

# Firm Dynamics and Pricing under Customer Capital Accumulation\*

Pau Roldan-Blanco<sup>†</sup>      Sonia Gilbukh<sup>‡</sup>

December 14, 2019

## Abstract

In a search model of firm dynamics, customer accumulation is shown to affect the lifecycle of firms and the dynamics of markups in response to aggregate demand shocks. In the model, sellers of different sizes and productivities post dynamic pricing contracts to strike a balance between attracting new customers and exploiting pre-existing ones. Calibrated to fit establishment-level and micro-pricing data from the U.S. retail sector, the model predicts that markups are procyclical to demand shocks, with a larger contribution by small sellers. The responses are weaker on impact, and more persistent in the transition, when customer-seller relationships last longer.

**JEL codes:** D21, D83, E2, L11.

**Keywords:** Customer Capital, Directed Search, Firm Dynamics, Dynamic Contracts.

---

\*We are greatly indebted to Ricardo Lagos, Jess Benhabib, Boyan Jovanovic, and Edouard Schaal, for their invaluable guidance and support. For insightful comments and suggestions, we thank Joseph Vavra (Associate Editor) and an anonymous referee, as well as Isaac Baley, Jaroslav Borovička, Luís Cabral, Laurent Cavenaile, Diego Daruich, Simon Gilchrist, Philipp Kircher, Julian Kozlowski, Espen Moen, Simon Mongey, Jesse Perla, Tom Schmitz, Nicholas Trachter, Javier Turen (discussant), Venky Venkateswaran, Ewout Verriest, and Basil Williams, as well as seminar audiences in various conferences and institutions. This research was financially supported by a McCracken Doctoral Fellowship from New York University and a Ph.D. Dissertation Internship from the FRB St. Louis. All estimates and analyses in this paper based on Information Resources Inc. (IRI) data are by the authors and not by IRI. The views expressed in this paper are those of the authors and do not necessarily coincide with those of the Banco de España or the Eurosystem.

<sup>†</sup>*Corresponding author:* [pau.roldan@bde.es](mailto:pau.roldan@bde.es). Banco de España, Calle Alcalá 48, 28014 Madrid.

<sup>‡</sup>CUNY Baruch College, Zicklin School of Business, 137 East 22nd Street, New York, NY 10010.

# 1 Introduction

Firms of different sizes and ages experience persistently different growth paths along their life cycle. Newly established businesses typically start out small relative to their more mature competitors, and this gap takes time to close (e.g. Dunne *et al.* (1988), Caves (1998), Cabral and Mata (2003)). A large theoretical literature, inspired by the seminal work of Jovanovic (1982) and Hopenhayn (1992), has attributed this evidence to a selection process on the basis of productivity differences across firms. Recently, however, this view has been challenged by a growing literature arguing that the productivity-based interpretation of firm heterogeneity may be confounding selection on technological productivity with selection on profitability (e.g. Foster *et al.* (2008)). Indeed, new empirical evidence from micro data has shown that there remains a large cross-sectional dispersion in firm revenue after controlling for heterogeneity in technological productivity. In light of this, the literature has suggested that the variation in firm performance in the cross-section and over time stems, to a great extent, from differences in idiosyncratic demand components. Further, the evidence suggests that this demand-side channel of variation is persistent.<sup>1</sup> These findings point to firm investment in demand as a key driver behind the existing differences in the lifecycle of businesses of similar productivity. Yet, few studies have formalized the firm dynamics implications of these investment decisions, or their consequences at the aggregate level.

This paper aims to fill this gap by developing a theory of firm dynamics in product markets with aggregate shocks. The model introduces a meaningful role for demand accumulation at the firm level, which is interpreted as the formation of a customer base. Section 2 presents a directed search model of the product market in which a fixed measure of ex-ante identical buyers must search for sellers of different productivities offering the same homogenous product. All sellers post, and commit to, long-term pricing contracts designed to attract new potential customers, and retain pre-existing ones. In the recursive formulation of the model, contracts are composed of two objects: a price level, which the seller charges to all of her incumbent customers, and a continuation utility, which she promises to deliver in expectation to all of its future customers. The seller's promise endogenously determines the rate at which new customers arrive, while being sufficiently generous to prevent any of the pre-existing buyers from voluntarily separating. This generates ex-post heterogeneity in seller size, even within the same productivity type, in line with the aforementioned empirical observations. On the other hand, sellers may lose customers or exit the market altogether due to exogenous separation shocks. When a seller loses all its customers, it may re-enter the product market by paying a fixed market-penetration cost. Overall, sellers move endogenously through the size space as a result of their pricing decisions.

In equilibrium, sellers strike a balance between instantaneous revenues (via high prices today)

---

<sup>1</sup> These findings have been documented in the U.S. for the retail sector (Hottman *et al.* (2016)), as well as for commodity-like markets (Foster *et al.* (2008, 2016)). Related evidence has been found for other countries, including Carlsson *et al.* (2017) (Sweden), Pozzi and Schivardi (2016) (Italy), Fitzgerald *et al.* (2017) (Ireland), Kaas and Kimasa (2018) (Germany), and Kugler and Verhoogen (2012) (Colombia). A related literature argues that markups are important contributors to revenue dispersion (e.g. Hsieh and Klenow (2009), DeLoecker (2011), Peters (2018)).

and future market shares (via lower prices in the future).<sup>2</sup> This inter-temporal trade-off determines the rate of seller growth. Section 2.4 argues that the way the trade-off is resolved depends on the size of the seller’s customer base. Because of the market penetration costs, sellers optimally decide to promise high continuation values to their buyers when they are small in order to generate a high rate of expansion, and thereby raise enough resources to grow out of the entry cost. As these sellers grow further, they lower their future promises and increase their prices as they increasingly prefer to extract rents from their existing base rather than from incoming customers. Therefore, markups increase as sellers grow in size, and the seller’s rate of growth slows down. As this growth process is counteracted by customer separation shocks, sellers converge to a stationary size. Therefore, on top of price dispersion, the model generates a well-defined and right-skewed distribution of sellers. Furthermore, the equilibrium is constrained-efficient. This allows us to interpret the model as a theory of efficient endogenous markup dispersion, in which sellers’ pricing decisions lead to a socially optimal allocation of customers across product markets.

In Section 3 we estimate the model in order to quantify the cross-sectional and lifecycle implications of customer capital accumulation. The model is computationally tractable because the equilibrium is block-recursive, a common feature of directed search models with aggregate shocks (e.g. Shi (2009), Menzio and Shi (2010, 2011)). This property implies that, in order to evaluate payoffs, agents need not keep track of the distribution of buyers and sellers over time.

The model is estimated to the U.S. retail sector (2001-2006) using a combination of establishment-level cross-sectional moments from the U.S. Census’ Business Dynamics Statistics (BDS), and information on sales and prices from the Information Resources Inc. (IRI) product-level data, which is aggregated to the establishment level. Using a revenue-weighted price index at the store level, we find empirical support for the main predictions of the model: (i) store size increases are associated with persistently higher price levels, and (ii) stores with lower prices experience persistently higher rates of sales growth. Further, the estimated model’s implications regarding cross-sectional and lifecycle dynamics are in line with the data, including right-skewness in the size and age distributions, cross-sectional dispersion in establishment growth rates, declining growth rates across establishment sizes, and declining exit hazard rates with establishment age.

Using the estimated model, Section 4 explores the aggregate implications of pricing under customer capital accumulation by analyzing the behavior of the economy in response to aggregate demand shocks. In this exercise, there are both level and distributional effects. First, similar to a productivity shock, a negative shock to aggregate demand lowers aggregate output because it depresses buyer search in equilibrium. In the wake of the shock, sellers choose to trade-off immediate losses to future market shares by inter-temporally shifting the burden of the shock between current and future buyers via an increasing path of continuation values for the transition. Since the price level falls with no changes in marginal costs, the average markup is procyclical. Empirically, the cyclicity of markups remains a topic of contention in macroeconomics, where results depend on the economic environment and the nature of the shock. Our analysis provides a rationale for studies

---

<sup>2</sup> Therefore, in our setting, the Diamond (1971) paradox does not hold: as sellers expect to keep serving the same customers in the future, they do not extract all the instantaneous surplus from them.

finding procyclical responses to aggregate demand shocks, such as [Nekarda and Ramey \(2013\)](#) and [Stroebel and Vavra \(2019\)](#), in contrast with the predictions of New Keynesian models.

A key contribution of our paper is to show that there also exist important distributional effects in markups during the transition. Through a decrease in the continuation promise of sellers, negative demand shocks lead to a decrease in the number of new matches, and sellers temporarily shrink in size. The contribution of smaller sellers to the average markup response is relatively stronger on impact and more persistent in the transition. This is because (i) on the extensive margin, there is a left-ward shift in the size distribution in response to the shock, and (ii) on the intensive margin, a small seller’s pricing policy is relatively more sensitive to size changes, for these sellers have a stronger incentive to grow in order to lower the risk of exiting. Finally, to unpack the channels behind the markup procyclicality, we illustrate the inter-temporal smoothing motive that disciplines the behavior of markups by showing that responses are weaker on impact and more persistent in the transition when seller-customer relationships last longer on average. Moreover, through a decomposition of the price level, we identify that an important driver of procyclicality in the average markup is the seller’s growth motive, i.e. procyclical adjustments in the price which allow the seller to fuel growth during the transition. These results show that, in response to adverse shocks, sellers use customers as a form of capital in order to ensure their continued growth.

**Related Literature** This paper is related to several strands of literature. First and foremost, it contributes to a growing literature introducing a role for firm intangibles into models of firm and industry dynamics.<sup>3</sup> A seminal contribution in this literature is by [Gourio and Rudanko \(2014b\)](#). Building on their work, our paper analyzes the dynamic implications of pricing for firm growth. [Rudanko \(2017\)](#) develops a similar idea, showing that when a firm cannot price-discriminate between old and new customers, its pricing may have an impact on firm growth, with smaller firms setting lower prices and growing faster. In our paper, in contrast, the existence of dynamics does not hinge on the discriminatory character of spot prices, because sellers write pricing contracts which are inter-temporal in nature.<sup>4</sup> Moreover, relative to both [Gourio and Rudanko \(2014b\)](#) and [Rudanko \(2017\)](#), our model (i) features competition for customers across sellers of different productivities, giving rise to a size distribution and to price dispersion; and (ii) incorporates aggregate shocks, aimed at exploring the dynamics of both the level and the cross-sectional distribution of prices and markups in the presence of aggregate fluctuations.<sup>5</sup> Both of these ingredients are important to capture features of the data. For example, [Stroebel and Vavra \(2019\)](#) find evidence for

---

<sup>3</sup> Intangibles are a substantial share of firms’ expenditures, with as much as 7.7% of U.S. GDP devoted to marketing ([Arkolakis \(2010\)](#)). Their effects have been widely studied in macroeconomics, including labor wedges ([Gourio and Rudanko \(2014a\)](#)), aggregate productivity ([McGrattan and Prescott \(2014\)](#), [McGrattan \(2017\)](#)), R&D and economic growth ([Cavenaile and Roldan-Blanco \(2019\)](#)), household behavior ([Hall \(2008\)](#)), and the lifecycle of industries ([Atkeson and Kehoe \(2005\)](#), [Perla \(2019\)](#)).

<sup>4</sup> As we shall see, the baseline model assumes no price discrimination because this allows us to uniquely pin down prices from seller size. However, Online Appendix B describes an extension of the model that allows for price discrimination, and shows that this assumption is innocuous for the existence of firm dynamics.

<sup>5</sup> See also [Kaas and Kimasa \(2018\)](#), who combine frictional product and labor markets to study the joint dynamics of prices, output, and employment.

procyclicality of markups in the retail sector in response to aggregate demand shocks, and argue that this might be due to shopper-supplier interactions, supporting our quantitative findings.<sup>6</sup>

Another related paper is [Dinlersoz and Yorukoglu \(2012\)](#), where customer acquisition is also a driver of firm and industry dynamics. In their paper, dynamics emerge from the slow dissemination of information about the firm’s fundamentals, while in our paper there is perfect information and instead dynamics emerge from forward-looking demand. [Paciello \*et al.\* \(2019\)](#) explore how idiosyncratic productivity affects the pricing decisions of firms with a customer accumulation incentive, whereas a key contribution of our paper is to understand heterogeneities among sellers of the same productivity, as motivated by the empirical evidence documented by [Foster \*et al.\* \(2008\)](#) and others.

Our model of pricing under customer capital relies on a large amount of survey evidence suggesting that customer dynamics are an important concern for firms’ pricing decisions in reality. For example, [Blinder \*et al.\* \(1998\)](#) show that the vast majority of firms report having implicit contracts with their customers.<sup>7</sup> These observations have given rise to a literature incorporating customer capital into macroeconomic models of firm pricing, to which the present paper also contributes. In early studies (e.g. [Phelps and Winter \(1970\)](#), [Bils \(1989\)](#), [Rotemberg and Woodford \(1991, 1999\)](#)), firms faced an exogenous law of motion for their customer base. Subsequent research has provided possible reasons for the emergence of such seller-customers relations. The two leading explanations are good-specific habits (as in [Ravn \*et al.\* \(2006\)](#), [Nakamura and Steinsson \(2011\)](#), and [Gilchrist \*et al.\* \(2017\)](#)) and switching costs (as in [Klemperer \(1987, 1995\)](#) and [Kleshchelski and Vincent \(2009\)](#)). Our interpretation is closer in spirit to the latter, but there are substantial differences in the modeling techniques.<sup>8</sup>

This paper is also related to the search literature on price dispersion, pioneered by [Varian \(1980\)](#) and [Burdett and Judd \(1983\)](#). [Menzio and Trachter \(2018\)](#) and [Kaplan \*et al.\* \(2019\)](#) show that dispersion can emerge from buyer heterogeneity in situations in which sellers can price-discriminate. In our model, in contrast, buyers are identical and it is ex-post differences between sellers which give rise to different price levels. While a similar argument is made in [Burdett and Coles \(1997\)](#) and [Menzio \(2007\)](#), these papers do not discuss the implications of customer capital accumulation for the evolution of the firm size distribution or for aggregate dynamics. [Luttmer \(2006\)](#) and [Fishman and Rob \(2003\)](#) discuss the implications for the firm size distribution, but in neither of those papers is there a meaningful role for prices.

From a methodological standpoint, our paper is related to search-and-matching models with large firms. We embed directed search ([Moen \(1997\)](#)) into a model of firm dynamics in the spirit of [Elsby and Michaels \(2013\)](#) and [Kaas and Kircher \(2015\)](#). The paper combines two technical

---

<sup>6</sup> Relatedly, [Kaplan and Menzio \(2016\)](#) also find procyclical markups in a model with product market frictions, though through a different mechanism from ours.

<sup>7</sup> Similar observations have been made by [Hall \*et al.\* \(1997\)](#), [Cason and Friedman \(2003\)](#), [Renner and Tyran \(2004\)](#), [Fabiani \*et al.\* \(2005\)](#), and [Apel \*et al.\* \(2005\)](#). There is also a large literature in Marketing showing that there exists a large degree of persistence in consumers’ brand preferences (e.g. [Bronnenberg and Dubé \(2017\)](#)).

<sup>8</sup> In [Kleshchelski and Vincent \(2009\)](#), the assignment of new customers to sellers is random, and the probability is proportional to the size of the customer base. In our model, by contrast, prices serve to direct buyer search, and the probability of assignment is endogenous to the chosen price. This matters because it gives rise to realistic lifecycle dynamics, one of the main focuses of our paper.

insights from this literature. First, the property of block-recursivity, which allows for a tractable characterization of the firm distribution and its dynamics. Second, the use of dynamic long-term contracts (e.g. [Moscarini and Postel-Vinay \(2013\)](#), [Schaal \(2017\)](#)), which reduce the dimensionality of the state space into an amenable recursive form. The closest papers in terms of theory are [Kaas and Kircher \(2015\)](#) and [Schaal \(2017\)](#). Relative to these, the paper makes two contributions: (i) it shows that a continuous-time setting with Markov shocks yields further analytical tractability, with a single optimality condition (equation (6)) encompassing the relevant trade-offs, and an analytical decomposition of the price level (equation (9)); and (ii) the equilibrium exhibits size-dependent firm growth rates even when firms do not have a decreasing returns to scale technology (due to, for example, convex adjustment costs).

