Firm Dynamics and Pricing under Customer Capital Accumulation*

Pau Roldan-Blanco† Sonia Gilbukh‡

June 24, 2019

Abstract

This paper analyzes the macroeconomic implications of customer capital accumulation at the firm level. We build an analytically tractable search model of firm dynamics in which firms compete for customers by posting pricing contracts in the product market. Cross-sectional price dispersion emerges in equilibrium because firms of different sizes and productivities use different pricing strategies to strike a balance between attracting new customers and exploiting incumbent ones. Using micro-pricing data from the U.S retail sector, we calibrate the model to match moments from the cross-sectional distribution of sales and prices, and use our estimated model to explain sluggish aggregate dynamics and cross-sectional heterogeneity in the response of markups to aggregate shocks. We find that there is incomplete price pass-through leading to procyclicality in the average markup, with smaller firms being more responsive to shocks than larger firms.

JEL codes: D21, D83; E2; L11

Keywords: Customer Capital; Product Market Frictions; Directed Search; Firm Dynamics; Dynamic Contracts; Price Dispersion.

*We are indebted to Ricardo Lagos, Jess Benhabib, Boyan Jovanovic, and Edouard Schaal, for their invaluable guidance and continued discussions. We have benefited from comments by 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, and Basil Williams; from seminar audiences at Banco de España, BI Norwegian Business School, Carlos III, CEMFI, EUI, IESE, NYU, FRB St. Louis, UAB, University of Essex, and University of Toronto Rotman; and from participants at SED 2017, the 2017 Search and Matching Network conference, the 2017 Summer Meetings of the Econometric Society, SAEe 2017, the 12th Joint French Macro Workshop, T2M 2019, the 2019 Barcelona GSE Summer Forum, and CEBRA 2019. 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.

†Banco de España, Calle Alcalá 48, 28014 Madrid. Corresponding author. Email: pau.roldan@bde.es.

‡CUNY Baruch College, Zicklin School of Business, 137 East 22nd Street, New York, NY 10010. Email: sophia.gilbukh@baruch.cuny.edu.

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 traditionally attributed this evidence to selection on the basis of productivity differences among 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)). In particular, new empirical evidence based on micro data has shown that a large cross-sectional dispersion in firm revenue remains even after controlling for heterogeneity in technological productivity. In light of this fact, the literature has suggested that differences in firm performance in the cross-section and over time stem, to a great extent, from differences in firms' idiosyncratic demand-side components. Further, the evidence suggests that this demand-side channel of variation is persistent. These facts thus point to firm investment in demand as a key factor behind the existing differences in the life-cycle of businesses of similar productivity.[1] Yet, few studies have formalized the firm dynamics implications of these investment decisions, and their consequences at the aggregate level.

In this paper, we aim to fill this gap by developing a theory of firm dynamics in product markets, in which there is a meaningful role for a demand accumulation process at the firm level. We interpret this process as the formation of a customer base. Section 2 presents a directed search model of a frictional product market in which a fixed measure of ex-ante identical buyers must search for sellers of different productivities that sell the same homogenous product. All sellers post, and commit to, long-term price contracts designed to (i) attract new potential customers, and (ii) retain pre-existing ones. Ex-post, *within* a productivity type, sellers become heterogeneous in the size of their customer base, since the pricing contract endogenously determines the rate at which demand accumulates. This generates price dispersion along two dimensions: productivity and size. Additionally, sellers may lose customers due to exogenous separation shocks. When a seller loses all its customers, it may (re-)enter the product market by paying a fixed entry (or "market penetration") cost, allowing it to attract its first buyer.

