notes: The figure plots firms’ prices as a function of idiosyncratic productivity, for the baseline and the counterfactual economy. The baseline economy is simulated using the parameter estimates in Table 2. The counterfactual economy’s parameters productivity process matches the same moments as in the baseline estimation (autocorrelation and volatility of the prices of customers’ baskets) but search is shut down ($\lambda \to \infty$). Therefore, in the counterfactual economy there is no extensive margin of demand. The parameter governing the intensive elasticity of demand is chosen for the counterfactual economy so that it matches the same overall elasticity of demand (intensive plus extensive margin) featured by the baseline economy.

Figure 2 displays the equilibrium prices in the baseline and counterfactual economies as a function of productivity. Prices are strictly decreasing in productivity in both economies. However, the pass-through of productivity shocks to prices is on average much lower in the baseline economy (17%) than in the counterfactual economy (77%).\footnote{Note that the pass-through is incomplete even in the counterfactual economy because, 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 $p$.} This explains why the support of productivity differs in the two scenarios: the counterfactual economy can generate the same dispersion in prices with lower variability of productivity shocks. The low
pass-through of idiosyncratic cost shocks predicted by the model matches, even though we did not explicitly targeted it, the pass-through of basket cost to basket price we observe in our data. Evidence of limited pass-through is also found in the literature analyzing the effect of exchange rates fluctuations on prices (Burstein and Jaimovich (2012)). The low pass-through induces positive correlation between markups with productivity, which has also been observed empirically (Petrin and Warzynski (2012)).

The shape of the curve in Figure 2 makes it apparent that the extent of pass-through is heterogenous. Firms with intermediate levels of productivity pass-through much less than both high and low productivity firms. Highly productive firms are sought after by customers; therefore they face lower customer retention concerns and behave almost like if they were only facing an the intensive margin elasticity of demand. Low productivity firms instead pass-through more because they have low margins and lower chances of retaining customers, so that trying to keep prices low by absorbing cost shocks is both costly and ineffective. Therefore, most of the effects of the competition for customers takes place for intermediate productivity ranges, where actually the vast majority of the firms lies, explaining why the pass-through of costs shocks is much lower on average than in the counterfactual economy.

Figure 3 plots the distribution of prices in the baseline and counterfactual economies. Recall that we imposed for these two distributions to have the same standard deviation but we put no restriction on their means, as we do not target markups in our estimation procedure. Therefore, in order to make the distributions comparable, we standardize them dividing by their mean.

The shape of the price distribution is not a result of the model: we estimated the parameters targeting the leptokurtic distribution of prices in our data. However, the rationale we just presented to explain heterogeneity in pass-through provides a good intuition of why our model can naturally generate leptokurtic price distributions and ultimately justifying our choice of the kurtosis as a key moment to identify the quantitative relevance of customer markets.

As we have seen, a large mass of firms with intermediate levels of productivity faces significant elasticity of their customer base. For these firms, setting a price higher than the competition is very costly; therefore in equilibrium their prices will be very close to each other. This results in a high density of prices close to the mean. At the same time, there are firms that do not try to stay close to the average price either because they are more

\[\text{We measure the pass-through by regressing households’ basket prices on their contemporaneous cost, on five lags of the cost and a full set of time dummies. The resulting cumulated pass-through, obtained summing the coefficients on current and lagged cost, is 16%.}\]

\[\text{See Kaplan and Menzio (2014) for additional evidence on the price distribution of homogeneous packaged goods.}\]
Notes: The figure plots the distribution of prices, standardized by their mean, for the baseline and the counterfactual economy. The baseline economy is simulated using the parameter estimates in Table 2. The counterfactual economy’s parameters productivity process matches the same moments as in the baseline estimation (autocorrelation and volatility of the prices of customers’ baskets) but search is shut down ($\lambda \to \infty$). Therefore, in the counterfactual economy there is no extensive margin of demand. The parameter governing the intensive elasticity of demand is chosen for the counterfactual economy so that it matches the same overall elasticity of demand (intensive plus extensive margin) featured by the baseline economy.