## 2 Model

### 2.1 Environment

Time is continuous, infinite, and indexed by  $t \in \mathbb{R}_+$ . The aggregate state of nature, denoted by  $\varphi \in \Phi \equiv \{\varphi < \dots < \bar{\varphi}\}$ , follows a homogenous continuous-time Markov chain with generator matrix  $\mathbf{\Lambda}_\varphi \equiv [\lambda_\varphi(\varphi'|\varphi)]$ .<sup>9</sup> The economy is populated by a measure-one continuum of risk-neutral, infinitely-lived, ex-ante identical buyers, and a continuum of risk-neutral sellers. The total measure of buyers is exogenous and normalized to one, whereas the total measure of sellers is determined endogenously. All agents share the same discount rate,  $r > 0$ .

There is a single homogenous, indivisible, and perishable good in the economy. Buyers and sellers must participate in a search-and-matching process in order to engage in trade. Possible interpretations are that there exist informational asymmetries regarding product characteristics, as in [Faig and Jerez \(2005\)](#) or [Gourio and Rudanko \(2014b\)](#), or that sellers face inventory or capacity constraints, as in [Burdett et al. \(2001\)](#). All buyers value the consumption of the good by the same flow utility,  $v(\varphi) > 0$ . The dependence of this parameter with the aggregate state  $\varphi$  is the source of aggregate demand fluctuations in the model.<sup>10</sup>

At any instant in time, a buyer is said to be active if it is matched with a seller and is consuming the good, and inactive if it is unmatched and searching for a seller at a flow cost,  $c$ . Active buyers separate from their supplier at an exogenous rate  $\delta_c > 0$ . Sellers belong to one of two groups: incumbents or potential entrants. At any given time, an incumbent seller has a customer base of  $n \in \mathbb{N} \equiv \{1, 2, 3, \dots\}$  customers, which is subsequently called the size of the seller. Besides size, incumbent sellers differ in their idiosyncratic productivity level  $z \in \mathcal{Z} \equiv \{z < \dots < \bar{z}\}$ , which follows a continuous-time Markov chain with generator matrix  $\mathbf{\Lambda}_z \equiv [\lambda_z(z'|z)]$ . An incumbent

---

<sup>9</sup> The transition rates satisfy standard conditions (e.g. [Norris \(1997\)](#)), namely  $\lambda_\varphi(\varphi|\varphi) \leq 0$ ,  $\lambda_\varphi(\varphi'|\varphi) \geq 0$  for all  $\varphi' \neq \varphi$ , and  $\sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) = 0$ . Additionally, we assume  $\sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) < +\infty$ , i.e. the economy always spends a non-zero measure of time in any given state, when such state is visited.

<sup>10</sup> This can be thought of capturing the cyclical nature of shopping behavior (e.g. [Petrosky-Nadeau et al. \(2016\)](#)). Online Appendix [E.1](#) suggests ways to endogenize the  $\delta_c$  separation rate. Online Appendix [E.2](#) shows a micro-foundation for  $v$  based on an environment with CES preferences.

seller's output is constrained by the size of its customer base. Since the good is indivisible, and because there is no benefit in leaving customers unserved, the number of units sold by the seller equals the number of customers in the base. In order to sell, the seller employs  $\ell$  workers at an exogenous wage  $w(z, \varphi) > 0$  using the technology  $n = \ell^\alpha$ ,  $\alpha \leq 1$ . We denote the total cost faced by a seller of size  $n$  by:

$$\mathcal{C}(n; z, \varphi) = w(z, \varphi)n^\psi$$

where  $\psi \equiv \frac{1}{\alpha} \geq 1$ . Incumbent sellers may exit the market (and enter the pool of potential entrants) at any point in time. This may occur for one of two reasons: either because they are hit by an exit shock (at rate  $\delta_f > 0$ ), or because they lose their last remaining customer. Potential entrants must incur a fixed entry cost  $\kappa > 0$ .<sup>11</sup> Upon entry, the seller draws an initial productivity level  $z_0 \in \mathcal{Z}$  from some distribution  $\pi_z$ , where  $\pi_z(z) \geq 0$ ,  $\forall z \in \mathcal{Z}$ , and  $\sum_{z \in \mathcal{Z}} \pi_z(z) = 1$ .

At every instant, all sellers publicly announce price contracts in order to attract buyers. Buyers can perfectly observe the posted contract and visit a seller posting it. For a customer-seller match formed at time  $t$ , a price contract is defined as a price sequence  $(p_{t+j} : j \geq 0)$ , conditional on no separation. Contracts are complete and contingent on the full history of states at each buyer tenure  $j$ , i.e.  $p_{t+j} = p(n^{t+j}; z^{t+j}, \varphi^{t+j})$ ,  $\forall j, t$ . On the demand side, there is no commitment, in that matched buyers can transition to inactivity if they so desire (we allow no seller-to-seller transitions). On the supply side, there are two assumptions. First, the seller fully commits to the contract that she posts. As contracts cannot be revised by the seller for the duration of the match, they must comply with the seller's initial promises. Second, there is anonymity among buyers, in that sellers are unable to price-discriminate between new and old customers.<sup>12</sup> As Online Appendix B formally shows, this second assumption is innocuous for the existence of firm dynamics in equilibrium. However, without it prices would be indeterminate.

Given these assumptions, a sufficient statistic for each contract is the life-time value  $x$  that the seller promises to deliver in expectation to each buyer. Let  $\mathcal{X} = [\underline{x}, \bar{x}] \subseteq \mathbb{R}_+$  be the set of feasible values, and assume that all sellers advertising the same  $x$  perfectly compete in all such contracts. Each seller can simultaneously post, and each buyer can simultaneously search for, at most one offer. For each  $\varphi \in \Phi$ , let  $B(x; \varphi)$  be the measure of buyers seeking to be matched under promised utility  $x$ , and  $S(x; \varphi)$  be the measure of sellers posting  $x$ . A sub-market  $x$  is said to be active if:

$$\theta(x; \varphi) \equiv \frac{B(x; \varphi)}{S(x; \varphi)} > 0$$

where  $\theta : \mathcal{X} \times \Phi \rightarrow [0, +\infty)$  denotes market tightness. The  $\theta$  schedule is taken as given by all agents for each  $\varphi$ . Denote the corresponding contact rates for buyers and sellers by  $\mu(\theta(x; \varphi)) \geq 0$

<sup>11</sup> This can be thought of the costs of maintaining idle product technology or, more broadly, as a cost of market penetration (e.g. [Arkolakis \(2010\)](#)).

<sup>12</sup> Commitment aims to capture a reputational concern, so that reneging on previous promises entails unaffordable costs for sellers (see [Nakamura and Steinsson \(2011\)](#) for an extended discussion). The anonymity assumption is meant to capture that, in largely populated markets with implicit relationships, buyer tenure is unknown to the seller.



and  $\eta(\theta(x; \varphi)) \geq 0$ , respectively, where  $\eta(\theta) = \theta\mu(\theta)$ . Then, we assume:<sup>13</sup>

- Assumption 1** (i)  $\eta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and  $\mu : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  are twice continuously differentiable;
- (ii)  $\eta$  is increasing and concave, and  $\mu$  is decreasing and convex;
- (iii) For some decreasing  $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , define  $f \equiv \eta \circ \mu^{-1} \circ h$ . Then, the function  $f(x)(\hat{x} - x)$  is concave for all  $x \in [0, \hat{x}]$  and  $\hat{x} > 0$ ;
- (iv)  $\eta(0) = \lim_{\theta \nearrow +\infty} \mu(\theta) = 0$ , and  $\lim_{\theta \nearrow +\infty} \eta(\theta) = \lim_{\theta \searrow 0} \mu(\theta) = +\infty$ .

The equilibrium concept is a symmetric Markov perfect equilibrium. Within this class of equilibria: (i) equilibrium policies depend solely on the seller's vector of payoff-relevant states  $(n, x; \mathbf{s})$ , where  $\mathbf{s} \equiv (z, \varphi)$ , and hence they are time-varying only insofar as they are state-dependent; and (ii) all sellers within the same product market  $x$  choose to post the same contract, as sellers perfectly compete within each sub-market. Because the dynamic contract is a price trajectory and thus a large and potentially complex object, we reformulate the model recursively. Define a recursive contract for a seller in state  $(n, x; \mathbf{s})$  as:

$$\omega \equiv \{p, \mathbf{x}'(n'; \mathbf{s}')\}$$

The contract is composed of two objects: (i) a price level  $p$ , which is to be charged to each one of the  $n$  incumbent customers of the seller; and (ii) a vector of continuation payoffs  $\mathbf{x}'(n'; \mathbf{s}') \subseteq \mathcal{X}$ , which are promised by the seller to each buyer conditional on every possible state, namely under any new size  $n' \in \{n-1, n+1\}$  (downsize or upsize) or exogenous state  $\mathbf{s}' \in \{(z', \varphi), (z, \varphi')\}$ .

## 2.2 Buyer and Seller Problems

Let  $U^B(\varphi)$  be the expected value of an inactive buyer in aggregate state  $\varphi \in \Phi$ . Inactive buyers search in the most profitable product markets, so  $U^B(\varphi) = \max_{x(\varphi) \in \mathcal{X}} u^B(x(\varphi); \varphi)$ , where:

$$ru^B(x; \varphi) = -c + \mu(\theta(x; \varphi)) \left( x - u^B(x; \varphi) \right) + \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi' | \varphi) \left( u^B(x; \varphi') - u^B(x; \varphi) \right)$$

The value of entering market  $x$  incorporates the search cost  $c > 0$  (first term), the option value of matching (second term), and the expected change in value due to an aggregate shock (third term).<sup>14</sup> Since inactive buyers choose the best market to search in, all active markets must be

<sup>13</sup> Restrictions (i) and (ii) guarantee that the problems of the buyer and the seller are well-defined, assumption (iii) guarantees that the price-posting problem of the seller has a unique interior solution, and (iv) is a transversality condition on the meeting rates.

<sup>14</sup> In all the equations to follow, notation is economized in two ways. First, since the value of inactivity is itself an equilibrium object,  $\theta(x; \varphi)$  is short-hand for  $\theta(x; \varphi, U^B)$ . Second, since the tightness schedule is taken as given,  $u^B(x; \varphi)$  is short for  $u^B(x; \varphi, \theta)$ , where  $\theta : \mathcal{X} \times \Phi \rightarrow \mathbb{R}_+$  is the equilibrium market tightness function.



equally attractive ex-ante. That is, for all  $(x, \varphi) \in \mathcal{X} \times \Phi$ , we have  $u^B(x; \varphi) \leq U^B(\varphi)$ , with equality if, and only if,  $\theta(x; \varphi) > 0$ . Therefore, in equilibrium:

$$\mu(\theta(x; \varphi)) \left( x - U^B(\varphi) \right) = \Gamma^B(\varphi) \quad (1)$$

for any active market  $x$ , where  $\Gamma^B(\varphi)$  denotes the opportunity cost of matching:

$$\Gamma^B(\varphi) \equiv c + rU^B(\varphi) + \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi' | \varphi) \left( U^B(\varphi) - U^B(\varphi') \right)$$

Equation (1) defines  $\theta(x; \varphi)$  as an increasing function of  $x$ : more ex-post profitable offers for buyers attract a larger measure of buyers per seller, while sellers offering less favorable contracts can expect to find a match sooner. In equilibrium, sellers design contracts for which a low meeting rate for buyers is compensated with higher expected promised values.

Consider now a buyer who is currently consuming from a seller of size  $n$  and productivity  $z$ , under contract  $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$ . The value for the buyer is given by the HJB equation:

$$\begin{aligned} rV^B(n, \omega; \mathbf{s}) = & v(\varphi) - p + (\delta_f + \delta_c) \left( U^B(\varphi) - V^B(n, \omega; \mathbf{s}) \right) \\ & + (n-1)\delta_c \left( x'(n-1; \mathbf{s}) - V^B(n, \omega; \mathbf{s}) \right) \\ & + \eta \left( \theta(x'(n+1; \mathbf{s}); \varphi) \right) \left( x'(n+1; \mathbf{s}) - V^B(n, \omega; \mathbf{s}) \right) \\ & + \sum_{z' \in \mathcal{Z}} \lambda_z(z' | z) \left( x'(n; z', \varphi) - V^B(n, \omega; \mathbf{s}) \right) \\ & + \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi' | \varphi) \left( x'(n; z, \varphi') - V^B(n, \omega; \mathbf{s}) \right) \end{aligned} \quad (2)$$

The right side of equation (2) has the following terms. In the first line,  $v - p$  is the net flow surplus at the agreed-upon price, and the second additive term adjusts for the event of match separation, due to either seller ( $\delta_f$ ) or buyer ( $\delta_c$ ) exit. The second line adjusts for the event that a customer (other than the one in question) separates. The third line is the expected change in value due to the seller successfully attracting a new customer, where the likelihood of such an event is endogenous to the choice of the contract by the seller. Finally, the last two lines reflect the change in value due to an exogenous shock, whether idiosyncratic or aggregate. Note that the customer is forward-looking as she internalizes size and productivity shocks through changes in the market's tightness that are due to the seller's choice of the contract.

The incumbent seller chooses the contract  $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$  that solves:

$$\begin{aligned} rV^S(n, x; \mathbf{s}) = \max_{\omega} \Bigg\{ & pn - \mathcal{C}(n; \mathbf{s}) + \delta_f \left( V_0^S(\varphi) - V^S(n, x; \mathbf{s}) \right) \\ & + n\delta_c \left( V^S(n-1, x'(n-1; \mathbf{s}); \mathbf{s}) - V^S(n, x; \mathbf{s}) \right) \end{aligned} \quad (3a)$$

$$\begin{aligned}
& + \eta \left( \theta(x'(n+1; \mathbf{s}); \varphi) \right) \left( V^S(n+1, x'(n+1; \mathbf{s}); \mathbf{s}) - V^S(n, x; \mathbf{s}) \right) \\
& + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left( V^S(n, x'(n; z', \varphi); z', \varphi) - V^S(n, x; \mathbf{s}) \right) \\
& + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left( V^S(n, x'(n; z, \varphi'); z, \varphi') - V^S(n, x; \mathbf{s}) \right) \Big\}
\end{aligned}$$

where  $V_0^S(\varphi)$  denotes the value of having no customers (derived below). The seller maximizes flow profits  $pn - \mathcal{C}(n; \mathbf{s})$ , taking into account expected changes in life-time value, including the events of exit (first line), customer separation (second line), customer attraction (third line), and exogenous shocks (last two lines). Importantly, the promised value  $x$  enters as a state in the seller's problem, because the following constraint must be satisfied:

$$V^B(n, \boldsymbol{\omega}; \mathbf{s}) \geq x \quad (3b)$$

Equation (3b) is a promise-keeping constraint: the value that each buyer of the seller obtains from contract  $\boldsymbol{\omega}$  must be weakly greater than the utility  $x$  that was promised to her. In equilibrium, this prevents voluntary customer separations, and renders long-lasting relationships.

Finally, inactive sellers have no customers (i.e.  $n = 0$ ) and, unlike incumbents, they must incur a flow set-up cost  $\kappa > 0$  in order to post a contract. Formally, their problem is:

$$rV_0^S(\varphi) = -\kappa + \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) v_0^S(z_0, \varphi) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left( V_0^S(\varphi') - V_0^S(\varphi) \right) \quad (4)$$

The entrant's ex-ante value  $V_0$  is composed of (i) the set-up flow cost  $\kappa$ ; (ii) the expected value of posting a contract, where:

$$v_0^S(z_0, \varphi) \equiv \max_{x_0 \in \mathcal{X}} \eta(\theta(x_0; \varphi)) \left( V^S(1, x_0; z_0, \varphi) - V_0^S(\varphi) \right)$$

is the expected value upon productivity draw  $z_0$ ; and (iii) changes due to an aggregate shock. As the total measure of sellers adjusts freely, an equilibrium with positive entry in all aggregate states requires that  $V_0^S(\varphi) = 0, \forall \varphi \in \Phi$ . The free-entry condition thus pins down the average market tightness among single-customer sellers in the cross-section of initial productivity draws.

## 2.3 Optimal Contract

To find the contract that solves problem (3a)-(3b), note that, because of monotonicity in preferences, a seller will always choose to offer the lowest possible values to their buyers so that her promises are still honored. Thus, the promise-keeping constraint (3b) must hold with equality in equilibrium. To economize on notation, therefore, for the remainder of the paper the value of a buyer is written simply as  $x$ , instead of  $V^B(n, \boldsymbol{\omega}; \mathbf{s})$ .