In equilibrium, sellers strike a balance between instantaneous revenues (via high prices today) and future market shares (via lower prices in the future).[2] This inter-temporal trade-off determines

[1] Hottman *et al.* (2016) show, for the U.S. retail market, that most variation in the firm size distribution is attributable to variation in demand components, while Foster *et al.* (2008, 2016) show, for U.S. commodity markets, that the (technological) productivity advantage of entrants is small and it dissipates over the first few years of operation, while idiosyncratic demand accounts for the bulk of the observed heterogeneity. Related studies using data for other countries include Carlsson *et al.* (2017) (Sweden), Pozzi and Schivardi (2016) (Italy), Hong (2017) (France), Kaas and Kimasa (2018) (Germany), and Kugler and Verhoogen (2012) (Colombia). A related literature argues that markups are important contributors to revenue dispersion, including Hsieh and Klenow (2009), DeLoecker (2011), and Peters (2018).

[2] Therefore, in our setting, the Diamond (1971) paradox does not hold, as sellers do not extract all the surplus from

the rate of firm growth. In Section 3, we argue that the way the trade-off is resolved depends on the size of the seller's customer base. Interestingly, the sign of the correlation between prices and firm size depends on the degree of frictions in the market. On the one hand, when market penetration costs are relatively high, small sellers optimally decide to promise high continuation values to the buyers in order to generate a high rate of expansion, and thereby raise enough resources to grow out of the entry cost. As these firms grow further, they lower their future promises and increase their prices as they increasingly prefer to exploit their base at the expense of future market shares. In this case, firms' markups tend to increase as they grow in size, and the firm's rate of growth slows down. On the other hand, when penetration costs are relatively low, the firm might instead be willing to lower its prices as it grows, because it has a weaker preference for rapid growth at the early stages of its life. In either scenario, firms converge to a stationary size as the endogenous customer acquisition process is counteracted by per-customer separation and exit shocks. Therefore, on top of price dispersion, the model generates a well-defined and right-skewed firm distribution. Furthermore, we show that 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 (contrained-)optimal allocation of customers across product markets.

In Section 4, we estimate our model in order to quantify the aggregate implications of firm-level customer accumulation. To obtain computational tractability, we exploit that the equilibrium is *block-recursive*, a common property in models of directed search (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 firm distribution across aggregate states and over time. We estimate the model using highly disaggregated product-level pricing data for the U.S. retail sector (2001-2007), and calibrate the model to moments of the distribution of relative prices and sales. We show that the model provides a good fit of the data, and that it is in line with standard micro-pricing statistics such as the frequency and the size of price changes.

Using the estimated model, in Section 5 we explore the aggregate implications of pricing under customer capital accumulation by analyzing the behavior of the economy in the presence of aggregate demand and aggregate supply fluctuations. In this exercise, we find both level and distributional effects. First, we show that the price pass-through of aggregate temporary supply shocks (e.g. marginal costs) is incomplete: in the wake of an adverse shock, firms choose to trade-off immediate losses to future market shares by inter-temporally transferring the burden of shocks onto their buyers via slightly higher prices today and lower promises for tomorrow. Since the price level reacts less than one for one to the increase in marginal costs, the markup is procyclical. Aggregate demand shocks also generate markup procyclicality: shocks that lower the marginal propensity to consume by buyers generate a bust in demand and instantaneously lower prices. Since the shock mean-reverts, firms depress their promises on impact but increase prices in the transition. Empirically, the cyclicality of markups remains a topic of contention in macroeconomics, with some studies finding countercyclicality (e.g. Galeotti and Schiantarelli (1998) and Bils *et al.* (2018)),

buyers because they expect to keep serving them in the future.

and some others arguing for procyclicality (e.g. Nekarda and Ramey (2013)). Our model provides a rationale for the latter type of observations.

A key contribution of our paper is to show that there also exist important distributional effects in markups during the transition after aggregate shocks. Through a decrease in the continuation promise of firms, both negative supply and negative demand shocks lead to a decrease in the number of new matches, and firms temporarily shrink in size. This implies that the pass-through is less incomplete for larger firms, because (i) along the extensive margin, there is a left-ward shift in the size distribution, and (ii) along the intensive margin, a small firm's pricing policy is relatively more sensitive to size changes, for these firms are more eager to grow. Moreover, we find that small firms experience a more persistent response, because during the transition the fraction of low-price, small-size firms relatively increases and takes time to adjust.