Efficient and it is optimal for them to set a lower price or because their productivity is so low that it is unfeasible for them to mimic the pricing of better competitors. These two groups account for the fat tails of the price distribution. The presence of both peakedness and fat tails is the distinctive characteristic of leptokurtic distributions. The fact that the presence of customer markets is the driver of the kurtosis of the price distribution is confirmed by looking at the shape of the price distribution in the counterfactual economy, where this margin is shut down. In that case, prices are almost normally distributed. Finally, it is worth noting that the shape of the price distribution in our model significantly departs from that of other models featuring equilibrium price dispersion like, for instance, Burdett and Judd (1983). In this case, however, the difference is not only due to the presence of customer retention concerns but also to other divergences in the modeling approach.
6 Exploring the predictions for aggregate dynamics

In this section we consider the pass-through of an unexpected persistent change in production cost affecting a subset of the firms in the economy on the grounds that this scenario is analogous to a shock to the real exchange rate. Simulating the effect of a real exchange rate shock allows to stack the predictions of our model against the findings of a large literature that has analyzed both empirically and theoretically the associated pass-through and the effects of persistent shocks on competition between foreign and domestic producers. A salient result of the exercise that we perform is that the response of the imports ratio -defined as the ratio of \( d \) goods bought from foreign firms to \( d \) goods home produced- to the aggregate shock is larger in the long run than in the short run. This is interesting as it is consistent with the empirical evidence in international macroeconomics (see Ruhl (2008)), and given the well-known inability of static demand models to address the issue.

A depreciation of the real exchange rate -defined as the price the domestic numeraire good \((n)\) in units of the foreign numeraire- implies de facto a decrease in labor cost for a subset of firms in the market (the foreign firms) relatively to another group (the domestic firms). Shocks affecting a subset of the suppliers of good \(d\) are particularly interesting in our environment as these shocks give rise to an increase in price dispersion and, therefore, in customers searching effort. We show that the general equilibrium feedback from customers searching to firms optimal pricing, as well as the inertia in the response of the customer base, substantially changes the transmission of these shocks with respect to static models of demand, such as those relying in monopolistic competition. In our model, aggregate shocks have a larger effect on the long run than in the short run, while the opposite is true in models of static demand. A similar result is obtained in the model of Drozd and Nosal (2012). Although both models deliver the result through a sluggish response of the customer base, the mechanism underpinning the inertia is quite different. In their model, firms invest in a marketing technology that affects the arrival rate of new customers, therefore affecting in a direct way the speed at which the customer base grows. Moreover, in their setup, firms negotiate individually the terms of trade with each customer, implying a useful disconnection.

---

27A shock to the effective cost of a subset of the players in the industry is the salient characteristic of a number of other real world scenarios. For instance, variations in state sales tax would affect local online sellers but not those located out-of-state, as the latter cannot be compelled to collect it (Einav et al. (forthcoming)). A similar effect is generated by the introduction of size-contingent employment protection legislation (Acemoglu and Angrist (2001), Schivardi and Torrini (2008)).

28Ruhl (2008) and Ghironi and Melitz (2005) extend models of static demand to address the empirical finding. Their extensions involve a disconnection of firm’s behavior with respect to the nature of the shock (i.e. transitory of permanent) as a result of a fixed cost of participation. In ours, this disconnect is not present.

29Given that we assume perfect competition in the market for the numeraire, the equilibrium wage coincides with the price of the numeraire.
for the purposes of their model, between customer dynamics and pricing. In ours, the unique choice of a firm is its current price which does not affect the speed at which customers arrive, rather affecting how many of those in contact with a given firm are willing to remain in its customer base. The fact that the unique choice of a firm is its price implies a tight connection between the build up of a customer base (extensive margin) and current profit maximization (intensive margin).

The specifics of the experiment that we perform are the following. We consider our economy in steady state at period \( t_0 \), as calibrated in Section 5. We assume that a fraction \( \alpha \in (0, 1) \) of the producers of good \( d \) are foreign firms denoted by \( j^* \), whereas the rest are domestic firms and are denoted by \( j \). Foreign firms produce in a symmetric country, characterized by the same parameter values of the domestic economy. For simplicity, we assume that the fraction and identity of firms exporting to the domestic country is a fixed and representative sample of the firms producing good \( d \) in the foreign country. The numeraire, i.e. good \( n \), is non-tradable. Given the symmetric structure, the steady state equilibrium of this simple open economy is identical to the steady state characterized in Section 5. We set the fraction of foreign firms in the economy to \( \alpha = 12\% \), which matches the share of imports into U.S. personal consumption expenditure (Hale and Hobijn (2011)). We hit foreign firms with an unexpected and unforeseen shock to production cost in the form of a scaling factor \( \tau \), so that the marginal cost of production of a foreign firm \( j^* \) goes from \( w_{t_0}^*/z_{t_0}^* \) to \( \tau w_{t_0}^*/z_{t_0}^* \). This shock is realized after the firm has learned about idiosyncratic productivity \( z \), but before pricing and customer’s exit decisions are taken, and dies out according to an AR(1) process, \( \tau_t = \rho \tau_{t-1} \) for \( t > t_0 \). Since the shock implies aggregate dynamics, we augment our economy with a simple equilibrium model of the labor market to capture the general equilibrium effects of the shock on wages and income. In particular, we assume that labor supply is set by a representative household in a perfectly competitive market, who then distributes labor income equally across the large mass of shoppers who take the consumption decisions. Given the numeraire good is sold in a perfectly competitive market, the equilibrium wage is pinned down by the marginal product of labor in the numeraire sector. See Section D for details.