Next, define the joint surplus in state  $(n, x; \mathbf{s})$  as the sum of the seller's expected value from the match,  $V^S(n, x; \mathbf{s})$ , and the aggregate expected value for all the  $n$  customers of the seller, that

is  $W(n, x; \mathbf{s}) \equiv V^S(n, x; \mathbf{s}) + nx$ . Online Appendix A.1 shows that the joint surplus solves:

$$\begin{aligned}
(r + \delta_f)W(n, x; \mathbf{s}) = \max_{\mathbf{x}'(n'; \mathbf{s}')} & \left\{ n \left( v(\varphi) + (\delta_f + \delta_c)U^B(\varphi) \right) - \left( \mathcal{C}(n; \mathbf{s}) + \eta \left( \theta(x'(n+1; \mathbf{s}); \varphi) \right) x'(n+1; \mathbf{s}) \right) \right. \\
& + \eta \left( \theta(x'(n+1; \mathbf{s}); \varphi) \right) \left( W(n+1, x'(n+1; \mathbf{s}); \mathbf{s}) - W(n, x; \mathbf{s}) \right) \\
& + n\delta_c \left( W(n-1, x'(n-1; \mathbf{s}); \mathbf{s}) - W(n, x; \mathbf{s}) \right) \\
& + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left( W(n, x(z', \varphi); z', \varphi) - W(n, x; \mathbf{s}) \right) \\
& \left. + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left( W(n, x(z, \varphi'); z, \varphi') - W(n, x; \mathbf{s}) \right) \right\} \quad (5)
\end{aligned}$$

Then, our main result is as follows:

**Proposition 1 (Joint Surplus Problem)** *The seller and joint surplus problems are equivalent:*

1. For any  $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$  that solves (3a)-(3b),  $\mathbf{x}'(n'; \mathbf{s}')$  is a solution to (5).
2. Conversely, for any vector  $\mathbf{x}'(n'; \mathbf{s}')$  that solves (5), there exists a unique  $p$  for which  $\{p, \mathbf{x}'(n'; \mathbf{s}')\}$  is a solution to (3a)-(3b).

For the proof, see Online Appendix A.1. Proposition 1 establishes that the contract that maximizes the seller's profits can be found by solving an alternative problem, given by (5). In this problem, the contract maximizes the buyers' total surplus (from consumption and separation), net of the seller's total costs (from serving customers and paying utility to the new customer, in expectation), conditional on the seller extracting rents from buyers up to the limit established by promise-keeping. Since the contract space is complete and all agents are risk-neutral, there always exists a menu of price and promised utilities that, for any future state, redistributes rents among the seller and its customers in a surplus-maximizing fashion. Moreover, as seen in Online Appendix A.1,  $W(n, x; \mathbf{s})$  is constant in  $x$  and  $p$ , and henceforth we may write  $W_n(\mathbf{s})$  instead. Indeed, since price and continuation promises map one-for-one, the maximized surplus is invariant to the rent-sharing components of the contract. Conveniently, this means that the optimal contract can be solved in two stages. The first stage solves for the continuation promises for upsizing, downsizing, and exogenous shocks, that maximize the size of the joint surplus. In the second stage, we find the price level that implements these promises such that promise-keeping binds in every state of nature.

**Stage 1: Continuation promises.** Let us begin with the upsizing choice,  $x'(n+1; \mathbf{s})$ . Taking the first-order condition of problem (5):<sup>15</sup>

$$\underbrace{\frac{\partial \eta(\theta(x'; \varphi))}{\partial x'}}_{\text{Change in attraction rate}} \cdot \underbrace{\left( W_{n+1}(\mathbf{s}) - W_n(\mathbf{s}) \right)}_{\text{Gain in joint surplus from additional customer}} = \underbrace{\frac{\partial \eta(\theta(x'; \varphi))}{\partial x'} \cdot x'}_{\text{Expected cost of paying } x' \text{ utils}} + \underbrace{\eta(\theta(x'; \varphi))}_{\text{Price adjustment by promise-keeping}} \quad (6)$$

<sup>15</sup> Sufficiency follows from Assumption 1 if, in the language of that assumption,  $h : x \mapsto \frac{\Gamma^B}{x - U^B}$  and  $\hat{x} \equiv W_{n+1} - W_n$ .

where  $\theta$  comes from equation (1):

$$\theta(x'; \varphi) = \mu^{-1} \left( \frac{\Gamma^B(\varphi)}{x' - U^B(\varphi)} \right) \quad (7)$$

Condition (6) states that the optimal upsizing choice  $x'$  must equate the expected marginal benefit in joint surplus of attracting an additional customer, to the expected marginal cost.<sup>16</sup>

Unlike the upsizing decision, the continuation values for downsizing,  $x'(n-1; \mathbf{s})$ , and for exogenous shocks,  $\{\mathbf{x}'(n; \mathbf{s}') : \mathbf{s}' \neq \mathbf{s}\}$ , cannot be pinned down via similar first-order conditions. This is because the joint surplus is invariant to these values, as they only affect the way in which the total surplus is shared ex-post. In particular, these objects must be consistent with the sorting behavior of inactive buyers. By symmetry, the downsizing choice of a size- $n$  seller must coincide with the upsizing choice of a seller of size  $(n-2)$ , found via condition (6). A similar symmetry argument holds for  $\mathbf{s}$  transitions, for given  $n$ . This allows us to inductively find the whole sequence of markets  $\{x'(n; \mathbf{s}) : n \geq 2\}$ , for any given  $\mathbf{s}$ . At the boundary  $n = 1$ , the free entry condition must hold. Using equation (4):

$$\kappa = \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \eta \left( \theta(x'(1; z_0, \varphi); \varphi) \right) \left( W_1(z_0, \varphi) - x'(1; z_0, \varphi) \right) \quad (8)$$

For illustration, Figures G.1 and G.2 in Online Appendix G depict, respectively, the equilibrium set of markets and all possible state transitions in equilibrium.

**Stage 2: Prices.** The equilibrium price can be backed out directly from the binding promise-keeping constraint. First, replace  $V^B(n, \boldsymbol{\omega}; z, \varphi) = x'(n; z, \varphi)$  and the components of  $\boldsymbol{\omega}$  in equation (2). Then, solving for  $p$  in the resulting equation yields:

$$\begin{aligned} p_n(\mathbf{s}) = & \underbrace{v(\varphi) - rx'(n; \mathbf{s})}_{\text{Baseline component}} + \underbrace{\delta_f \left( U^B(\varphi) - x'(n; \mathbf{s}) \right)}_{\text{Exit component}} + \underbrace{\eta \left( \theta(x'(n+1; \mathbf{s}); \varphi) \right) \left( x'(n+1; \mathbf{s}) - x'(n; \mathbf{s}) \right)}_{\text{Growth component}} \\ & + \underbrace{n \delta_c \left( \frac{U^B(\varphi) + (n-1)x'(n-1; \mathbf{s})}{n} - x'(n; \mathbf{s}) \right)}_{\text{Separation component}} \\ & + \underbrace{\sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left( x'(n; z', \varphi) - x'(n; \mathbf{s}) \right)}_{\text{Idiosyncratic-shock component}} + \underbrace{\sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left( x'(n; z, \varphi') - x'(n; \mathbf{s}) \right)}_{\text{Aggregate-shock component}} \end{aligned} \quad (9)$$

The optimal price level can be decomposed into additive parts. The first one (the baseline component) is the price that would prevail if customers were to stay matched forever.<sup>17</sup> Each of the remaining additive terms in (9) introduces the necessary adjustments in prices for each of

<sup>16</sup> These costs are given by the cost of promising utility to the new customer (first term on the right-hand side of equation (6)), and lowering the price on all the other customers to keep them from voluntarily separating (second term).

<sup>17</sup> Indeed, in that case we would have  $p = v - rx$ , that is  $x = \int (v - p) e^{-rt} dt$ .

the possible transitions (exit, growth, separation, and idiosyncratic and aggregate shocks) to yield payoffs compatible with the seller's prior promises.

It remains to describe the distribution dynamics of the economy. The model features heterogeneous agents making forward-looking decisions and sorting into markets in the presence of aggregate shocks. Yet, the characterization of the equilibrium has been silent thus far on the composition of agents, or the evolution of this distribution. This is because the equilibrium is block-recursive. In our model, block-recursive arises from the ex-ante revenue-equalization condition across all markets among inactive buyers (equation (1)), which implies that the equilibrium tightness on each market adjusts to be consistent with agents' beliefs.<sup>18</sup> As search is directed, market tightness serves as a sufficient statistic to evaluate payoffs, and agents need not forecast the evolution of aggregates over future states of the economy.

This allow us to describe the model's dynamics via simple flow equations. Let  $S_{n,t}(z) \geq 0$  be the total measure of sellers of size  $n$  and productivity  $z \in \mathcal{Z}$  at time  $t \geq 0$ . All such sellers seek to find new customers in market  $x'(n+1; z, \varphi)$ . Let  $B_{n+1}^I(z, \varphi)$  be the measure of inactive buyers searching within this market. Then, tightness must adjust so that:

$$B_{n+1,t}^I(z, \varphi) = \theta_{n+1}(z, \varphi) S_{n,t}(z) \quad (10)$$

where we henceforth use the short-hand notation  $\theta_n(z, \varphi) \equiv \theta(x'(n; z, \varphi); \varphi)$ . Let  $B_{n,t}^A(z)$  be the measure of customers matched with sellers of type  $(n, z)$  at time  $t$ . By construction, we have:

$$B_{n,t}^A(z) = n S_{n,t}(z) \quad (11)$$

at any  $t \geq 0$ . The unit measure of buyers must be either matched with a seller and consuming, or looking for a seller. Hence,  $\sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} (B_{n,t}^A(z) + B_{n,t}^I(z, \varphi)) = 1, \forall \varphi \in \Phi, t \geq 0$ . Note that market tightness jumps instantaneously in response to an aggregate shock.<sup>19</sup> This is because the mass of inactive buyers adjusts immediately to guarantee that the indifference condition among unmatched buyers (equation (1)) is met in all states of nature. However, by block-recursive, tightness remains constant over time for each aggregate state. The mass of potential entrants, denoted  $S_{0,t}(\varphi)$ , jumps following a  $\varphi$ -shock, and otherwise evolves smoothly due to sellers flowing in and out of inactivity. The distribution of sellers,  $\{S_{n,t}(z)\}$ , is slow-moving and continuous in time. Because of this slow adjustment, the model features sluggish aggregate dynamics.

Formally, the flows into and out of size  $n \geq 2$  for some  $z \in \mathcal{Z}$  are given by:

$$\partial_t S_{n,t}(z) = \eta(\theta_n(z, \varphi)) S_{n-1,t}(z) + (n+1) \delta_c S_{n+1,t}(z) + \sum_{\tilde{z} \neq z} \lambda_z(z|\tilde{z}) S_{n,t}(\tilde{z})$$

<sup>18</sup> [Kaas and Kircher \(2015\)](#) exploit similar insights to obtain tractability. An alternative approach would be to assume free entry of sellers across all product markets, as in [Menzio and Shi \(2010, 2011\)](#) and [Schaal \(2017\)](#).

<sup>19</sup> Our notation convention throughout this section is the following: (i) explicit dependence on  $\varphi$  denotes that a variable jumps with  $\varphi$ ; (ii) a subscript  $t$  denotes that the variable is time-varying for given  $\varphi$ .

$$-\left(\delta_f + n\delta_c + \eta(\theta_{n+1}(z, \varphi)) + \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z)\right) S_{n,t}(z) \quad (12)$$

where  $\partial_t$  denotes a time derivative. Inflows (first line) are given by sellers that acquire their  $n$ -th customer, the share of size- $(n+1)$  incumbents that lose a customer, and those that transition into productivity  $z$  from some  $\tilde{z} \neq z$ . Outflows (second line) are given by sellers that either exit, lose or gain a customer, or transition to a distinct productivity state,  $\tilde{z} \neq z$ . For size  $n = 1$ , the flow equation is identical except that the first term on the first line must be replaced by the flow of entrants into productivity  $z$ , i.e.  $\pi_z(z)\eta(\theta_1(z, \varphi))S_{0,t}(\varphi)$ . Finally, the measure of potential entrants,  $S_{0,t}(\varphi)$ , obeys the following ODE:

$$\partial_t S_{0,t}(\varphi) = \delta_f \mathcal{S}_t + \delta_c \sum_{z \in \mathcal{Z}} S_{1,t}(z) - \sum_{z_0 \in \mathcal{Z}} \pi_{z_0}(z_0) \eta(\theta_{1,t}(z_0, \varphi)) S_{0,t}(\varphi) \quad (13)$$

where  $\mathcal{S}_t \equiv \sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} S_{n,t}(z)$  is the total measure of incumbent sellers. These equations offer a full characterization of the model's distribution dynamics. The stationary distribution of sellers is the solution to  $\partial_t S_{n,t}(z) = 0$ ,  $\forall(n, z)$ . Online Appendix F.1 shows how to find the aggregate measures of agents on the stationary solution.

We are now ready to define an equilibrium, and establish the efficiency result:

**Definition 1** *A Recursive Equilibrium is (i) a set of value functions  $U^B(\varphi)$ ,  $V^B(n, \omega; z, \varphi)$ ,  $V^S(n, x; z, \varphi)$ , and  $W(n, x; z, \varphi)$ ; (ii) a decision rule  $x(\varphi)$  for inactive buyers; (iii) seller promises  $x'(n+1; z, \varphi)$ ,  $x'(n-1; z, \varphi)$ ,  $\{x'(n; z', \varphi) : z' \in \mathcal{Z}\}$ , and  $\{x'(n; z, \varphi') : \varphi' \in \Phi\}$ ; (iv) prices  $p_n(z, \varphi)$ ; and (v) a market tightness function  $\theta(x, \varphi)$ ; such that conditions (1)-(9) are satisfied.*

**Proposition 2 (Efficiency)** *A Recursive Equilibrium is efficient.*

For the proof, see Online Appendix A.2. In our environment, the social planner, constrained by the product market frictions, chooses distributions of buyers and sellers, as well as market tightness levels, in order to maximize consumption gains by active buyers, net of search costs by inactive buyers, and production and entry costs by sellers. Using this definition of welfare, the model features efficient firm dynamics and efficient pricing behavior. Intuitively, price dispersion is necessary in the model to optimally distribute trade gains among buyers and sellers, as this serves to efficiently direct buyer search toward specific product markets.

## 2.4 Exploring the Mechanism

This section examines the model's qualitative implications for cross-sectional heterogeneity and lifecycle patterns. The intuitions will be useful for the quantitative part of the paper, where we will show that a calibrated version of the model is able to replicate salient features of the retail sector in the U.S. in terms of both cross-sectional features and the lifecycle profile of retail establishments.

In the model, sellers grow as a result of their pricing decisions. Formally, the law of motion of seller size is:

$$n_{t+\Delta} - n_t = \begin{cases} 1 & \text{with probability } \eta(\theta_{n_{t+1}})\Delta + o(\Delta) \\ -1 & \text{with probability } n_t\delta_c\Delta + o(\Delta) \\ -n_t & \text{with probability } \delta_f\Delta + o(\Delta) \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

for a small time lapse  $\Delta > 0$ , where  $o(\Delta)$  captures higher-order terms. By equation (7), the customer attraction rate can be written as  $\eta(\theta_{n+1}) = \eta \circ \mu^{-1}\left(\frac{\Gamma^B}{x'_{n+1} - U^B}\right)$ , where  $x'_{n+1}$  is the solution to equation (6). Note that  $\eta(\theta_{n+1})$  is an increasing function of  $x'_{n+1}$  by Assumption 1. Intuitively, by offering better contracts, sellers increase the rate at which they accumulate new customers. Therefore, to understand the mechanics of growth, it becomes important to understand how optimal promises  $x'_{n+1}$  depend on seller size  $n$  in equilibrium. We address to this issue first. Later on, we will look at how optimal promises shape the relation between prices and seller size, as the price level is ultimately the only seller choice variable that is observable in the data.

**From size to continuation values** In equilibrium, the interdependence of size and growth comes from two mechanisms: (i) decreasing returns in technology (i.e.  $\psi > 1$ ); and (ii) frictions in the product market. Considered in isolation, the first mechanism bounds seller growth, as it implies that there exists an optimal scale for sellers. As marginal costs are increasing in size, smaller sellers find it optimal to set high continuation values while they are small to generate fast growth, only to lower their promises as they approach the optimal scale. Similar decreasing returns-to-scale mechanisms have been used in the search-and-matching firm-dynamics literature to generate a notion of firm size (e.g. [Elsby and Michaels \(2013\)](#), [Schaal \(2017\)](#), [Gavazza \*et al.\* \(2018\)](#)). In our quantitative exercise, this mechanism will help us pin down the average size of sellers in the model. However, as this channel is not new to our paper, this section abstracts from it by imposing  $\psi = 1$  for the sake of illustration.