Finally, through a novel decomposition of the price level, in Section 5.2 are able to identify the main margins of adjustment in prices in response to shocks. After adverse supply shocks, the main driver of the pass-through incompleteness is what we label the "firm-growth component" of prices, namely a force pushing for lower prices today in order to compensate customers for the increase in firm growth in the transition, which the customers dislike. Absent this channel, pass-through would be much less incomplete, leading to a less procyclical response in markups. Upon demand shocks, by contrast, most variation occurs purely through tastes: a large part of the price reaction is simply due to the fall in buyers' marginal utility. The channels of price adjustment are therefore different depending on the nature of the shock.

Related Literature To build our model of pricing under customer capital, we rely on a large amount of survey evidence suggesting that the customer base of firms are an important dimension for pricing decisions in reality. For example, Blinder *et al.* (1998) show that the vast majority of firms report having implicit contracts with their customers.[3] Our model with dynamic pricing contracts with commitment on the seller's side aims to capture this observation directly. In doing so, the paper speaks to several strands in the literature. First, it is related to the literature that incorporates customer capital into macroeconomic models of firm pricing. Among early studies, Phelps and Winter (1970), Bils (1989), and Rotemberg and Woodford (1991, 1999), all analyzed pricing behavior under customer retention concerns. In these papers, firms face an exogenous law of motion for their customer base. Subsequent research has provided possible reasons for the emergence of 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 we substantially differ in the modeling techniques.[4]

[3] Similar observations have been made in Hall *et al.* (1997), Cason and Friedman (2003), Renner and Tyran (2004), Fabiani *et al.* (2004), 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)).

[4] For example, 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 allows us to

More generally, this paper is related to a literature introducing a role for firm intangibles into models of firm and industry dynamics.[5] Within this literature, the paper that we most relate to is Gourio and Rudanko (2014b). Building on their work, we analyze 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, since pricing contracts are inter-temporal in nature.[6] As a result, prices may increase or decrease in firm size, depending on parameters. In this sense, our model does not take a stance on the empirical debate regarding the dynamics of firm-level prices, where the literature has found mixed evidence.[7] Moreover, relative to both Gourio and Rudanko (2014b) and Rudanko (2017), our model (i) incorporates competition for customers across firms of different productivities, giving rise to a firm size distribution and to price dispersion; and (ii) features 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.[8] Another related paper is Dinlersoz and Yorukoglu (2012), where customer acquisition is also a driver of firm and industry dynamics. In their paper, though, dynamics emerge from the dissemination of information about the firm's fundamentals, while in our paper there is perfect information about the sellers' attributes and dynamics emerge from forward-looking contracts. Paciello *et al.* (2017) explore how idiosyncratic productivity affects the pricing decisions of firms with customer accumulation concerns, 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 paper is also related to the search literature on price dispersion, pioneered by Varian (1980) and Burdett and Judd (1983).[9] Menzio and Trachter (2018) and Kaplan *et al.* (2018) 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 firms 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

speak to the firm-dynamics patterns emphasized by Foster *et al.* (2008) and others.

[5] 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 on macroeconomic aggregates have been widely studied, from labor wedges (Gourio and Rudanko (2014a)) and aggregate productivity (McGrattan and Prescott (2014), McGrattan (2017)), to household behavior (Hall (2008)) and the evolution of industries (Atkeson and Kehoe (2005), Perla (2017)).

[6] In the baseline model we assume no price discrimination because this allows us to uniquely pin down prices from firm size (equation (15)). However, in Appendix C we extend the model to allow for price discrimination, and show that this assumption is innocuous for firm dynamics.