In Figure 4 and Figure 5 we plot the responses of several variables of interest to a 1% increase in the productivity of foreign firms which, given the assumed process for the shock, implies a half life of approximately a quarter for the aggregate shock. Figure 4 displays the responses of the total demand of good \( d \) in the domestic economy, as well as the two components of demand satisfied by domestic production, i.e. \( M \), and by foreign production, i.e. \( M^*D^* \), where \( M^* \equiv \int m^* dj^* \) and \( M \equiv \int m dj \) denote the measures of customers of foreign and domestic firms respectively, and where \( D^* \) and \( D \) denote the analogous for the demand per customer. Figure 5 plots the response of the weighted average price of good \( d \),
as well as the weighted average price responses of domestic and foreign firms. Both total demand and average price experience modest changes, from their steady state values, as a result of the shock. The fact that total demand barely increases is a consequence of the modest decrease in the average price. There are two reasons behind the modest fall in prices. The first one is mechanical: Because foreign firms account for a small share of aggregate output in the economy, their price changes have small effect on the average economy-wide price. The second reason is due to the presence of customer markets. As the productivity shock is persistent, customers would like to be part of the customer base of foreign firms as they will likely set lower prices in the future than domestic firms. As a result, on average, foreign firms face a slacker extensive margin of demand, and therefore they pass-through to customers only a portion of the improvement of their production cost.

Figure 4: The response of demand to an exchange rate shock in the baseline economy

Notes: The figure plots the response of different measures of total quantity demanded of good d to a 1% increase in the productivity of “foreign firms” in our baseline economy, expressed in % deviations from the steady state. The productivity increase dies out according to a weekly autocorrelation coefficient of 0.95. The blue line refers to the response of total demand. The red dashed line refers to the response of demand from foreign producers. The black line refers to demand from domestic producers.

However small in magnitude, the price change of foreign firms have important conse-
Figure 5: The response of prices to an exchange rate shock in the baseline economy

![Graph showing the response of different measures of prices of good d to a 1% increase in the productivity of “foreign firms” in our baseline economy, expressed in % deviations from the steady state. The productivity increase dies out according to a weekly autocorrelation coefficient of 0.95. The blue line refers to the response of the average price paid by all customers in the economy. The red dashed line refers to the response of the average price paid by customers at foreign firms. The black line refers to the response of average price paid by customers at domestic firms.]

Notes: The figure plots the response of different measures of prices of good d to a 1% increase in the productivity of “foreign firms” in our baseline economy, expressed in % deviations from the steady state. The productivity increase dies out according to a weekly autocorrelation coefficient of 0.95. The blue line refers to the response of the average price paid by all customers in the economy. The red dashed line refers to the response of the average price paid by customers at foreign firms. The black line refers to the response of average price paid by customers at domestic firms.

The persistence of the effect of the shock is intimately related to the extensive margin of demand. In Figure 6 we decompose the response of the imports ratio to the aggregate shock with respect to the margin of adjustment: Intensive or extensive. This decomposition follows from the formula for the imports ratio. Notice that we can rewrite its expression as \((M^*/M)(D^*/D)\), where the ratio \(M^*/M\) accounts for the effect of the extensive margin of demand for the distribution of demand between foreign and domestic firms. The plot shows that, although the ratio of prices barely moves, there are persistent changes in the imports ratio. On impact, the ratio increases by approximately 0.2%, and then persistently increases, at a diminishing rate, as the aggregate shock fades out. Interestingly, this shows that a transitory aggregate shock can have persistent effects on the economy, substantially altering the composition of demand, displaying a larger elasticity of the imports ratio to an exogenous aggregate shock in the long run than in the short run.
demand, through the movement of customers from domestic to foreign firms, while $D^*/D$ accounts for the intensive margin of demand. As the plot shows, the intensive margin explains most of the effect on the short run, while the extensive margin accounts for the long term effects of the shock. The fact that the intensive margin is the main driver of the short run effects is a consequence of both the demand being downward slopping -and foreign firms posting lower prices on impact-, and the process of reallocation of customers being slow. The fact that the extensive margin explains the long run behavior of the imports ratio is also clear. Because the shock is persistent, customers understand that foreign firms will charge, on average, lower prices than domestic firms for some time. As a result, they want to reallocate their shopping activities towards foreign firms. As the process of finding foreign firms takes time, this margin is less relevant in the short run, but explains a great deal of the change in the import ratio in the long run. As the productivity shock vanishes, customers that reallocated their shopping activities towards foreign firms do not have incentives to return to domestic firms as in steady state both economies share the same productive properties. This implies that the transitory shock to productivity of foreign firms has persistent consequences for the relative importance of imports, and this occurs through the extensive margin of demand.