Let us turn to the second growth mechanism, related to the product market frictions. To build intuition, consider a Cobb-Douglas matching function, where  $\mu(\theta) = \theta^{\gamma-1}$  (with  $\gamma < 1$ ), which will be used in the quantitative part of the paper. In this case, one can show:

**Proposition 3 (Solution to the Joint Surplus)** *For each exogenous state  $\mathbf{s} \in \mathcal{Z} \times \Phi$ :*

(a) *The joint surplus  $W_n(\mathbf{s})$  solves the following second-order difference equation in  $n$ :*

$$W_{n+1}(\mathbf{s}) - W_n(\mathbf{s}) - U^B(\varphi) = \left(\frac{\Gamma^B(\varphi)}{\gamma}\right)^\gamma \left(\frac{\Gamma_n^S(\mathbf{s})}{1-\gamma}\right)^{1-\gamma}$$

*where  $\Gamma_n^S(\mathbf{s})$  is a function of  $n$ ,  $W_n$ ,  $W_{n-1}$ , and parameters.*

(b) *Promised utility is given by  $x'_{n+1}(\mathbf{s}) = \gamma(W_{n+1}(\mathbf{s}) - W_n(\mathbf{s})) + (1-\gamma)U^B(\varphi)$ .*



For the proof, see Online Appendix A.3. Proposition 3 shows that, in spite of the rich dynamics of the model, the joint surplus can be expressed in a tractable form. First, one can show that:<sup>20</sup>

$$\Gamma^B = \mu(\theta_{n+1}) \underbrace{(x'_{n+1} - U^B)}_{\text{Ex-post buyer net gains from new match}} \quad \text{and} \quad \Gamma_n^S = \eta(\theta_{n+1}) \underbrace{(W_{n+1} - W_n - x'_{n+1})}_{\text{Ex-post seller net gains from new match}} \quad (15)$$

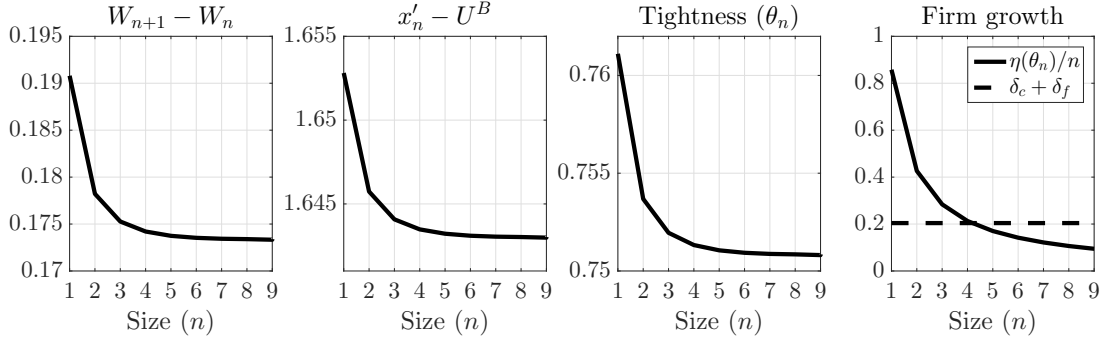
The left-hand side of each equation represents the opportunity cost of search for the buyer and seller, respectively. Equation (15) states that, in equilibrium, these opportunity costs equal the ex-ante net gains from matching for the buyer and the seller, respectively. The buyer's ex-ante net gain is constant because of the buyers' indifference condition (equation (1)). For the seller, ex-ante net gains depend on size: the seller's surplus equals the total gain in joint surplus ( $W_{n+1} - W_n$ )  $> 0$ , net of the value  $x'_{n+1}$  that was promised to the new consumer. With this in mind, part (a) of Proposition 3 says that the equilibrium marginal net gain in joint surplus from each new match (i.e.  $W_{n+1} - W_n - U^B$ ) is a convex combination of the ex-ante net match gains that accrue to the new customer and her seller. Part (b), on the other hand, shows the ex-post surplus sharing rule between the seller and the new customer. The customer's ex-post net gains can be written as:

$$x'_{n+1} - U^B = \gamma(W_{n+1} - W_n - U^B) \quad (16a)$$

showing that a fraction  $\gamma$  of the total gains in joint surplus are absorbed by the new incoming buyer. On the other hand, for the seller:

$$\underbrace{V_{n+1}^S - V_n^S}_{\text{Ex-post seller net gains from new match}} = \underbrace{(1 - \gamma)(W_{n+1} - W_n - U^B)}_{\text{[A] Surplus extracted out of new customer}} + \underbrace{n(x'_n - x'_{n+1})}_{\text{[B] Surplus transferred between seller and pre-existing customers}} \quad (16b)$$

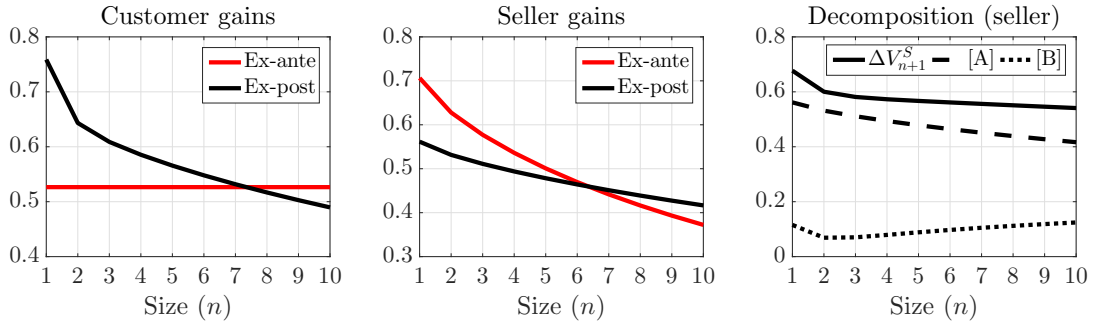
which shows that, when the seller obtains a new buyer, the seller absorbs the remainder share  $(1 - \gamma)$  of the total net gain in joint surplus (term [A]). In addition, there is a transfer between the seller and each of the pre-existing customers, given by term [B].



**Figure 1:** Change in joint surplus, net customer ex-post gains, market tightness, and seller growth, as a function of size, in a numerical example with constant marginal cost (i.e.  $\psi = 1$ ). Seller growth (equation (14)) has been decomposed into the per-customer attraction (solid line) and attrition (dashed line) rates.

<sup>20</sup> For the remainder of this section, we suppress the dependence on  $\mathbf{s}$  everywhere to alleviate notation. This does not fundamentally affect the intuitions to follow.

Figures 1 and 2 offer a graphical representation of the equilibrium. Figure 1 shows the net gain in joint surplus, the customer ex-post gains, market tightness, and the seller's per-customer attraction and attrition rates, as functions of size. The first two panels show that total gains in joint surplus and continuation promises are both decreasing in size. Intuitively, this is due to the combination of search frictions and the entry cost,  $\kappa$ . Every time a seller exits the product market, it must pay the fixed cost  $\kappa$  to re-enter. When a new seller enters into the market, it finds it optimal to set a high continuation promise in order to ensure fast growth initially. By doing so, the seller can raise sufficient resources from future customers quickly to recoup the initial entry cost, and thereby grow out of the risk of exiting. Due to the matching frictions, setting a high promise generates a large inflow of customers (third panel in Figure 1) and high per-customer attraction rates (solid line in the fourth panel) for small sellers. As the seller grows in size, it prefers to start extracting rents from the growing base rather than from newcomers, and therefore lowers continuation promises. As the per-customer attrition rate is constant in size by assumption (dashed line in the fourth panel of Figure 1), there is an optimal size for sellers (given by the intersection of the two lines). To the left (or right) of this point, sellers are growing (or shrinking) in expectation. Therefore, there is cross-sectional dispersion in sizes, and a well-defined size distribution.<sup>21</sup>



**Figure 2:** Ex-ante and ex-post gains for the new buyer and the seller (left and central panels), and decomposition of ex-post seller gains, across seller size  $n$  (right panel), for a numerical example. On the first two panels, the red lines are  $\Gamma_n^B$  and  $\Gamma_n^S$ , and the black lines are  $(x'_{n+1} - U^B)$  and  $(W_{n+1} - W_n - x'_{n+1})$ , respectively. The right-most panel shows the decomposition in equation (16b).

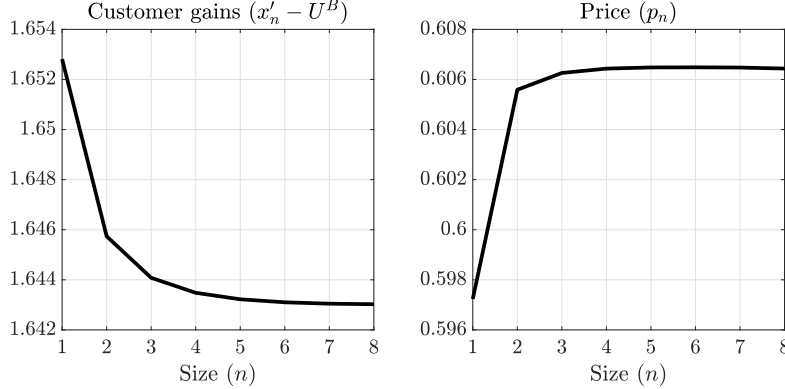
Figure 2 shows how gains from matching are shared between buyers and sellers of different sizes. Recall that promised utilities decline with size in equilibrium. For buyers (left panel of Figure 2), this implies that the ex-post gains are higher when her seller is smaller (see equation (16a)). Sellers (central panel), on the other hand, extract an ever-shrinking surplus from each new customers as they grow larger (term [A] in equation (16b) declines with  $n$ ), and obtain a positive transfer from their pre-existing customers (term [B] is positive). Furthermore, this transfer is increasing in size (right panel of Figure 2): as sellers grow larger, they extract less surplus from new customers (term [A] declines), and more from their existing base (term [B] increases).

**From continuation values to prices** To understand how these results translate to prices, take the partial derivative of the price level (equation (9)) around the equilibrium promise  $x'_n$ :

<sup>21</sup> In the model with exogenous shocks  $(z, \varphi)$ , there is additionally cross-sectional dispersion in productivities (for given size), as sellers with different productivities face different incentives to attract new buyers.

$$\left. \frac{\partial p_n}{\partial x} \right|_{x=x'_n} = -\left(r + \delta_f + \eta(\theta_{n+1}) + n\delta_c\right) < 0$$

There is a negative relationship between the price level and the utility that the seller promises to each of her incumbent buyers. This is a direct consequence of the binding promise-keeping constraint: an increase in prices leads to a reduction in buyer value because the seller extracts surplus up to the buyers' indifference point. Since smaller sellers make more generous promises (recall Figure 1), prices are an increasing function of size in equilibrium (see Figure 3).<sup>22</sup>



**Figure 3:** Net customer ex-post gains, and price level, as a function of size, in a numerical example with constant marginal cost (i.e.  $\psi = 1$ ).

Taking stock, sellers' pricing decisions shape growth and generate ex-post dispersion in prices and sizes. Upon entry, small sellers set low prices in order to recoup the entry cost and generate a high initial rate of growth. As sellers mature, they increase their price as they prefer to extract rents from their expanding base, which slows down growth. In the next section, we estimate the model to study whether this mechanism can generate dynamics similar to those seen in the data.

## 3 Estimation

### 3.1 Mapping the Model to the Data

In this section, we estimate the model to match key features of the U.S. retail sector. Here, sellers are retail establishments (stores), and buyers are shoppers. The retail sector is a good application of our model because buyers are likely to develop lasting relationships with their sellers (as shown in [Paciello \*et al.\* \(2019\)](#)). In addition, retail establishments face a potentially large

<sup>22</sup> There exists a debate in the literature regarding the dynamics of prices on firm tenure. Some studies have found that prices are increasing ([Foster \*et al.\* \(2008, 2016\)](#) and [Piveteau \(2017\)](#)) in firm age, while others have found they are slightly decreasing ([Berman \*et al.\* \(2019\)](#)) or even flat ([Fitzgerald \*et al.\* \(2017\)](#)). More specifically, recent studies in retail (e.g. [Argente \*et al.\* \(2018\)](#) and [Fitzgerald and Priolo \(2018\)](#)) unveil rich dynamics in prices and quantities, including cannibalization effects from newly introduced products and differences in product qualities within the firm. Our model does not feature multi-product firms, and is therefore not equipped to capture these dimensions of within-firm variation. However, our focus throughout is not to explain price dynamics at such levels of granularity, but rather to explore the forces of customer accumulation that may occur at the level of the establishment.

number of customers, so the customer anonymity assumption in the model is likely fitting. In this context, the buyer’s search cost may be interpreted as a proxy for the transport, information, or utility costs associated to finding, or switching away from, trusted suppliers.

In the model, pricing decisions are made by retailers. In the data, these decisions could be made at the retailer or the manufacturer level. Unfortunately, this information is not directly available to us. However, [Stroebel and Vavra \(2019\)](#) find that, in response to demand shocks, the price adjustment is made mostly at the retail rather than the wholesale level. Since our main quantitative exercise in Section 4 will explore the behavior of markups in response to aggregate demand shocks, we believe that the model adheres well to the interpretation of sellers as retailers and not as manufacturers.

We use two sources of data to estimate the model. The first one is the Business Dynamics Statistics (BDS) data from the U.S. Census Bureau, restricted to the retail sector (2001-2006).<sup>23</sup> These data are used to calibrate the cross-sectional features of the model. The BDS reports information at both the firm and the establishment levels. Since the process of customer accumulation and price setting in retail is likely more relevant at the establishment level, in what follows a seller in the model corresponds to a retail establishment in the data.

The second data source allows us to explore the model’s performance in terms of the lifecycle of establishment sizes and prices. We use micro-pricing scanner data from Information Resources, Inc. (IRI) for the period 2001-2006.<sup>24</sup> Our IRI sample includes weekly information on revenues and quantities at the product level for 2,036 retail (drug and grocery) stores distributed across 50 metropolitan statistical areas (MSA) in the United States. Products are defined by their barcode, or Universal Product Code (UPC), and are grouped into 31 different categories.<sup>25</sup> The average weekly price of each product is computed by taking the ratio of weekly sales to quantity sold, i.e.  $p_{gst} = \frac{Sales_{gst}}{Q_{gst}}$ , for product  $g$  in store  $s$  and week  $t$ . To aggregate up to the establishment, and to accommodate for the multiproduct nature of grocery retailing ([Smith \(2004\)](#)), we construct a store-week-specific price index given by the average price of a composite bundle of goods that are sold in the store throughout the period, weighted by their share of overall sales:

$$p_{st} = \sum_{g \in G_s} \omega_{gs} p_{gst}, \quad \omega_{gs} = \frac{\sum_t Sales_{gst}}{\sum_{g \in G_s} \sum_t Sales_{gst}} \quad (17)$$

where  $G_s$  is the set of all goods sold by store  $s$  across all weeks. Note that, since weights are time-invariant, any dynamics in store-level prices come exclusively from changes in the individual prices.<sup>26</sup> In turn, the overall sales of the store is computed by  $Sales_{st} = \sum_{g \in G_s} Sales_{gst}$ .

<sup>23</sup> The BDS data are publicly available at <https://www.census.gov/programs-surveys/bds.html>.

<sup>24</sup> The IRI data are available for request at [www.iriworldwide.com/en-US/solutions/Academic-Data-Set](http://www.iriworldwide.com/en-US/solutions/Academic-Data-Set). For documentation, see [Bronnenberg et al. \(2008\)](#). Recent studies in economics using the IRI data include [Alvarez et al. \(2016\)](#), [Gagnon and López-Salido \(2014\)](#), [Coibion et al. \(2015\)](#), [Paciello et al. \(2019\)](#), and [Stroebel and Vavra \(2019\)](#).