[7] On the one hand, Foster *et al.* (2008, 2016) and Piveteau (2017) find that prices are increasing in the firm's tenure in the market. On the other hand, Berman *et al.* (2017) find that they are slightly decreasing. Finally, Fitzgerald *et al.* (2017) find no dynamics.

[8] See also Kaas and Kimasa (2018), who combine frictional product and labor markets to study the joint dynamics of prices, output, and employment.

[9] These papers have inspired a vast literature. For instance, Kaplan and Menzio (2016) also find procyclical markups in a model with product market frictions, though through a mechanism different from ours.

and Rob (2003) discuss the implications for the firm size distribution, but in neither of those papers is there a meaningful role for prices.

Finally, from a methodological standpoint, our paper is related to search-and-matching models with large firms. We embed *directed* search (after Moen (1997)) into a model of firm dynamics in the spirit of Elsby and Michaels (2013) or Kaas and Kircher (2015). Particularly, we combine two technical insights from this literature. First, we exploit the property of block recursivity, which allows for a tractable characterization of the firm distribution and its dynamics. Secondly, we make 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, we make two technical contributions: (i) we show that a continuous-time setting with Markov shocks yields further analytical tractability, with a single optimality condition (equation (12)) encompassing the relevant trade-offs, an analytical formulation of the joint surplus (Proposition 3), and an analytical decomposition of the price level (equation (15)); (ii) we show that we can generate size-dependent firm growth rates even when firms do not have a decreasing returns to scale technology (such as convex adjustment costs).

2 Model

2.1 Environment

Time is continuous, infinite, and indexed by $t \in \mathbb{R}_+$. The aggregate state of the economy is indexed by a time-varying random variable φ taking values in a discrete and finite support $\Phi \equiv \{\underline{\varphi} < \cdots < \overline{\varphi}\}$, with cardinality $k_\varphi \equiv |\Phi| \geq 2$. The aggregate state is the source of exogenous aggregate demand and/or supply fluctuations. φ follows a homogenous continuous-time Markov chain with generator matrix $\mathbf{\Lambda}_\varphi \equiv \left[\lambda_\varphi(\varphi'|\varphi)\right]$, where $\lambda_\varphi(\varphi'|\varphi)$ denotes the intensity rate of a φ-to-φ' transition.[10]

The economy is populated by a measure-one continuum of risk-neutral, infinitely-lived, ex-ante identical *buyers*, and a continuum of risk-neutral *firms* (also referred to as *sellers*). The total mass of buyers is exogenous and normalized to unity, though the composition of buyers across aggregate states and between types is endogenous. The total measure of firms is endogenous. Agents discount future payoffs with a common and exogenous 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 market in order to engage in trade because the product market is frictional: searchers cannot coordinate into finding a match with certainty at any given instant. The product market frictions are meant to capture congestion effects in

[10] For all $\varphi \in \Phi$, the following properties hold: $\lambda_\varphi(\varphi|\varphi) \leq 0$, $\lambda_\varphi(\varphi'|\varphi) \geq 0$ for any $\varphi' \neq \varphi$, and $\sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) = 0$. These properties are definitional of continuous-time Markov processes (e.g. Norris (1997), Chapters 2 and 3). Additionally, $\sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) < +\infty$, $\forall \varphi$ (i.e. the economy always spends a non-zero measure of time in any given state, when visited).

product markets with customer anonymity. One interpretation is that there exist informational asymmetries regarding product characteristics, or some aspects of supply that are unknown to the potential customer (e.g. the exact location of seller-price pairs). Another interpretation is that sellers may face inventory or capacity constraints, and are unable to simultaneously serve a large amount of buyers (as in Burdett *et al.* (2001)). These demand considerations lead businesses to invest in reputation-building in order to overcome those frictions.[11]