Finally, in Figure 7 we compare the response of the imports ratio in our baseline economy -displaying an active extensive margin- to the response implied by the economy with the inactive extensive margin developed in Section 5.2. Although indistinguishable with respect to total demand elasticity, both economies differ a great deal on how they respond to the aggregate shock. As the figure shows, the inelastic economy displays a transitory increase in the imports ratio that peaks on impact, while in our baseline economy the imports ratio exhibits a persistent increase that peaks in the long run. What drives the differences is the extensive margin of demand. In the inelastic economy, foreign firms passthrough a large portion of the shock to prices as their sales level is regulated solely by the intensive margin elasticity, while in the baseline economy foreign firms passthrough less of the shock to prices, as a result of a slacker extensive margin of demand.

7 Conclusions

TBA

References

Figure 6: The response of imports to an exchange rate shock in the baseline economy

Notes: The figure plots the response of different statistics of imports of good $d$ to a 1% increase in the productivity of “foreign firms” in our baseline economy, expressed in % deviations from the steady state. The productivity increase dies out according to a weekly autocorrelation coefficient of 0.95. The blue line refers to the response of the ratio of total imports to total sales of domestic producers. The red dashed line refers to the response of the ratio of total customers buying from foreign firms to total customers being from domestic producers. The black line refers to the ratio of average demand of customers from foreign firms to average demand of customers from domestic producers.


Figure 7: The response of imports to an exchange rate shock: Baseline vs Counterfactual

Notes: The figure plots the response of the ratio of total imports to total sales of domestic producers of good \(d\) to a 1% increase in the productivity of “foreign firms” in our baseline economy, expressed in % deviations from the steady state. The productivity increase dies out according to a weekly autocorrelation coefficient of 0.95. The blue line refers to our baseline economy. The red dashed line refers to the counterfactual economy with no extensive margin (\(\lambda \to \infty\)), and parametrized to be observationally equivalent with respect to total demand elasticity, volatility and autocorrelation of prices.


Coibion, O., Gorodnichenko, Y. and Hong, G. H. (2012). The cyclicality of sales, regular


A Proofs

A.1 Proof of Lemma 1

Customer’s decisions are sequential: first she decides if to incur in the search cost $\psi$ 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 $\bar{V}(p, z) \leq \bar{V}(p(z'), z')$, 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 $\psi$. Because the value function in the case of searching is decreasing in $\psi$ and the value function in the case of not searching does not depend on $\psi$, 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 $\hat{V}(p, z)$ 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 $\hat{V}(p, z)$ 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 $\frac{\partial \hat{V}(p, z)}{\partial p} = -v'(p) + \frac{\partial \bar{V}(p, z)}{\partial p} \geq 0$ by the definition of $\tilde{V}(p, z)$ and $v(p)$ being decreasing in $p$. Also, we have that

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

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 $\hat{V}(\bar{p}(z), z) = \max_{z \in [\hat{p}(z), \bar{p}(z)]} \{ \hat{V}(P(z), z) \}$. 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 $\bar{p}(z) \in [\bar{p}(z), p^*(z)]$ if $\bar{p}(z) < p^*(z)$, and $\bar{p}(z) = p^*(z)$ otherwise. Next, we show that $\bar{p}(z) < p^*(z)$ if $\bar{p}(z) < p^*(z)$. The results follows because at $p = \bar{p}(z)$ $W(p, z)$ is strictly decreasing in $p$ as by
definition of $\hat{p}(z)$, 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 $Q(x) = \Gamma \int^x G \left( \int^\infty_s (u-x) dH(u) \right) dH(u)$ is continuously differentiable in $p$. Thus, $\Delta(p, z)$ 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:

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

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 (9) can be rewritten as

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

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

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

A direct implication of nonnegative prices is that $\varepsilon_q(p, z) - 1 + \varepsilon_m(p, z)x(p, z) \geq 0$, so that sign $[\partial \mu(p, z)/\partial \varepsilon_m(p, z)] = \text{sign}[x(p, z)](1 - \mu(p, z))]$. 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 $x(p, z) > 0$ 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 $x(p, z) < 0$ and, therefore, $x(p, z) < 0$. As a result, $\partial \mu(p, z)/\partial \varepsilon_m(p, z) < 0$. 