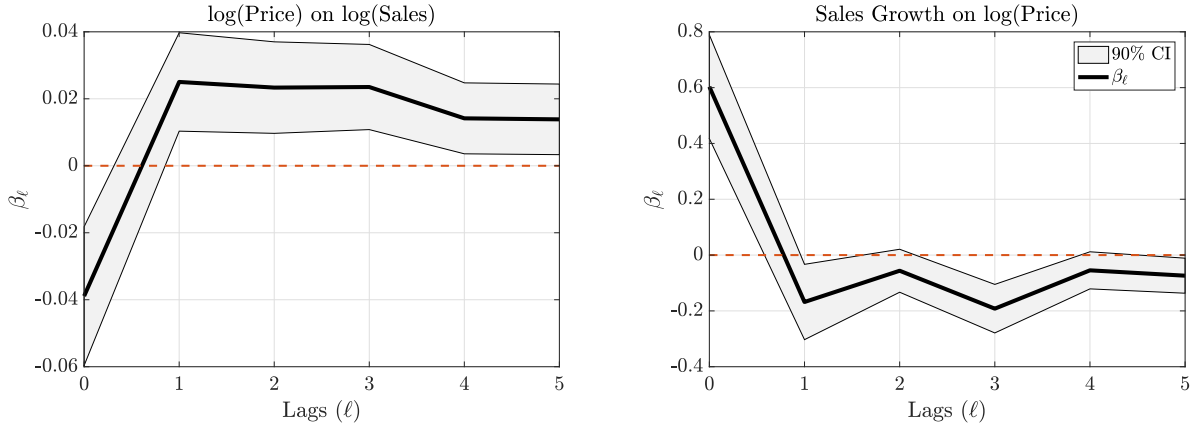
<sup>25</sup> The categories include a variety of food, beverages, personal hygiene, and cleaning products. The data represent about 15% of household spending in the Consumer Expenditure Survey.

<sup>26</sup> As only products that are sold by the store in all weeks are being considered, the price index is not affected by changes in the composition of goods within the store. This reduces our sample of products, from 78,667 to 26,022 distinct UPC codes. The number of stores is not affected by this restriction.

These measures can be used to find support in the data for the two basic predictions of the model. In particular, we check (i) whether smaller size is associated with a lower price index, and (ii) whether lower prices predict faster store growth. The basic regression specification is:

$$y_{st} = \alpha + \sum_{\ell=0}^L \beta_{\ell} x_{s,t-\ell} + \alpha_s + \alpha_t + \varepsilon_{st}$$

for store  $s$  in week  $t$ . In the first regression, the dependent variable is  $\log(p_{st})$ , and the regressors are lags of  $\log(\text{Sales}_{st})$ . Thus,  $\beta_{\ell}$  captures the effect of larger store size on prices  $\ell = 0, \dots, L$  weeks ahead. The second regression is week-on-week sales growth on lags of  $\log(p_{st})$ , so  $\beta_{\ell}$  captures the effect of increasing prices on sales growth  $\ell$  weeks ahead.<sup>27</sup> The regression introduces lags because the mechanism described in the model (i.e. pricing to build up a customer base) implies, by its very nature, that the effects are likely to accumulate over time. As one cannot make comparisons across stores (because the  $G_s$  sets may contain different products for different stores), we exploit within-store variation by adding store fixed effects ( $\alpha_s$ ) on top of week fixed effects ( $\alpha_t$ ).



**Figure 4:** Coefficients  $\{\beta_{\ell}\}_{\ell=0}^L$  for the store-level regression analysis. Left panel:  $\log(p_{st})$  on  $\log(\text{Sales}_{st})$ . Right panel: sales growth on  $\log(p_{st})$ . IRI data (2001-2006). The figures show the point estimates (solid black line), together with 90% confidence bands, on a horizon of  $L = 5$  weeks (i.e. one month). Standard errors are clustered at the MSA level (50 clusters in total).

Figure 4 plots  $\{\beta_{\ell}\}$  up to  $L = 5$  weeks, together with the corresponding confidence bands, for each of the two regressions. Our results are in line with the predictions of the model. The left panel shows that an increase in store size is associated with a significant increase in prices within the store for five consecutive weeks (i.e. a full month). Through the lens of the model, this occurs because the seller sets higher prices as it successfully attracts new customers. Contemporaneously, the correlation is negative ( $\beta_0 < 0$ ), potentially reflecting the intensive margin of demand (higher prices mean lower quantity demanded from incumbent customers), which the model is not designed to capture. The effect in subsequent weeks, though, is positive and significant, suggesting that the extensive margin of demand exhibits persistence, perhaps due to inertia in shopping behavior, or because households learn slowly about new deals. The right panel shows that lower prices predict faster growth within the store in the future, with effects that accumulate for several weeks.

<sup>27</sup> Throughout, the growth rate is defined as  $g_{t+1}^{\text{sales}} = \frac{\text{Sales}_{s,t+1} - \text{Sales}_{st}}{0.5(\text{Sales}_{s,t+1} + \text{Sales}_{st})}$  to correct for extreme values.

Potentially because of a similar intensive-margin channel, the contemporaneous effect is positive. However, subsequently there exist significantly negative effects which accumulate for at least one month. All in all, these findings provide suggestive evidence in support of the predictions of the model: lower prices at the store level are associated with smaller and faster growing stores.

### 3.2 Calibration Strategy

The model is parametrized as follows. First, the shifter in the cost function  $\mathcal{C}(n; z, \varphi) = w(z, \varphi)n^\psi$  is given by  $w(z, \varphi) = we^{z+\varphi}$ , where  $w > 0$  is a parameter. This specification is isomorphic to multiplicative idiosyncratic and aggregate TFP shocks in the production function, which is standard in the literature (e.g. [Kaas and Kircher \(2015\)](#)). The model is solved on a grid  $\mathcal{N} \equiv \{1, 2, \dots, \bar{n}\}$  for seller size, for some large upper bound  $\bar{n} < +\infty$ . One can then show that the model's dynamics are convergent:

**Proposition 4 (Stability)** *The dynamic system given by equations (12)-(13) is stable, and converges to an invariant distribution for each aggregate state  $\varphi \in \Phi$ .*

For the proof, see Online Appendix A.4. Further, the structure of the exogenous shocks,  $(z, \varphi)$ , must be specified. In principle, this requires the estimation of  $k_s(k_s - 1)$  transition rates for each shock  $s \in \{z, \varphi\}$ , a potentially very large number of parameters. To reduce the dimensionality, each shock is assumed to follow an Ornstein-Uhlenbeck process in logs:<sup>28</sup>

$$\begin{aligned} d \log(z_t) &= -\rho_z \log(z_t)dt + \sigma_z dB_t^z \\ d \log(\varphi_t) &= -\rho_\varphi \log(\varphi_t)dt + \sigma_\varphi dB_t^\varphi \end{aligned}$$

where  $(B_t^z, B_t^\varphi)$  are standard Brownian motions. Conveniently, this reduces the estimation of shocks to only two parameters per shock: a persistence  $\rho$ , and a volatility  $\sigma$ . For full details, see Online Appendix C.1. Finally, the entrants' productivity distribution  $\pi_z$  is taken as the ergodic distribution associated with the estimated Markov chain for  $z$ .

The model has 11 free parameters to be identified. Of these, 9 are deep parameters: the value of consumption ( $v$ ), the time discount rate ( $\rho$ ), the exit rate of sellers ( $\delta_f$ ) and separation rate of customers ( $\delta_c$ ), the scale and curvature parameters of the cost function ( $w$  and  $\psi$ ), the entry cost for new sellers ( $\kappa$ ), the matching elasticity ( $\gamma$ ), and the search cost for inactive buyers ( $c$ ). In addition, we must set values for the persistence and dispersion parameters of the exogenous productivity state process:  $(\rho_z, \sigma_z)$ . Sellers are assumed to accumulate at most  $\bar{n} = 50$  customers.<sup>29</sup> As the model is estimated at its steady-state equilibrium, for now aggregate shocks  $\varphi$  are turned off, and re-introduced in Section 4.

<sup>28</sup> An OU process is a specific continuous-time Markov chain, the analogue of an AR(1) process in continuous time.

<sup>29</sup> In the estimated model this is a sufficiently large bound: about 92.5% of sellers have 10 customers or less, and the share of sellers with exactly 50 customers is very small (about 0.005%). Moreover, as seen later, this bound is sufficiently large to match the right tail of the size distribution (see left panel of Figure 5).

**External identification** The parameters  $(v, r, \gamma, \delta_f)$  are calibrated outside the model. The value of consumption is normalized to  $v = 1$ , so that the consumption good serves as the numeraire of the economy. The discount rate is set to  $r = 0.05$ , corresponding to a discount factor of approximately 95% annually. The matching elasticity parameter  $\gamma$  is set based on the opportunity cost of buying versus selling from the American Time Use Survey of the Bureau of Labor Statistics (BLS), used in [Gourio and Rudanko \(2014b\)](#). Following their paper, the time spent by consumers in shopping activities relative to selling activities is 25%. In the model, this relative opportunity cost pins down a matching elasticity of  $\gamma = 0.2$ .<sup>30</sup> Finally, in the model  $\delta_f$  is the rate at which establishments exit the market, regardless of their size or productivity. In practice, one reason this may happen is because of exit of the firm. Following this interpretation, we set  $\delta_f = 0.1016$ , the average firm exit rate in the retail sector over the 2001-2006 period from the BDS.

**Internal identification** The remaining 7 parameters are estimated jointly via Simulated Method of Moments (SMM). This procedure is implemented using an algorithm that draws quasi-random parameter vectors over a 7-dimensional space, and then employs an unweighted minimum-distance criterion that compares empirical moments to their model counterparts. The model's moments are computed from both the stationary solution as well as from simulated data. For the stationary solution, we solve a fixed-point problem that uses value function iteration on  $W$  and a bisection method to solve for the value of inactivity  $U^B$  that satisfies the free-entry condition. Online Appendix C.2 outlines the details of this solution method.<sup>31</sup> To obtain moments from simulated data, we generate histories for 1,000 distinct stores over  $T = 100$  periods. As the BDS data comes at the yearly frequency, a period in the model is a year. Moreover, because the IRI data is at a weekly frequency, each period of simulated data is discretized into steps of size  $\Delta = 1/52$ .<sup>32</sup>

The set of targeted moments can be grouped into two categories: (i) cross-sectional moments from retail sector (from BDS data); and (ii) store-level moments related to the distribution of sales and prices (from IRI data). Due to the model being highly non-linear, identification is challenging. However, we can provide an interpretation for how different parameters are informative about each moment. Online Appendix D.1 provides a more rigorous identification test for each parameter-moment pair discussed in the text.

At the aggregate level, the targets are the entry rate, the average size of establishments, and the average markup. First, we target an entry rate of 11.62%, which is computed from the BDS data as the average over the 2001-2006 period of the ratio of establishments aged one year or less to the total number of existing establishments in the retail sector. The entry rate in the model is

---

<sup>30</sup> Note from equation (A.3.4) in Online Appendix A.3 that the aggregate relative opportunity cost of time for sellers versus buyers is constant across  $(n, z)$ , and given by  $\frac{S_n(z)\Gamma_n^S(z)}{B_{n+1}^I(z)\Gamma_n^B} = \frac{1}{\gamma} - 1$ .

<sup>31</sup> Further, Online Appendix F.2 shows that, for a given  $U^B$ , the value function exists and is unique.

<sup>32</sup> The simulation moments are computed using only the second half of the time sample, to allow for convergence. All sellers are drawn from the stationary distribution at time  $t = 0$  and evolve endogenously through simulated Markov chains that replicate equations (12)-(13).



defined as the ratio of entering sellers to the total measure of incumbents. By equation (13):

$$EntryRate = \frac{S_0}{S} \sum_{z_0 \in \mathcal{Z}} \pi_{z_0}(z_0) \eta(\theta_1(z_0)) \quad (18)$$

Matching the entry rate is relevant because an important share of aggregate dynamics in the model is driven by small sellers, as shown in the next section. A relevant parameter to identify the entry rate is  $\delta_c$ , the rate of customer turnover. This is because, in the stationary solution, the establishment entry and exit rates coincide, with the latter given by  $ExitRate = \delta_f + \delta_c \frac{\sum_z S_1(z)}{S}$ , on which  $\delta_c$  acts as a shifter. Second, to calibrate store size, we target the average size of retail establishments from the BDS, equal to 18.06 employees on average over 2001-2006. In the model, the mapping between customers and employees is given by  $\ell = n^\psi$ , which is used to compute the average employment in the cross-section of simulated sellers. This moment helps us identify  $\psi$  (the curvature in technology), as this parameter controls for the optimal scale of sellers in the model. Third, as our main quantitative exercise will consider the cyclicalities of markups, it is important to accurately predict the cross-sectional average markup of the economy. Because measuring markups in the data usually requires a stand on market structure and the demand curve faced by firms, in the literature estimates vary substantially depending on the empirical methodology used, the industry of consideration, and the overall characteristics of the sample. Using firm-level data, typical estimates range from about 10% to as much as 50% or more (e.g. [Christopoulou and Vermeulen \(2008\)](#), [DeLoecker and Warzynski \(2012\)](#), [DeLoecker et al. \(2016\)](#)). Following [Faig and Jerez \(2005\)](#), we impute the retail markup from the average ratio of gross margins to sales in the retail sector, which gives us 38.4%.<sup>33</sup> In the model, the average markup is computed as the sales-weighted average of the ratio of price to marginal cost:

$$\bar{m} = \sum_{n \in \mathbb{N}} \sum_{z \in \mathcal{Z}} s_n(z) m_n(z); \quad \text{with } m_n(z) \equiv \frac{p_n(z)}{mc_n(z)} \quad (19)$$

where  $s_n(z) = \frac{np_n(z)}{\sum_{n,z} np_n(z)}$  is the sales share of sellers of type  $(n, z)$ , and  $mc_n(z) \equiv \mathcal{C}(n; z) - \mathcal{C}(n-1; z)$  is the marginal cost. Though many parameters affect the average markup, the cost scale  $w$  is arguably the most relevant one, as it acts as a shifter in the marginal cost.

At the store level, we target moments related to the distribution of prices and sales from the IRI data. First, using indirect inference, we target the  $\beta_1$  coefficient from the regression on the left panel of Figure 4,  $\beta_1 = 0.025$ . This moment helps us identify the seller entry cost  $\kappa$ , as this parameter controls for the dependence between prices and seller size in equilibrium (recall our discussion in Section 2.4). The second target is the degree of price dispersion, computed as the average standard deviation in normalized prices. Unfortunately, our store-level measure of prices (equation (17)) precludes us from making meaningful comparisons across stores, as different

---

<sup>33</sup> This number is obtained from the latest Annual Retail Trade Report of the U.S. Census Bureau (<https://www.census.gov/retail/index.html>). The average gross margin is about 27.75%, implying an average markup of  $.2775/(1 - .2775) \approx .384$ . For comparison, [Hottman \(2017\)](#) estimates a slightly lower average markup in the U.S. retail sector, in the range 29-33%.

goods may be in assortment in different stores. Therefore, we rely on the evidence from [Kaplan and Menzio \(2015\)](#), who compute that the degree of price dispersion that is due to store-related components is equal to 0.057. This moment is informed by the search cost  $c$  because, by the buyer free-entry condition (equation (1)), this parameter controls for buyer participation in different markets, and it is therefore a driver of ex-post differences in prices across sellers. Finally, for the persistence and dispersion parameters  $(\rho_z, \sigma_z)$ , we target the average dispersion in store-level sales week-on-week growth rates (0.138), and the autocorrelation in normalized prices (0.995).

**Estimation Results** Table 1 presents the full set of calibrated parameters, and the result of the calibration exercise in terms of moment-matching. The model matches all moments well. To match the entry rate, the model predicts a buyer separation rate of 17.9% annually. This value falls right in the middle of the range reported by [Gourio and Rudanko \(2014b\)](#) of 10% to 25%, and implies that customer-seller relationships last for 5.58 years on average. To match the average size of retail establishments, the model requires a cost curvature of  $\psi = 1.1205$ , implying that a store of 10 customers faces productions costs which are 1.32 times higher than the combined costs of 10 stores of one customer each. To match the average markup, the cost scale must be  $w = 0.4248$ , corresponding to the marginal cost faced by a size-one store with median productivity.<sup>34</sup> Moreover, the model implies that customers must pay 5.11% of the product’s valuation for the opportunity to search, and that potential entrants face a fixed cost of entry that is 64% higher than the buyer’s instantaneous valuation of the good.