Buyers value the consumption of the good by the same fixed utility flow, $v > 0$.[12] At any instant in time, a buyer is said to be *active* if it is matched with a firm and is consuming the good, and *inactive* if it is unmatched and searching for a seller at a cost, c. The parameters (c, v) possibly depend on the aggregate state of nature, φ, which incorporates the possibility of aggregate and exogenous demand fluctuations into the model.[13]

Sellers belong to one of two groups: incumbent (or *active*) sellers, and potential entrant (or *inactive*) sellers. At any given time t, a typical incumbent seller has a customer base of $n_t \in \mathbb{N} \equiv \{1, 2, 3, \dots\}$ customers, which we subsequently call the *size* of the seller. Besides size, seller also differ in their idiosyncratic productivity level z, which takes values on a discrete and finite support $\mathcal{Z} \equiv \{\underline{z} < \cdots < \overline{z}\}$ of cardinality $k_z \equiv |\mathcal{Z}| \geq 2$. Like the aggregate state, the idiosyncratic state follows a continuous-time Markov chain with generator matrix $\mathbf{\Lambda}_z \equiv \left[\lambda_z(z'|z)\right]$, where $\lambda_z(z'|z)$ denotes the transition rate from z to z'.[14]

An incumbent 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, with each customer consuming one unit. The seller also faces operating variable flow costs $\mathcal{C}(n; z, \varphi)$ for serving customers. These depend on the idiosyncratic state (n, z), as well as possibly the aggregate state φ. Furthermore:

Assumption 1 *(i) $\mathcal{C}$ is continuous and increasing in n, with $\mathcal{C}(n; z, \varphi) \geq 0$ and $\mathcal{C}(0; z, \varphi) = 0$.*
(ii) $\mathcal{C}(n; z, \varphi)$ is weakly convex in n.

Assumption 1 imposes mild regularity conditions on the firm's selling technology. It states that firm profits are continuous in firm size. The curvature of $\mathcal{C}$ with respect to n determines the degree of returns to scale. Here, we need not make an explicit assumption besides *weak* convexity. Indeed, as we shall see, equilibrium prices depend non-linearly on sizes (giving rise to non-proportional firm growth) even when marginal costs are constant in n.

Besides serving their customers, incumbent sellers post prices in the product market. Posting a price bears no explicit cost for an incumbent seller. We assume, though, that incumbent sellers

[11] Informational frictions in the product market are the preferred interpretation of Faig and Jerez (2005), Gourio and Rudanko (2014b), and Foster *et al.* (2016), among others.
[12] Appendix F.1 shows a micro-foundation for v based on a more familiar and general environment with CES preferences.
[13] The source of variation in shopping disutility can be thought of as reflecting the cyclical nature of household shopping behavior, documented by Petrosky-Nadeau *et al.* (2016) for the United States.
[14] The usual conditions apply. For all $z \in \mathcal{Z}$: $\lambda_z(z|z) \leq 0$; $\lambda_z(z'|z) \geq 0$, $\forall z' \neq z$; $\sum_{z' \in \mathcal{Z}} \lambda_\varphi(z'|z) = 0$; and $\sum_{z' \neq z} \lambda_\varphi(z'|z) < +\infty$.

may exit the market (and enter the pool of potential entrants). This may occur in one of two ways: (i) if they go bankrupt and lose all customers at once, at an exogenous rate $\delta_f > 0$, or (ii) if they separate from their last remaining customer (because the buyer leaves), at an exogenous rate $\delta_c > 0$.[15] These events are mutually independent, and orthogonal to the (z, φ)-shocks.

Like incumbent firms, inactive firms are posting prices in order to attract customers and start operating in the product market. However, they must incur an entry cost $\kappa > 0$, which possibly depends on the aggregate state of nature, φ. The fixed cost κ can be thought of as a proxy for the costs of maintaining idle product technology or, more broadly, as a cost to market penetration in the sense of Arkolakis (2010). Sellers who successfully attract their first customer (and thus start operating with $n = 1$) draw an initial productivity level $z_0 \in \mathcal{Z}$ from some distribution π_z, where $\pi_z(z) \geq 0$, $\forall z \in \mathcal{Z}$, and $\sum_{z\in\mathcal{Z}} \pi_z(z) = 1$. We assume free entry of firms into the product market.