41
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 \( \hat{p}(z) \geq p^*(\tilde{z}) \), we have that \( \bar{p}(z) = p^*(\tilde{z}) \) for each \( z \). From Proposition 1 we know that, for a given \( z \), the optimal price \( \hat{p}(z) \) belongs to the set \( [p^*(\tilde{z}), \ p^*(z)] \). Over this set, the objective function of the firm,

\[
W(p, z) = \Delta(p, z) (\pi(p, z) + \beta \text{ constant}) , \tag{11}
\]

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 \( z_1, z_2 \in [\tilde{z}, \bar{z}] \) 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 \( \bar{p}(z) \) 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 \( \bar{p}(z) \) is upper hemi-continuous and \( W(\hat{p}(z), z) \) is continuous in \( z \). Given that \( \hat{p}(z) \) is non-increasing in \( z \) it follows that \( \hat{p}(z) \) has a countably many discontinuity points. We proceed as follows. Let \( \hat{P}(z) \) be the set of prices that maximize the firm problem. Whenever a discontinuity arises at some \( \tilde{z} \) (so that \( \hat{P}(\tilde{z}) \) is not a singleton) we modify the optimal pricing rule of the firm and consider the convex hull of the \( \mathcal{P}(\tilde{z}) \) as the set of possible prices chosen by the firm with productivity \( \tilde{z} \). The constructed mapping from \( z \) to \( \hat{P}(z) \) is then upper-hemicontinuous, 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 decreasing and differentiable in $z$ in any neighborhood of the first order condition. Finally, given that $\hat{p}(z)$ has a countably many discontinuity points, it has countably many points where it is not differentiable, and the first order condition does not apply at those points, but applies everywhere else. These points have measure zero with respect to the density of the $z$ and therefore $\hat{p}(z)$ is almost everywhere differentiable.

**Points (i)-(ii).** 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 follows by contradiction. Consider, in the contradiction, that $\hat{p}(z_1) = \hat{p}(z_2) = \bar{p}$ for some $z_1, z_2 \in \mathbb{Z}_{\geq 0}$. Also, without loss of generality, assume that $z_1 < z_2$. Given that we already established the necessity of the first order condition presented in equation (7) -when prices are monotonic, suppose that it is satisfied at the duple $\{z_2, \bar{p}\}$. Notice that, because of the assumed i.i.d. structure of productivity shocks together with $\pi_z(p, z) < 0$, it is not possible that the first order condition is also satisfied at the duple $\{z_1, \bar{p}\}$. Moreover, because the first order condition is necessary and we already established that $\hat{p}(z)$ cannot be increasing at any $z$, we conclude that the optimal price at $z_1$ is strictly larger than at $z_2$. That is, $\hat{p}(z_1) > \hat{p}(z_2)$. Notice that this verifies the conjecture used to prove the necessity of the first order condition, which in turn validates the use of equation (7) here.\(^{30}\)

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

\(^{30}\)If prices are not strictly decreasing this argument cannot be used as the first order condition is not necessary. However, it is possible to prove that $\hat{p}(z)$ is strictly decreasing in $z$ for some region of $z$. The argument follows by contradiction. Suppose, in the contradiction, that $\hat{p}(z)$ is everywhere constant in $z$ at some $\bar{p}$. Then $\hat{p}(z) = \bar{p}$ for all $z$. If $\bar{p} > p^*(\bar{z})$, then $\bar{p}$ would not be optimal for firm with productivity $\bar{z}$ which would choose a lower price. If $\bar{p} = p^*(\bar{z})$, then continuous differentiability of $G$ together with $H = G = Q = 0$ at the conjectured constant equilibrium price, imply that the first order condition is locally necessary for an optimum, and a firm with productivity $z < \bar{z}$ would have an incentive to deviate according to equation (7), and set a strictly higher price than $\bar{p}$. 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 $\hat{p}(z) \geq \bar{p}(z)$ and $\hat{p}(z) = \bar{p}(z)$ for all $z$. 
We now prove part (ii). \( \hat{\psi}(p, z) \geq 0 \) immediately follows its definition. The fact that \( \hat{V}(z) \)

is strictly in \( z \), together with Lemma 2, immediately implies that \( \hat{\psi}(\hat{p}(z), \hat{z}) = 0 \) and that 

\( \hat{\psi}(\hat{p}(z), z) \) is strictly increasing in \( z \). Finally, Lemma 3 implies that \( \Delta(\hat{p}(z), z) \) is increasing

in \( z \). Because of price dispersion, some customers are searching, which guarantees that 

\( \Delta(\hat{p}(z), \hat{z}) > 1 \). Likewise, \( \Delta(\hat{p}(z), \hat{z}) < 1 \).