Parameter		Value	Target (Source)	Model	Data
Buyer separation rate	$\delta_c$	0.1790	Establishment entry rate (Census)	11.744	11.620
Cost curvature	$\psi$	1.1205	Average establishment size (Census)	18.148	18.060
Cost scale	$w$	0.4248	Average markup (Census)	38.624	38.408
Buyer search cost	$c$	0.0511	Price dispersion ( <a href="#">Kaplan and Menzio (2015)</a> )	0.055	0.057
Seller entry cost	$\kappa$	1.6368	$\beta_1$ in $\log(p)$ -on- $\log(\text{sales})$ regression (IRI)	0.037	0.025
$z$ -shock volatility	$\sigma_z$	0.2948	Dispersion of sales growth (IRI)	0.121	0.138
$z$ -shock persistence	$\rho_z$	0.0379	Autocorrelation of normalized prices (IRI)	0.983	0.995
<i>Period: 1 year</i>					

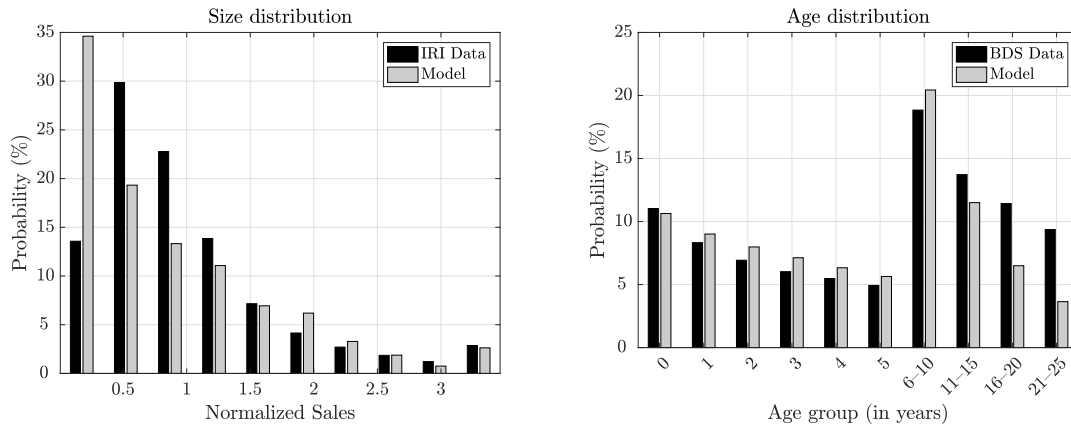
**Table 1:** Set of internally calibrated parameters and model fit. *Notes:*  $(\rho_z, \sigma_z)$  are Euler-Maruyama parameters of the Ornstein-Uhlenbeck process for  $z$  (see Online Appendix C.1 for details).

### 3.3 Cross-Sectional and Lifecycle Implications

According to a recent literature, surveyed in the Introduction, a large degree of cross-sectional and lifecycle differences among businesses in the data stems from idiosyncratic demand-related

<sup>34</sup> One could also consider a sales-weighted harmonic average markup,  $\bar{m}^h \equiv \left( \sum_{n,z} s_n(z) [m_n(z)]^{-1} \right)^{-1}$ , as the arithmetic average is known to overpredict the aggregate markup (e.g. [Baqaee and Farhi \(2019\)](#)). In the calibrated model, the difference between the two is small:  $\bar{m} = 38.62\%$  versus  $\bar{m}^h = 37.43\%$ .

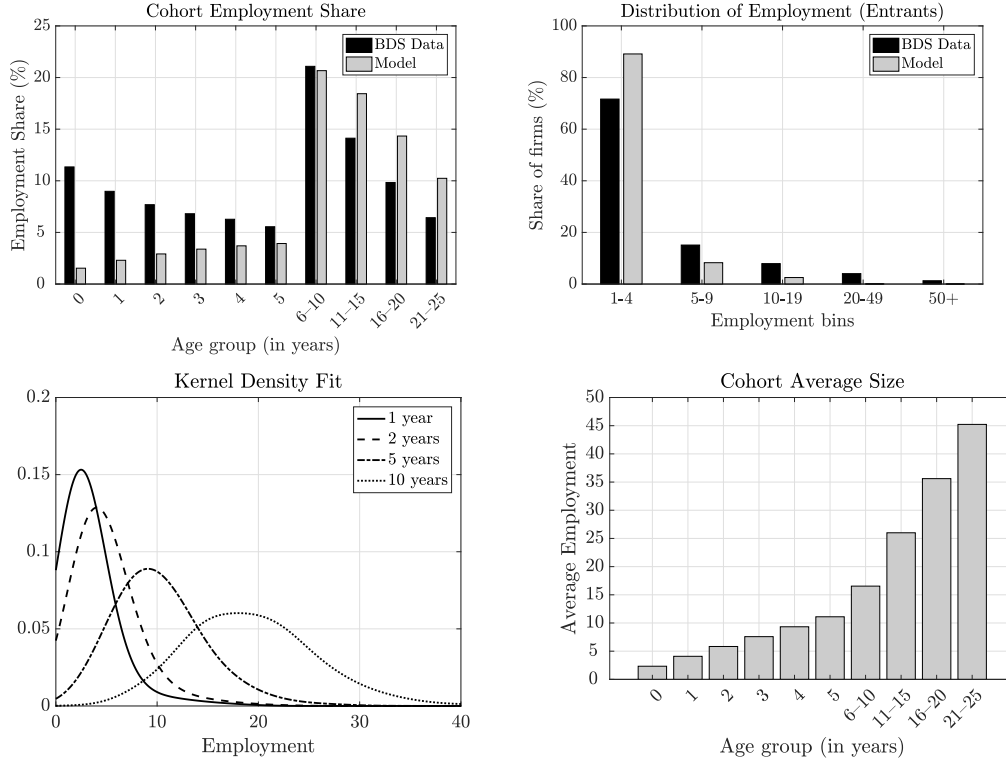
components (see [Hottman \*et al.\* \(2016\)](#) and [Foster \*et al.\* \(2008, 2016\)](#)). In our model, heterogeneity comes from differences in both productivity ( $z$  shocks) and demand (prices). This section discusses the predictions of the estimated model regarding cross-sectional and lifecycle dynamics, and compare them to the data. We document that the stylized firm-dynamics facts commonly found for the economy as a whole (e.g. [Clementi and Palazzo \(2016\)](#)) hold within the retail sector as well: namely, the size and age distributions are right-skewed, smaller establishments exhibit faster growth rates, and exit hazard rates decay with establishment age. As we have no data on the number of customers of establishments, we resort to an observable proxy of size which can be measured in the same way in the model and the data. We use sales for the IRI, and employment for the BDS, both of which are standard measures of size in the literature. Moreover, the IRI does not provide store age information, so we use the BDS for moments related to establishment age.



**Figure 5:** Distributions of establishment size and age: data versus model. *Notes:* Left panel: Store-level sales data from the IRI, where the right-tail of the sales distribution in the data has been winsorized at the 2% level. Right panel: Establishment-level age data from the BDS, where each age group is the time-series average over the 2001-2006 period of the share of establishments in that group.

First, we verify that the estimated model provides correct predictions for the establishment size and age distributions. The left panel of Figure 5 plots the distribution of normalized sales in the model and the IRI data, showing that there is a good fit of the right skewness in this distribution and of the rate of decay in sales. The model, however, puts too much mass on the lower tail of the distribution. The reason for this disagreement between data and model is that, by construction, in the model there is too much incumbent turnover among sellers with very few customers. The right panel shows the establishment age distribution in the model versus the data. Since the IRI does not provide reliable establishment age information, the BDS data is used instead. The fit for this distribution is very good, suggesting that the correlation between size and age implied by the parameters governing seller transitions is close to that in the data.

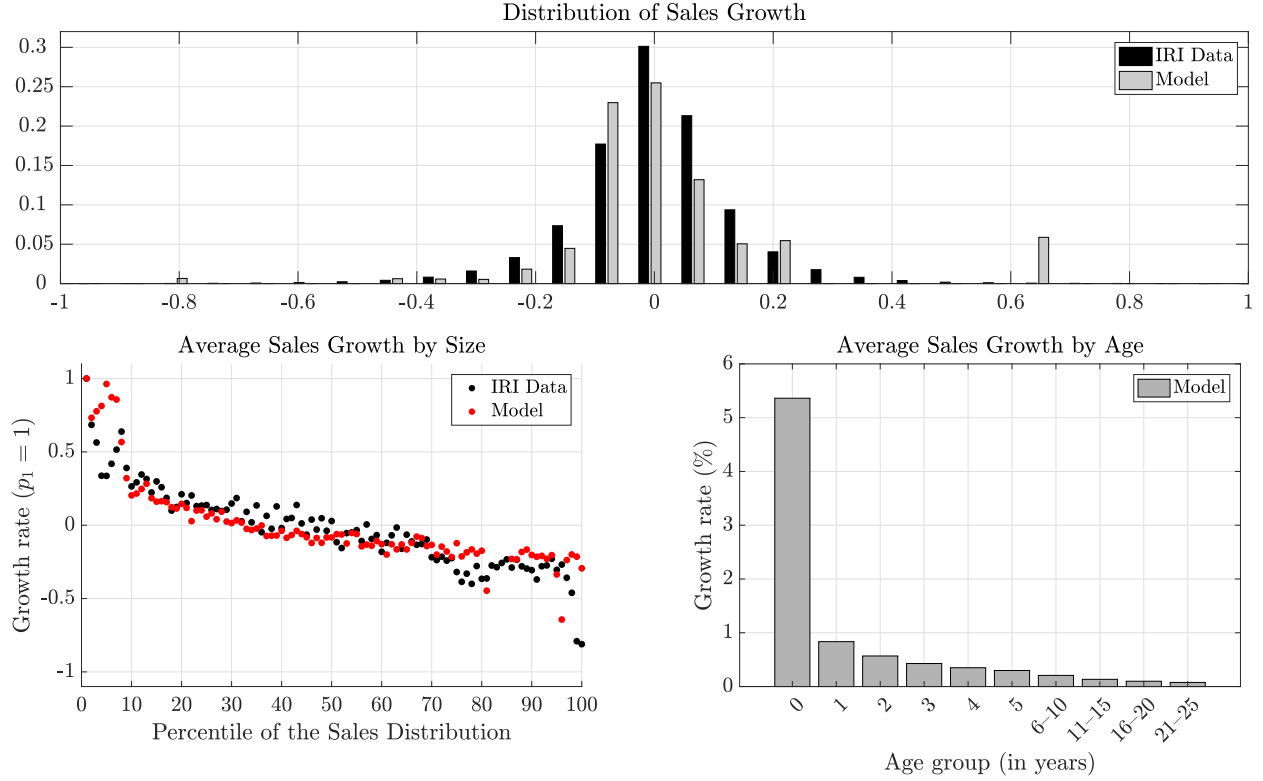
To explore the relationship between size and age in the model and the data more carefully, the upper-left panel in Figure 6 shows the share of total employment by age cohort compared to the BDS data. The model offers a good match for mature establishments. For the younger ones, though the model underpredicts employment shares, it nevertheless captures quite well the distribution of employment. This can be seen in the upper-right panel of Figure 6, which shows that a disproportionate share of entering establishments (aged less than one full year) are small,



**Figure 6:** Cross-sectional establishment moments: data versus model. *Notes:* Top panels: employment shares by establishment age group (left), and distribution of employment among entrants (right), in the model and the BDS data. Bottom panels: Kernel density estimates for the employment distribution by age cohort (left), and the average size of each cohort (right), in the estimated model.

both in the model and in the BDS data. The lower-left panel presents kernel density estimates for the employment distribution of establishments of different age cohorts in the model, showing that the employment distribution shifts and becomes less right-skewed as cohorts age. Similar features have been observed in the data for other sectors, predominantly manufacturing (e.g. [Cabral and Mata \(2003\)](#)). Finally, the lower-right panel shows the evolution of the average employment, in the model, as a function of establishment age, showing that there is a positive correlation between size and age. Both decreasing returns to scale ( $\psi > 1$ ) and the pricing mechanism explained in Section 2.4 slow down growth as sellers age, giving rise to selection effects: a few old sellers become very large, which accounts for the positive skewness in the size distribution.

The next figures explore the model's performance in terms of establishment lifecycle dynamics. The top panel of Figure 7 shows the cross-sectional distribution of non-zero sales growth rates in the IRI data and the model. Both distributions are symmetric about zero, and exhibit long tails. The bottom-left panel depicts average growth rates within different percentiles of the sales distribution in the IRI data and the model (with the first percentile normalized to one), showing that smaller establishments tend to grow faster. The model matches the rate of decline in growth rates with seller size very well. The right panel shows that, in the model, most growth in sales is fueled by young establishments, relative to older cohorts. Though the IRI data does not allow us to check this prediction, as it does not contain store age information, the result is consistent with well-known empirical regularities from various U.S. sectors (e.g. [Evans \(1987\)](#) and [Hall \(1987\)](#)),



**Figure 7:** Establishment lifecycle moments: data versus model. *Notes:* Top panel: Distribution of non-zero week-on-week sales growth (model vs. IRI data). Bottom-left: sales growth by size percentile, with first percentile normalized to one (model vs. IRI data). Bottom-right: average growth by age cohort in the model.

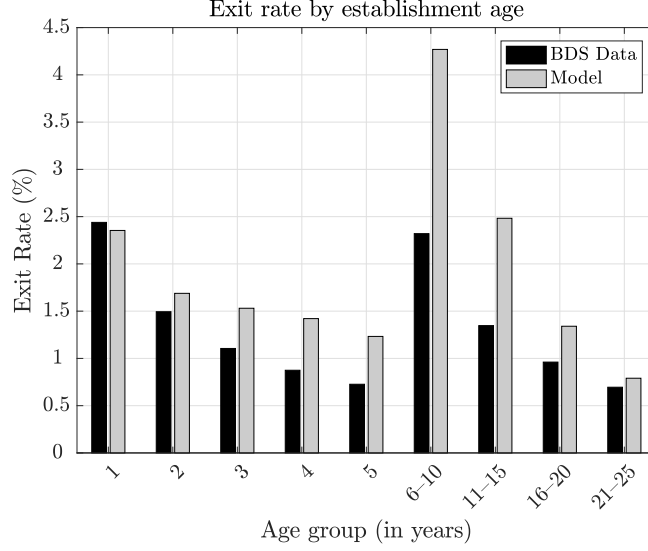
and it also holds in other countries (e.g. [Coad \(2009\)](#)).

Finally, the model provides a good fit for the exit dynamics of different age cohorts. Figure 8 shows that exit hazard rates decline with establishment age in the model. This is again in line with our BDS data, and with empirical observations made for the economy as a whole (e.g. [Dunne \*et al.\* \(1989\)](#) and [Mukoyama and Lee \(2015\)](#)). Canonical models of technology-driven firm dynamics in the spirit of [Hopenhayn \(1992\)](#) deliver this prediction because of a process of selection on productivity: on average, entrants are less productive than incumbents and, as firms mature, less productive incumbents are forced to exit. In our model, the selection channel operates even in the absence of productivity differences: survivors generate high expected growth through low prices, and this lowers the exit hazard of those establishment that succeed in accumulating customers.

In sum, the model provides cross-sectional and lifecycle predictions that are in line with the data. Fitting these dimensions of the data is relevant for our quantitative analysis in the next section, which shows that establishments of different sizes may contribute differently to the response of the prices and markups to perturbations in aggregate demand.

## 4 Aggregate Demand Shocks with Customer Capital

Using the estimated model, this section studies the effects of aggregate demand fluctuations at both the macroeconomic level as well as in the cross-section of sellers. A recent literature has



**Figure 8:** Establishment exit rates by age: model vs. BDS data.

emphasized the relevance that consumer shopping behavior may have for macroeconomic dynamics (e.g. [Bai \*et al.\* \(2017\)](#), [Petrosky-Nadeau and Wasmer \(2015\)](#) and [Michaillat and Saez \(2015\)](#)). In these papers, a shock to demand has an impact on search incentives, which may lead to persistent output effects. This section argues that the underlying size heterogeneity and the forces of entry and exit can provide additional insights into the response of markups to aggregate demand shocks.