Pricing Contracts

At every instant, sellers announce price contracts in order to attract buyers. Buyers can perfectly observe the posted contract and visit the seller posting it. For a customer-seller match formed at time t, a *price contract* is defined as a sequence $(p_{t+j} : j \geq 0)$, which specifies the price level at each tenure $j \geq 0$, conditional on no separation. Contracts are complete and state-contingent. Thus, every element p_{t+j} of the contract is contingent on the history of aggregate and the firm's idiosyncratic states up to date $t + j$. Since all the relevant states are public, then $p_{t+j} = p(n^{t+j}; z^{t+j}, \varphi^{t+j})$, $\forall j, t$.

The contractual environment is as follows. On the demand side, we assume no commitment to the contract, in that matched buyers can costlessly transition to inactivity if they so desire (though in equilibrium this will not occur because of the subsequent additional cost c of re-sampling firms). On the sellers' side, we make two key assumptions. First, the seller fully commits to the contract that is posted. This means that contracts with captive customers cannot be revised by the firm for the duration of the match, and contracts have to comply with the firm's prior promises.[16] Second, we assume anonymity among buyers, in that the firm is unable to discriminate between new and old customers, and thus cannot index the contract to the buyer's identity.[17] This helps us pin down the price uniquely as a function of firm size. In Appendix C we show, however, that the existence of firm dynamics in equilibrium does not hinge on this assumption.

Product Markets

A sufficient statistic for each long-term pricing contract is the promised life-time value x that the contract delivers in expectation to the buyer at the point in time when the match is formed and the contract is initiated. We let $\mathcal{X} = [\underline{x}, \overline{x}] \subseteq \mathbb{R}_+$ be the set of feasible values, and assume that all sellers

[15] In Appendix F.2 we suggest ways to endogenize this customer separation rate.

[16] An interpretation of this assumption is that firms have a reputational concern, so that reneging on previous promises entails unaffordable costs for them. We discuss this assumption further in Section 3.

[17] Intuitively, this is meant to capture that, in densely populated markets where implicit relationships develop, buyer valuations are unknown to the seller.

advertising the same x compete in all such contracts. Thus, x effectively indexes a segment of the product market. Each seller can simultaneously post, and each buyer can simultaneously search, in at most one market segment. For each realization $\varphi \in \Phi$ of the aggregate state, let $B(x;\varphi) \in [0,1]$ be the mass of buyers seeking to be matched under promised utility x, and $S(x;\varphi) \geq 0$ be the mass of sellers posting x. A market is said to be *active* if:

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

where $\theta(x;\varphi)$ is the buyer-to-seller ratio in market segment x, or the *market tightness*. Agents take the mapping $\theta : \mathcal{X} \times \Phi \to [0, +\infty)$ as given when directing their search toward specific offers. In a typical $x \in \mathcal{X}$, a buyer obtains offer x at the endogenous Poisson arrival rate $\mu\big(\theta(x;\varphi)\big) \geq 0$, while a seller successfully finds a buyer for offer x at the Poisson arrival rate $\eta\big(\theta(x;\varphi)\big) \geq 0$, where $\eta(\theta) = \theta\mu(\theta)$. Further, we assume:

Assumption 2 *(i) $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ and $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ are twice continuously differentiable;*

(ii) η is increasing and concave; μ is decreasing and convex;

(iii) For some decreasing $h : \mathbb{R}_+ \to \mathbb{R}_+$, define $f = \eta \circ \mu^{-1} \circ h$. Then, the function $f(x)(\widehat{x} - x)$ is concave for all $x \in [0, \widehat{x}]$ and $\widehat{x} > 0$;

(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 first two restrictions guarantee that the problems of the buyer and the seller are well-defined; assumption *(iii)* is a restriction on the composition $\eta \circ \mu^{-1}$ guaranteeing that the price-posting problem of the seller has a unique interior solution; finally, part *(iv)* is a transversality condition on the meeting rates.