A.6 Proof of Remark 1

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

Part (2). From Proposition 2 we have that, in equilibrium, the highest price is \( \hat{p}(z) \). Moreover, under the assumptions of the Proposition 2, the first order condition is a necessary condition for optimality of prices. We will use this to show that, as \( n \) approaches zero, \( \hat{p}(z) \) has to approach \( \hat{p}(\hat{z}) = p^*(\hat{z}) \).

In equilibrium, it is possible to rewrite the first order condition (i.e. equation (7)) evaluated at \( \{\hat{p}(z), \hat{z}\} \) as \( LHS(\hat{p}(\hat{z}), n) = RHS(\hat{p}(\hat{z}), n) \), where

\[
LHS(\hat{p}(\hat{z}), n) \equiv G' \left( \frac{\hat{\psi}(\hat{p}(\hat{z}), \hat{z})}{n} \right) \frac{\hat{\psi}_p(\hat{p}(\hat{z}), \hat{z})}{n} \]

\[
+ \left( G \left( \frac{\hat{\psi}(\hat{p}(\hat{z}), \hat{z})}{n} \right) H'(\hat{V}(\hat{p}(\hat{z}), \hat{z})) + \frac{1}{\Gamma} Q'(\hat{V}(\hat{p}(\hat{z}), \hat{z})) \right) \hat{V}_p(\hat{p}(\hat{z}), \hat{z}) ,
\]

\[
RHS(\hat{p}(\hat{z}), n) \equiv - \frac{\pi_p(\hat{p}(\hat{z}), \hat{z})}{\Pi(\hat{p}(\hat{z}), \hat{z})} \left( 1 - G \left( \frac{\hat{\psi}(\hat{p}(\hat{z}), \hat{z})}{n} \right) \right) ,
\]

given that \( H(\hat{V}(\hat{p}(\hat{z}), \hat{z})) = Q(\hat{V}(\hat{p}(\hat{z}), \hat{z})) = 0 \).

Suppose that as \( n \downarrow 0 \), \( \hat{\psi}(\hat{p}(\hat{z}), \hat{z}) \) does not converge to zero. Then, \( G \left( \frac{\hat{\psi}(\hat{p}(\hat{z}), \hat{z})}{n} \right) \uparrow 1 \) as \( n \downarrow 0 \). This implies that \( \lim_{n \downarrow 0} RHS(\hat{p}(\hat{z}), n) > 0 \).

Consider now the function \( LHS(\hat{p}(\hat{z}), n) \). Again, suppose that as \( n \downarrow 0 \), \( \hat{\psi}(\hat{p}(\hat{z}), \hat{z}) \) does not converge to zero. Notice that the second term of the function approaches a finite number as \( \hat{V}_p(\hat{p}(\hat{z}), \hat{z}) \) is bounded by assumptions on \( v(p) \) and \( H'(\hat{V}(\hat{p}(\hat{z}), \hat{z})) \) and \( Q'(\hat{V}(\hat{p}(\hat{z}), \hat{z})) \) being bounded as a result of Proposition 2. Moreover, as long as \( \hat{p}(\hat{z}) > \bar{p}(z) = p^*(\hat{z}) \), we have that \( \hat{\psi}_p(\hat{p}(\hat{z}), \hat{z}) > 0 \) so that \( \hat{\psi}_p(\hat{p}(\hat{z}), \hat{z})/n \) diverges as \( n \) approaches zero. This means that
$G'\left(\frac{\hat{\psi}(\hat{p}(\bar{z}), \bar{z})}{n}\right) \frac{\hat{\psi}(\hat{p}(\bar{z}), \bar{z})}{n}$ is divergent and therefore the first order condition cannot be satisfied.