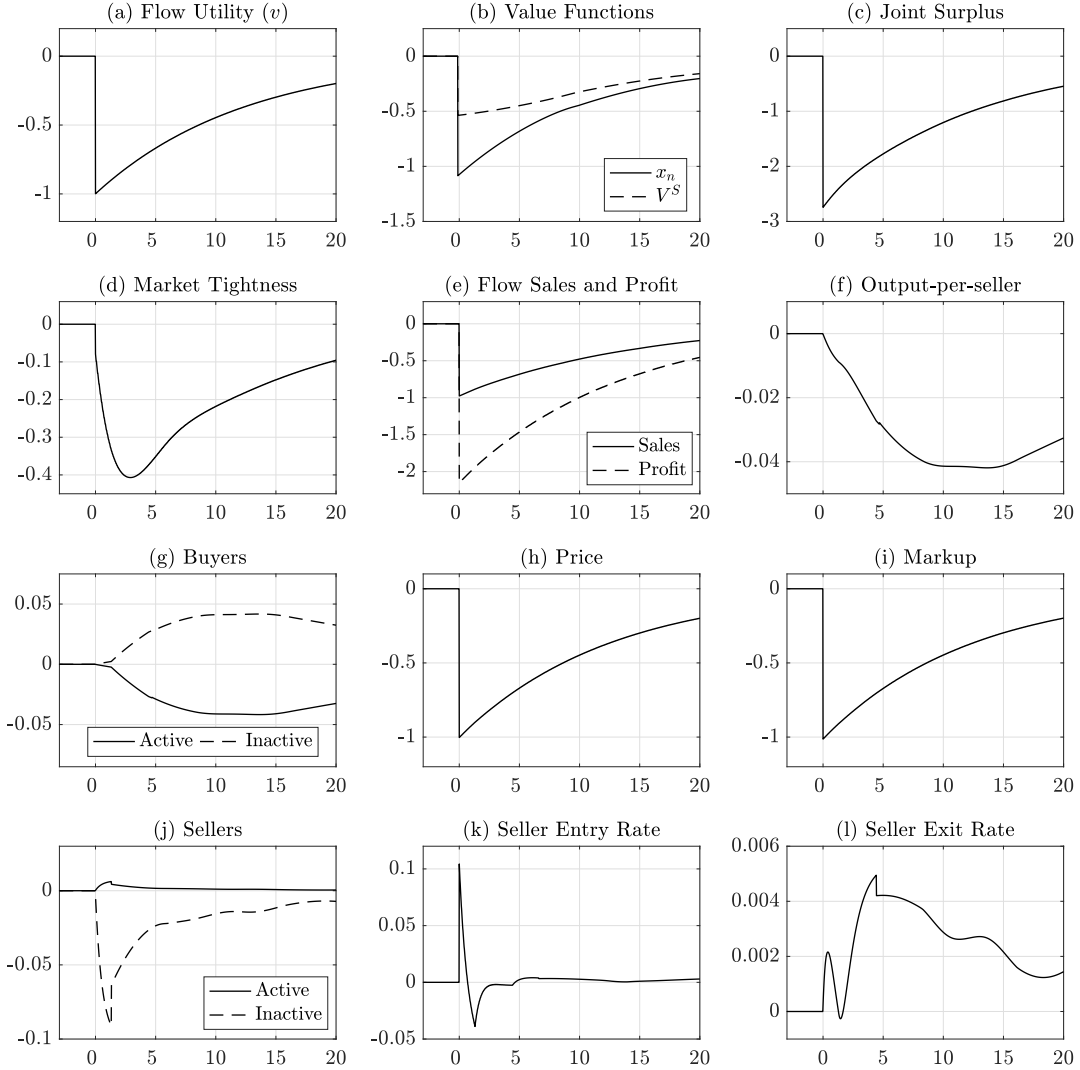
We implement a 1% negative shock to the instantaneous utility from consumption ( $v$ ), and let the process mean-revert without any further shocks.<sup>35</sup> Because of block-recursivity, aggregate dynamics are internalized by agents and anticipated fully. This makes demand “shocks” akin to business cycles rather than unexpected perturbations to the steady state equilibrium.

## 4.1 Aggregate Dynamics

Figure 9 presents the effects of a negative shock to the flow utility from consumption ( $v$ ) on a number of equilibrium variables. Panels (b)-(i) show the effects of the shock on cross-sectional averages. A negative shock to demand (panel (a)) leads to a decrease in the number of buyers looking for a seller, since consumption is now worth less to consumers. Because the buyers’ outside option has relatively improved, sellers can lower the continuation utility that they promise to deliver to each pre-existing customer going forward (panel (b)) in order to mitigate the adverse effects of the shock on their own profits. A large share of the burden of the shock is passed onto the customer: while the seller’s value decreases by half a basis point (dashed line), it is the sharp decrease in the value of the buyer (solid line), double in magnitude, which accounts for most of the instantaneous drop in joint surplus (panel (c)). As consumption is worth less and sellers offer worse deals to their customers, sellers attract fewer inactive buyers, and the average tightness in the market falls

<sup>35</sup> We simulate the economy on  $\mathbb{T} = \{\Delta, 2\Delta, \dots, \frac{T}{\Delta}\}$ , and choose a path for  $\{\varphi_t : t \in \mathbb{T}\}$  such that  $\varphi_t = \hat{\varphi}$  (its mean), for  $t \in \{\Delta, \dots, t_0 - \Delta\}$ ,  $\varphi_{t_0} = \underline{\varphi}$ , and  $\log \varphi_t = (1 - \rho_\varphi \Delta) \log \varphi_{t-\Delta}$ , for  $t > t_0$ , where  $v(\hat{\varphi}) = 1$  and  $v(\underline{\varphi}) = 0.99$ . We choose  $\rho_\varphi = 0.0807$ , implying the shock has a half-life of about 8.5 years.

(panel (d)). As a result, sellers generate lower sales and flow profit (panel (e)), the average output per seller declines (panel (f)), and the overall measure of active customers in the economy falls (panel (g)).<sup>36</sup> In these respects, the demand shock is akin to a productivity shock, in line with the intuition developed in [Bai \*et al.\* \(2017\)](#): by depressing buyer search, negative demand shocks lead to a contraction in economic activity.



**Figure 9:** Impulse responses of selected variables to a one-time -1% shock to marginal utility ( $v$ ). *Notes:* Series presented in %-deviations from the steady state where  $v(\varphi) = 1$ . The shock hits at date  $t_0 = 0$  and the time period is a year. All average variables have been computed using the theoretical distribution of sellers over states, and interpolated using cubic splines.

Panels (h) and (i) show the implications of the demand shock for average prices and markups. Since the incumbent customers of all sellers in the economy now value consumption relatively less, sellers must lower prices on impact. Note that the size of the instantaneous fall in prices is smaller

<sup>36</sup> As each buyer consumes one unit, output per seller is computed as the average number of buyers per seller, defined by  $\left(\sum_{n,z} \frac{1}{n} L_n(z, \varphi)\right)^{-1}$ , where  $L_n(z, \varphi) \equiv \frac{n S_n(z, \varphi)}{\sum_{n,z} n S_n(z, \varphi)}$  is the fraction of active buyers that are customers of seller  $(n, z)$  in aggregate state  $\varphi$ .



than that on continuation promises (solid line in panel (b)), reflecting the fact that, in response to the adverse shock, sellers transfer the cost of the shock away from their current customers and onto their future ones. In the transition, this burden is slowly shifted back from future to pre-existing buyers: as flow utility mean-reverts, sellers begin to offer better deals to their new customers (continuation values increase) at the expense of their pre-existing ones (spot prices rise). As there are no changes in marginal costs, the average markup exhibits a similar behavior.

All in all, the model predicts that prices and markups move procyclically in response to aggregate demand shocks. This is in sharp contrast to New Keynesian models, where markups are countercyclical in response to aggregate demand shocks because nominal rigidities prevent some firms from adjusting their prices to changes in marginal costs. In contrast, in our model markups are procyclical because sellers are forward-looking in their pricing decisions: in the wake of an adverse shock to their profits, they decide to inter-temporally shift resources between current and future demand, as they are constrained by the promises they made to their pre-existing buyers.<sup>37</sup>

## 4.2 Distributional Effects

Behind these aggregate dynamics lie important distributional effects. Panels (j), (k), and (l) of Figure 9 show the response of the measure of sellers, and the entry and exit rates, respectively. As a result of the shock, the average size of sellers declines for a few periods (panel (f)), until this trend is eventually reversed by the continued increase in promised utilities, causing seller growth to pick up in the transition. In the first recovery phase, therefore, the size distribution is gradually shifting to the left. As the risk of exiting becomes higher (because buyers' search intensity has declined), the exit rate goes up. Interestingly, the entry rate experiences a short-lived increase on impact, which is sufficiently strong for the overall measure of inactive sellers to decline.<sup>38</sup>

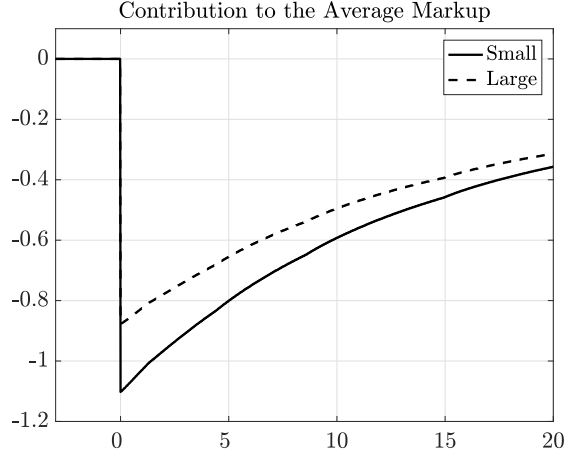
These shifts in the distribution imply cross-sectional differences in the markup response of different sellers. Figure 10 shows that the contribution of small sellers to the average markup is both stronger and more persistent than that of large sellers.<sup>39</sup> There are two reasons for these results. First, prices are relatively more elastic to size changes for smaller sellers because the pricing policy function is steeper for these sellers (recall Figure 3). Second, the entry of new sellers and the thinning of the right tail of the distribution, induced by the demand shock, amplify the relative contribution of small sellers to the average markup. Subsequently, the slow adjustment in the distribution implies that these cross-sectional differences persist in time, albeit the gap shrinks

---

<sup>37</sup> Note that, unlike the New Keynesian model, our set-up does not feature a general-equilibrium effect whereby demand shocks increase marginal costs for firms, as wages are exogenous. However, in Online Appendix D.2 we show that markups are procyclical in response to negative shocks to marginal costs as well. We argue that this is due to the same inter-temporal pricing trade-off discussed above.

<sup>38</sup> The reason for this behavior in the entry rate is that, unlike in all other markets, in the entry market tightness increases as a result of the shock. The intuition is as follows. In order to enter into the economy, potential sellers must raise resources from future customers to pay for the fixed market penetration cost,  $\kappa$ . Since inactive buyers' ex-ante match value has worsened, sellers must now set the initial promised compensation sufficiently high to guarantee that the costs of entry are still being recouped in expectation. As a result, the buyer-to-seller ratio spikes up.

<sup>39</sup> These observations are consistent with the empirical results in Hong (2017), who finds substantial heterogeneity in markup cyclicity across firms of different sizes.

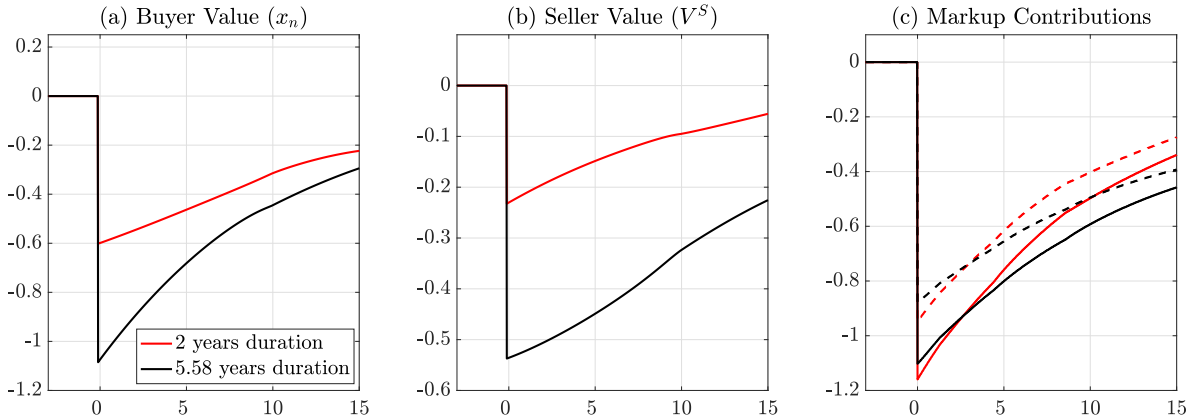


**Figure 10:** Impulse response to a one-time  $-1\%$  shock to marginal utility ( $v$ ): contribution to the average markup by small versus large sellers. *Notes:* More details in Figure 9.

as the distribution converges. These results suggest that taking into account the lifecycle patterns of establishments may be an important dimension to understand the cyclicity of markups.

### 4.3 Inspecting Markup Procyclicality

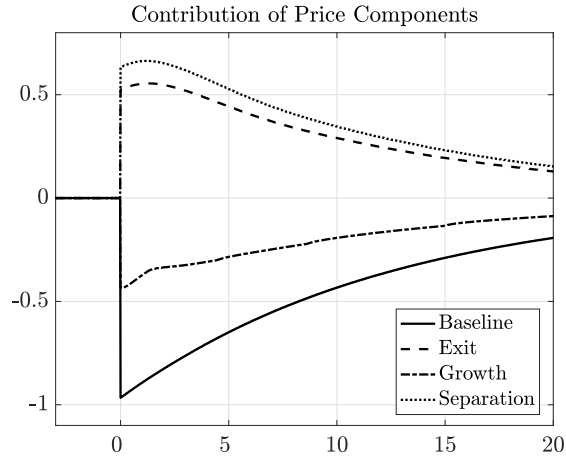
Taking stock, prices and markups are procyclical to aggregate demand perturbations, with stronger effects for smaller sellers. This section inspects different channels that help explain this behavior in the model. The key channel of markup adjustment are sellers' forward-looking pricing decisions. To illustrate this key dimension, we show that the amplitude and persistence of the markup response is shaped by the duration that sellers expect to be matched with their buyers. This margin is controlled by the customer turnover parameter,  $\delta_c$ . This parameter was pinned down using the establishment entry rate in the data, so it plays an important role in the heterogeneous response of the markup in the cross-section.



**Figure 11:** Impulse responses of selected variables to a one-time  $-1\%$  shock to marginal utility ( $v$ ), for different values of  $\delta_c$ . *Notes:* The figure shows the responses for an economy with high  $\delta_c$  (red lines) and the estimated  $\delta_c$  (black lines). On the right-most panel, the solid (respectively, dashed) lines are the contributions to the average markup by small (respectively, large) sellers. More details in Figure 9.

To see this, Figure 11 plots the responses to the demand shock of the buyer and seller value functions (left and central panels), together with the responses of the contributions to the markup by small and large sellers (right panel), under two different parametrizations. The black lines correspond to the estimated economy, where  $\delta_c = 0.1790$ , implying that the average seller-customer relationship lasts 5.58 years. The red lines depicts an economy in which this duration is 2 years (i.e.  $\delta_c = 0.5$ ). The overall responses on buyer and seller values are amplified by the expected length of relationships. Intuitively, when sellers expect their customers to remain captive for a shorter amount of time, they care more about immediate profits, so there is less inter-temporal shifting between customers. As a result, promised utility decreases less on impact (panel (a)), and the adverse effects of the shock on the seller’s valuation (panel (b)) are weaker.

Importantly, the duration of customer relationships has implications for the amplification and persistence of demand shocks on the average markup of the economy (panel (c)). When  $\delta_c$  is higher, markups fall more on impact for all sellers, and they mean-revert faster. Intuitively, when sellers expect to develop longer relationships with their customers, they prefer to shift the effect of the shock farther into the future, implying a more persistent effect on markups. For instance, the half-life of the markup response for smaller sellers is about 11 years when the expected duration of relationships is 5.58 years, and only about 8 years when the expected duration is 2 years. For large sellers, these numbers are approximately 13 and 6 years, respectively.



**Figure 12:** Impulse responses to a one-time  $-1\%$  shock to marginal utility ( $v$ ): decomposition of the contributions to the average price response of the components identified in equation (9). Notes: More details in Figure 9.