This analysis concluded that, if $\hat{\psi}(\hat{p}(\bar{z}), \bar{z})$ does not converge to zero as $n$ becomes arbitrarily small, the first order condition, i.e. equation (7), cannot be satisfied. This occurs because $LHS(\hat{p}(\bar{z}), n)$ would diverge to infinity, while $RHS(\hat{p}(\bar{z}), n)$ would remain finite. It then follows that, as $n$ approaches zero, a necessary condition is that $\hat{\psi}(\hat{p}(\bar{z}), \bar{z})$ also approaches zero. This condition can be restated as requiring that $\hat{p}(\bar{z})$ approaches $\bar{p}(z)$ as $n$ approaches zero. Moreover, given the assumptions of Proposition 2, $\bar{p}(z) = \hat{p}(\bar{z}) = p^*(\bar{z})$.

In the end, if $\hat{p}(\bar{z})$ approaches $p^*(\bar{z})$ as $n$ becomes arbitrarily small (so that $\hat{\psi}(\hat{p}(\bar{z}), \bar{z}) \to 0$ and $\hat{\psi}_p(\hat{p}(\bar{z}), \bar{z}) \to 0$), we have that $\lim_{n \to 0} LHS(\hat{p}(\bar{z}), n) < \infty$ and $\lim_{n \to 0} RHS(\hat{p}(\bar{z}), n) < \infty$ as $\pi_p(p^*(\bar{z}), \bar{z})$ is bounded as $\pi(p^*(\bar{z}), \bar{z}) > 0$. However, if $\hat{p}(\bar{z})$ does not approach $p^*(\bar{z})$ as $n$ becomes arbitrarily small, we have that $LHS(\hat{p}(\bar{z}), n)$ diverges as $n$ approaches zero, while $LHS(\hat{p}(\bar{z}), n)$ remains finite. As the first order condition has to be satisfied in equilibrium, a necessary condition is that, as $n$ approaches zero, the highest price in the economy, i.e. $\hat{p}(\bar{z})$, has to approach the lowest price in the economy, i.e. $p^*(\bar{z})$.

### B Data sources and variables construction

#### B.1 Data and selection of the sample

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. This implies that we can observe our agents behavior only when they shop with the chain; on the other hand, cash register data contain significantly less measurement error than databases relying on home scanning (Einav et al. (2010)).

The main data source contains 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. Data are recorded through usage of the loyalty card; 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.

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. This has important implications for us as our definition of basket requires us to be able to attach a price to each of the item composing week in every week, even when the customer does not shop. The issue can be solved using another dataset with information on weekly store revenues and quantities between January 2004 and December 2006 for a panel of over 200 stores. For each good (identified by its UPC) carried by the stores in those weeks, the data report total amount grossed and quantity sold. Exploiting this information we can 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 main 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.

On top of reporting revenues and quantities for each store-week-upc triplet, the store level data also contain a measure of cost. This variable is constructed on the basis of the estimated markup imputed by the retailer for each item and includes more than the simple wholesale cost of the item (the share of transportation cost, etc.). Eichenbaum et al. (2011) suggest to think about it as a measure of replacement cost, i.e. the cost of placing an item on the shelf to replace an analogous one just grabbed by a consumer. We use this measure to construct our instrument of the basket price.

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

A final piece of the data is represented by the IRI-Symphony database. We use store level data on quantities and revenues for each UPC in 30 major product categories for a large sample of stores (including small and mom&pop ones) in 50 Metropolitan Statistical Areas in the US. The data allow to construct unit value prices for all the stores competing with the chain who provided the main dataset. However, the coarse geographic information prevents us from matching each customer with the stores closer to her location (in the same zipcode, for instance) and forces us to adopt the MSA as our definition of a market.

**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

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

**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 purchased at store \(j\) in week \(t\) is computed as:

\[
p_{ijt} = \sum_{k \in K_i} \omega_{ik} p_{kt} , \quad \omega_{ik} = \frac{\sum_t E_{ikt}}{\sum_k \sum_t E_{ikt}} ,
\]

where \(K^i\) is the set of all the UPC’s \((k)\) purchased by household \(i\) during the sample period, \(p_{kt}\) is the price of a given UPC \(k\) in week \(t\) at the store \(j\) where the customer shops. \(E_{ikt}\) represents expenditure by customer \(i\) in UPC \(k\) in week \(t\) and the \(\omega_{ik}\)’s are a set of household-UPC specific weights. The construction of the cost of the basket follows the same procedure where we substitute the unit value price with the measure of replacement cost provided by the retailer.
We choose to calculate the weights using the total expenditure in the UPC by the household over the two years in the sample. This can lead to some inaccuracy in identifying the goods the customer cares for at a given point in time. For example, if a customer bought only Coke during the first year and only Pepsi during the second year of data, our procedure would have us give equal weight to the price of Coke and Pepsi throughout the sample period. If we used a shorter time interval, for example using the expenditure share in the month, we would correctly recognize that she only cares about Coke in the first twelve months and only about Pepsi in the final twelve months. However, weights computed on short time intervals are more prone to bias induced by pricing. For example, a two-weeks promotion of a particular UPC may induce the customer to buy it just because of the temporary convenience; this would give the UPC a high weight in the month. The effect of promotion is instead smoothed when we compute weights using expenditure over the entire sample period.

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

\[
p_{it}^{\text{competitors}} = \sum_{z \in M^i} s_z \sum_{k \in K_i} \omega_{ik} p_{kzt}, \quad \omega_{ik} = \frac{\sum_t E_{ikt}}{\sum_k \sum_t E_{ikt}}, \quad s_z = \frac{\sum_t R_{zt}}{\sum_{z'} \sum_t R_{z't}}
\]

where \(M^i\) is the MSA of residence for customer \(i\) and \(R_{zt}\) represents revenues of store \(z\) in week \(t\). In other words, in the construction of the competitors’ price index, stores with higher (revenue based) market shares weight more.

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, \(C \equiv \left( \sum_{k=1}^{K} c_k^{\theta} \right)^{\frac{1}{\theta}}, \theta > 1\). 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

$$c_k(p_k) = \frac{1}{P} \left( \frac{\nu}{P} \right)^{-\theta} \text{, where } P \equiv \left( \sum_{k=1}^K p_k^{1-\theta} \right)^{-\frac{1}{\theta-1}} \text{ is a price level that solves } P \ C = I.$$  

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 $b'(p_k) = \theta b(p_k) \left( b(p_k) - b(p_k)^{\frac{1}{\theta}} \right)$. Moreover, in a symmetric equilibrium, where $p_k = p$ for all $k$, $b(p) = K^{\frac{\theta}{\theta-1}}$ and $b'(p) = \frac{\theta}{kp} \left( K^{\frac{\theta}{\theta-1}} - K^{\frac{1}{\theta-1}} \right)$.  

**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,

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

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

which, in a symmetric equilibrium, reduce to

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

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

where we also used that, in a symmetric equilibrium, $P = K^{\frac{1}{\theta-1}} p$. 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.
Indirect utility function, \( u(C(p_k)) \). This is analogous, through the lens of the model, to evaluating \( v(p) \). Using the definition of \( c_k(p_k) \) together with \( P C = I \), we can obtain the following expressions,

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

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

\[
\frac{\partial^2 u(C)}{\partial p_k^2} = -C^{-\gamma-1} \left( \frac{\partial C}{\partial p_k} \right)^2 \left[ \gamma - \theta \left( 1 - b(p_k) \frac{1-\theta}{\theta} \right) + 2 \right]
\]

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

\[
\varepsilon_d(p_k) = -\frac{\partial \log c_k(p_k)}{\partial \log p_k} = \theta - (\theta - 1) \frac{c_k(p_k)p_k}{I},
\]

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 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,

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

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

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

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 $\partial \pi(p, z)/\partial p = 0$. 
characterizes its maximizer.

D A simple model of the labor market

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 $\Gamma$ of shoppers/customers, and a representative worker. The preferences of the household are given by

$$E_t \left[ \int_0^\Gamma V_t(p_i^t, z_i^t, \psi_i^t) \, di - J_t \right],$$

where $V_t(p_i^t, z_i^t, \psi_i^t)$ 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 $\ell_T$. 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 sub-script $t$ for the value function.

The worker chooses the path of $\ell_t$ 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

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

s.t. $$p_t d + q_t n \leq I_t,$$

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

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

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 \( P_{t+s}^i = ((p_{t+s}^i)^{1-\theta} + (q_{t+s})^{1-\theta})^{1/\theta} \) 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 \( y_n = Z_t \ell_n \), where \( Z_t \) is aggregate productivity, and \( \ell_n \) 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 \( y_j = Z_t z_j \ell_j \), where \( Z_t \) is aggregate productivity, and \( \ell_j \) 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 \( \ell_t = \ell_n + \int_0^1 \ell_j \, dj \).

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 Numerical solution of the model

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

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 \( \mathcal{P}(z) \in [p^*(z), p^*(\bar{z})] \) 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 \( \Delta(p(z), z) > 0 \) ensure the existence of a unique \( K(z) \) that solves...
equation (8). Finally, given $Q(\cdot)$, $H(\cdot)$, $\mathcal{P}(z)$ and $\tilde{V}(p, z)$, we solve the firm problem and
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 $z \in \left[\underline{z}, \bar{z}\right]$. 