Finally, to shed light on the channels through which these markup adjustments occur, we decompose the average price response (panel (h) of Figure 9) into the contributions of the different components identified in equation (9). This decomposition allows us to understand the degree to which prices vary due to changes in fundamentals (“baseline” component) vis-à-vis changes in the incentives to compensate buyers for the prospect of seller exit, seller growth, or customer separation.<sup>40</sup> Figure 12 shows the response to the aggregate demand shock of the contribution of each of

<sup>40</sup> The “exit”, “growth”, and “separation” components of the price level are all negative in the estimated model. Intuitively, in order to remain matched, captive buyers require a price compensation for the potential loss in value

these components to the average price response. The procyclicality in the response of the average price (and therefore markup) is driven by the “baseline” and, to a lesser extent, the “growth” components. Intuitively, the first channel is due to changes in fundamentals: prices decrease because the value of consumption for captive customers is now lower. At the same time, sellers grow more concerned with expanding their base in the aftermath of the shock, so they implement a price compensation to their pre-existing customers in order to afford higher continuation promises to future buyers, as these will generate higher rates of subsequent growth. The contributions of the “exit” and “separation” components, by contrast, are countercyclical, though not strong enough to overturn the aforementioned procyclical forces.

To recap, in response to adverse shocks, markups decrease both because sellers are forced to reduce prices as buyers value consumption less, and because sellers expect to grow during the transition, for which they must compensate their buyers today via lower prices. This inter-temporal motive is shaped by the degree by which sellers expect to continue interacting with their buyers in the future. Intuitively, sellers use customers as a form of capital to smooth out the effects of demand shocks on markups over time.

## 5 Conclusion

Empirical evidence suggests that a major source of variation in the performance of businesses is due to heterogeneity across idiosyncratic demand components. Further, price differences are important drivers of revenue differences for firms that operate within the same product markets, even among those firms with similar productivity levels. These observations may help understand the behavior of markups at both the firm and the aggregate levels.

To address this question, this paper has presented a search model of demand accumulation through pricing with aggregate and idiosyncratic shocks and a relevant scope for firm dynamics. We have studied the connection of customer capital with the dynamics of establishments at the microeconomic level, and with the dynamics of the average markup at the macroeconomic level. Using establishment-level data for the U.S. retail sector, the estimated model can replicate patterns of growth and cross-sectional dispersion, such as the right-skewness in the size distribution, the distribution of sales growth, and declining growth and hazard exit rates with establishment age. We have then studied the response of the economy to aggregate demand shocks, and found both level and compositional effects. In response to adverse shocks, sellers inter-temporally smooth out the effects of the shock on prices by transferring the burden onto their future buyers, giving rise to markup procyclicality. Moreover, smaller sellers contribute relatively more to the average response, and the overall response has large amplitude and lower persistence when relationships are shorter.

Our analysis shows that incorporating micro-founded pricing behavior into quantitative macro models can shed light on firm and aggregate dynamics. Investigating the power of customer markets to rationalize these and other dimensions of the data remains an exciting avenue for future research.

---

due to the seller exiting the market or growing (both of which lower each buyer’s ex-post match value), and due to customer attrition (which lowers the average value for customers).

# References

- ALVAREZ, F., LeBIHAN, H. and LIPPI, F. (2016). The real effects of monetary shocks in sticky price models: a sufficient statistic approach. *American Economic Review*, **106** (10), 2817 – 2851.
- APEL, M., FRIBERG, R. and HALLSTEN, K. (2005). Micro foundations of macroeconomic price adjustment: Survey evidence from Swedish firms. *Journal of Money, Credit and Banking*, **37** (2), 313 – 338.
- ARGENTE, D., LEE, M. and MOREIRA, S. (2018). How do firms grow? The life cycle of products matters. *Working Paper, University of Chicago*.
- ARKOLAKIS, C. (2010). Market penetration costs and the new consumer margin in international trade. *Journal of Political Economy*, **118** (6), 1151 – 1199.
- ATKESON, A. and KEHOE, P. J. (2005). Modeling and measuring organization capital. *Journal of Political Economy*, **113** (5), 1026 – 1053.
- BAI, Y., RIOS-RULL, V. and STORESLETTEN, K. (2017). Demand shocks as productivity shocks. *Working Paper, University of Pennsylvania*.
- BAQAEE, D. R. and FARHI, E. (2019). Productivity and misallocation in general equilibrium. *NBER Working Paper 24007*.
- BERMAN, N., REBEYROL, V. and VICARD, V. (2019). Demand learning and firm dynamics: Evidence from exporters. *Review of Economics and Statistics*, **101** (1), 91 – 106.
- BILS, M. (1989). Pricing in a customer market. *The Quarterly Journal of Economics*, **104** (4), 699 – 718.
- BLINDER, A. S., CANETTI, E. R. D., LEBOW, D. E. and RUDD, J. B. (1998). *Asking About Prices: A New Approach to Understanding Price Stickiness*. Russell Sage Foundation.
- BRONNENBERG, B. J. and DUBÉ, J.-P. (2017). The formation of consumer brand preferences. *The Annual Review of Economics*, **9**, 353 – 382.
- , KRUGER, M. W. and MELA, C. F. (2008). The IRI marketing data set. *Marketing Science*, **27** (4), 745 – 748.
- BURDETT, K. and COLES, M. G. (1997). Steady state price distributions in a noisy search equilibrium. *Journal of Economic Theory*, **72**, 1 – 32.
- and JUDD, K. L. (1983). Equilibrium price dispersion. *Econometrica*, **51** (4), 955 – 969.
- , SHI, S. and WRIGHT, R. (2001). Pricing and matching with frictions. *Journal of Political Economy*, **109** (5), 1060 – 1085.
- CABRAL, L. and MATA, J. (2003). On the evolution of the firm size distribution: Facts and theory. *American Economic Review*, **93** (4), 1075 – 1090.
- CARLSSON, M., MESSINA, J. and SKANS, O. N. (2017). Firm-level shocks and labor flows. *Working Paper, Uppsala University*.
- CASON, T. N. and FRIEDMAN, D. (2003). Buyer search and price dispersion: A laboratory study. *Journal of Economic Theory*, **112** (2), 232 – 260.

- CAVENAILE, L. and ROLDAN-BLANCO, P. (2019). Advertising, innovation, and economic growth. *Working Paper, University of Toronto*.
- CAVES, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature*, **36** (4), 1947 – 1982.
- CHRISTOPOULOU, R. and VERMEULEN, P. (2008). Markups in the Euro Area and the U.S. over the period 1981-2004. *European Central Bank, Working Paper Series*, **856**.
- CLEMENTI, G. L. and PALAZZO, B. (2016). Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics*, **8** (3), 1 – 41.
- COAD, A. (2009). *The Growth of Firms: A Survey of Theories and Empirical Evidence*. Northampton, MA: Edward Elgar Publishing.
- COIBION, O., GORODNICHENKO, Y. and HONG, G. H. (2015). The cyclicality of sales, regular, and effective prices: Business cycle and policy implications. *American Economic Review*, **105** (3), 993–1029.
- DELOECKER, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, **79** (5), 1407 – 1451.
- , GOLDBERG, P. K., KHANDELWAL, A. K. and PAVCNIK, N. (2016). Prices, markups, and trade reform. *Econometrica*, **84** (2), 445 – 510.
- and WARZYNSKI, F. (2012). Markups and firm-level export status. *American Economic Review*, **102** (6), 2437 – 2471.
- DIAMOND, P. A. (1971). A model of price adjustment. *Journal of Economic Theory*, **3**, 156 – 168.
- DINLERSOZ, E. M. and YORUKOGLU, M. (2012). Information and industry dynamics. *American Economic Review*, **102** (2), 884 – 913.
- DUNNE, T., ROBERTS, M. J. and SAMUELSON, L. (1988). Patterns of firm entry and exit in U.S. manufacturing industries. *The RAND Journal of Economics*, **19** (4), 495 – 515.
- , — and — (1989). The growth and failure of U.S. manufacturing plants. *Quarterly Journal of Economics*, **104** (4), 671 – 698.
- ELSBY, M. W. L. and MICHAELS, R. (2013). Marginal jobs, heterogeneous firms, and unemployment flows. *American Economic Journal: Macroeconomics*, **5** (1), 1 – 48.
- EVANS, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *The Journal of Industrial Economics*, **35** (4), 567 – 581.
- FABIANI, S., DRUANT, M., HERNANDO, I., KWAPIL, C., LANDAU, B., LOUPIAS, C., MARTINS, F., MATHÄ, T. Y., SABBATINI, R., STAHL, H. and STOKMAN, A. C. (2005). The pricing behavior of firms in the Euro Area: New survey evidence. *European Central Bank, Working Paper Series*, (535).
- FAIG, M. and JEREZ, B. (2005). A theory of commerce. *Journal of Economic Theory*, **122** (1), 60 – 99.
- FISHMAN, A. and ROB, R. (2003). Consumer inertia, firm growth and industry dynamics. *Journal of Economic Theory*, **109**, 24 – 38.

- FITZGERALD, D., HALLER, S. and YEDID-LEVI, Y. (2017). How exporters grow. *Working Paper, Federal Reserve Bank of Minneapolis*.
- and PRIOLO, A. (2018). How do firms build market share? *Working Paper, Federal Reserve Bank of Minneapolis*.
- FOSTER, L., HALTIWANGER, J. and SYVERSON, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, **98** (1), 394 – 425.
- , — and — (2016). The slow growth of new plants: Learning about demand? *Economica*, **83**, 91 – 129.
- GAGNON, E. and LÓPEZ-SALIDO, D. (2014). Small price responses to large demand shocks. *Federal Reserve Board Discussion Paper*.
- GAVAZZA, A., MONGEY, S. and VIOLANTE, G. L. (2018). Aggregate recruiting intensity. *American Economic Review*, **108** (8), 2088 – 2127.
- GILCHRIST, S., SCHOENLE, R., SIM, J. and ZAKRAJŠEK, E. (2017). Inflation dynamics during the financial crisis. *American Economic Review*, **107** (3), 785 – 823.
- GOURIO, F. and RUDANKO, L. (2014a). Can intangible capital explain cyclical movements in the labor wedge? *American Economic Review*, **104** (5), 183 – 188.
- and — (2014b). Customer capital. *Review of Economic Studies*, **81** (3), 1102 – 1136.
- HALL, B. H. (1987). The relationship between firm size and firm growth in the U.S. manufacturing sector. *Journal of Industrial Economics*, **35** (4), 583 – 606.
- HALL, R. E. (2008). General equilibrium with customer relationships: A dynamic analysis of rent-seeking. *Working Paper, Stanford University*.
- HALL, S., WALSH, M. and YATES, A. (1997). How do U.K. companies set prices? *Bank of England Quarterly Bulletin*, pp. 180 – 192.
- HONG, S. (2017). Customer capital, markup cyclicity, and amplification. *Working Paper, Federal Reserve Bank of St. Louis*.
- HOPENHAYN, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, **60** (5), 1127 – 1150.
- HOTTMAN, C. (2017). Retail markups, misallocation, and store variety across U.S. cities. *Board of Governors of the Federal Reserve System, Working Paper*.
- , REDDING, S. J. and WEINSTEIN, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, **131** (3), 1291 – 1364.
- HSIEH, C.-T. and KLENOW, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, **124** (4), 1403 – 1448.
- JOVANOVIĆ, B. (1982). Selection and the evolution of the industry. *Econometrica*, **50** (3), 649 – 670.



- KAAS, L. and KIMASA, B. (2018). Firm dynamics with frictional product and labor markets. *Working Paper, Goethe University Frankfurt*.
- and KIRCHER, P. (2015). Efficient firm dynamics in a frictional labor market. *American Economic Review*, **105** (10), 3030 – 3060.
- KAPLAN, G. and MENZIO, G. (2015). The morphology of price dispersion. *International Economic Review*, **56** (4), 1165 – 1206.
- and — (2016). Shopping externalities and self-fulfilling unemployment fluctuations. *Journal of Political Economy*, **124** (3), 771 – 825.
- , —, RUDANKO, L. and TRACHTER, N. (2019). Relative price dispersion: Evidence and theory. *American Economic Journal: Microeconomics*, **11** (3), 68 – 124.
- KLEMPERER, P. (1987). Markets with consumer switching costs. *The Quarterly Journal of Economics*, **102** (2), 375 – 394.
- (1995). Markets with consumer switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *Review of Economic Studies*, **62**, 515 – 539.
- KLESHCHELSKI, I. and VINCENT, N. (2009). Market share and price rigidity. *Journal of Monetary Economics*, **56** (3), 344 – 352.
- KUGLER, M. and VERHOOGEN, E. (2012). Prices, plant size, and product quality. *Review of Economic Studies*, **79**, 307 – 339.
- LUTTMER, E. G. (2006). Consumer search and firm growth. *Working Paper, Federal Reserve Bank of Minneapolis*.
- MCGRATTAN, E. R. (2017). Intangible capital and measured productivity. *NBER Working Paper 23233*.
- and PRESCOTT, E. C. (2014). A reassessment of real business cycle theory. *American Economic Review*, **104** (5), 177 – 182.
- MENZIO, G. (2007). A search theory of rigid prices. *PIER Working Paper 07-031*.
- and SHI, S. (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, **145** (4), 1453 – 1494.
- and — (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, **119** (3), 468 – 510.
- and TRACHTER, N. (2018). Equilibrium price dispersion across and within stores. *Review of Economic Dynamics*, **28** (2), 205 – 220.
- MICHAILLAT, P. and SAEZ, E. (2015). Aggregate demand, idle time, and unemployment. *The Quarterly Journal of Economics*, **130** (2), 507 – 569.
- MOEN, E. R. (1997). Competitive search equilibrium. *Journal of Political Economy*, **105** (2), 385 – 411.

- MOSCARINI, G. and POSTEL-VINAY, F. (2013). Stochastic search equilibrium. *Review of Economic Studies*, **80** (4), 1545 – 1581.
- MUKOYAMA, T. and LEE, Y. (2015). Entry and exit of manufacturing plants over the business cycle. *European Economic Review*, **77** (C), 20 – 27.
- NAKAMURA, E. and STEINSSON, J. (2011). Price setting in forward-looking customer markets. *Journal of Monetary Economics*, **58** (3), 220 – 233.
- NEKARDA, C. J. and RAMEY, V. A. (2013). The cyclical behavior of the price-cost markup. *NBER Working Paper 19099*.
- NORRIS, J. (1997). *Markov Chains*. Cambridge University Press.
- PACIELLO, L., POZZI, A. and TRACHTER, N. (2019). Price dynamics with customer markets. *International Economic Review*, **60** (1), 413 – 446.
- PERLA, J. (2019). A model of product awareness and industry life cycles. *Working Paper, University of British Columbia*.
- PETERS, M. (2018). Heterogeneous markups, growth and endogenous misallocation. *Working Paper, Yale University*.
- PETROSKY-NADEAU, N. and WASMER, E. (2015). Macroeconomic dynamics in a model of goods, labor, and credit market frictions. *Journal of Monetary Economics*, **72**, 97 – 113.
- , — and ZENG, S. (2016). Shopping time. *Economics Letters*, **143**, 52 – 60.
- PHELPS, E. S. and WINTER, S. G. (1970). Optimal price policy under atomistic competition. *Microeconomic Foundations of Employment and Inflation Theory*, pp. 309 – 337.
- PIVETEAU, P. (2017). An empirical dynamic model of trade with consumer accumulation. *Working Paper, Johns Hopkins University*.
- POZZI, A. and SCHIVARDI, F. (2016). Demand or productivity: What determines firm growth? *RAND Journal of Economics*, **47** (3), 608 – 630.
- RAVN, M., SCHMITT-GROHÉ, S. and URIBE, M. (2006). Deep habits. *Review of Economic Studies*, **73**, 195 – 218.
- RENNER, E. and TYRAN, J.-R. (2004). Price rigidity in customer markets: An empirical study. *Journal of Economic Behavior and Organization*, **55** (4), 575 – 593.
- ROTEMBERG, J. J. and WOODFORD, M. (1991). Markups and the business cycle. *NBER Macroeconomics Annual*, **6**, 63 – 140.
- and — (1999). The cyclical behavior of prices and costs. *NBER Working Paper 6909*.
- RUDANKO, L. (2017). Pricing in a frictional product market. *Mimeo, Federal Reserve Bank of Philadelphia*.
- SCHAAL, E. (2017). Uncertainty and unemployment. *Econometrica*, **85** (6), 1675 – 1721.
- SHI, S. (2009). Directed search for equilibrium wage-tenure contracts. *Econometrica*, **77** (2), 561 – 584.

- SMITH, H. (2004). Supermarket choice and supermarket competition in market equilibrium. *Review of Economic Studies*, **71** (1), 235 – 263.
- STROEBEL, J. and VAVRA, J. (2019). House prices, local demand, and retail prices. *Journal of Political Economy*, **127** (3), 1391 – 1436.
- VARIAN, H. R. (1980). A model of sales. *American Economic Review*, **70** (4), 651 – 659.