Recursive Formulation

We seek to solve for the *symmetric Markov perfect equilibrium* of this economy. We narrow attention to this class of equilibria in the following sense. (i) *Markov-perfection* means that the equilibrium policies depend solely on the firm's vector of payoff-relevant states $(n, x; \mathbf{s})$, where henceforth we denote $\mathbf{s} \equiv (z, \varphi)$.[18] (ii) We look for a *symmetric* equilibrium in the sense that all firms within the same product market x choose to post the same contract. This is a consequence of the assumption that there is competition within each market segment, and the fact that the firm's states are fully observable. (iii) Finally, we restrict our attention to a *stationary* environment, in which policies are time-varying only insofar as they are state-dependent. Thus, subsequently we drop time subscripts unless otherwise needed.

[18] Assuming Markov perfection in an environment without price discrimination allows us to obtain unique predictions for prices. However, in these type of models with dynamic contracts and commitment, real variables are uniquely determined even when Markov perfection is not imposed. The reason is that sellers find it equivalent to change current prices for future promised utilities.

Because the dynamic contract is a price trajectory and thus a large and potentially complex object, we exploit the property of stationarity to propose the following recursive formulation. We define a *recursive dynamic contract* for a firm in state $(n, x; \mathbf{s})$ as:

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

The elements of a recursive contract $\boldsymbol{\omega}$ are the following. First, the contract specifies the price p that is to be charged to each one of the n incumbent customers of the firm. Second, the contract specifies the vector $\mathbf{x}'(n'; \mathbf{s}') \subseteq \mathcal{X}$ of continuation payoffs that are promised by the firm to each buyer in the next stage, i.e. under every possible size $n' \in \{n-1, n, n+1\}$ and exogenous state $\mathbf{s}' \in \big\{(z', \varphi), (z, \varphi')\big\}$. By stationarity, contracts are rewritten every time the seller changes sizes or productivity, or if an aggregate shock hits, and they remain in place otherwise (i.e. $x'(n'; \mathbf{s}') = x$ when $n' = n$ and $\mathbf{s}' = \mathbf{s}$). Notice, finally, that the contract is not indexed to the aggregate distribution of agents across states. This is an implication of the property of *block recursivity*, which we take as given and we discuss in some detail in Section 2.4.

2.2 Value Functions

Inactive Buyers

Let $U^B(\varphi)$ be the expected value of an inactive buyer in state $\varphi \in \Phi$. The buyer enters the market segment that offers the highest valuation, and therefore:

$$U^B(\varphi) = \max_{\widehat{x}(\varphi) \in \mathcal{X}} \; u^B\big(\widehat{x}(\varphi); \varphi\big) \tag{1}$$

where $u^B\big(x; \varphi\big)$ is the value of searching in market x, satisfying the HJB equation:

$$ru^B\big(x; \varphi\big) = -c(\varphi) + \mu\big(\theta(x; \varphi)\big)\Big(x - u^B\big(x; \varphi\big)\Big) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)\Big(u^B\big(x; \varphi'\big) - u^B\big(x; \varphi\big)\Big) \tag{2}$$

for any $x \in \mathcal{X}$. Equation (1) states that the inactive buyer searches in the product market that promises the highest expected value. The value of entering market x incorporates the search cost $c(\varphi) > 0$, and the option value of matching with a firm within said market. The last additive term in equation (2) incorporates the expected change in value due to a change in the aggregate state.[19]

Since inactive buyers choose the best market to search in, all active markets must be equally attractive ex-ante. Therefore:

$$\forall (x, \varphi) \in \mathcal{X} \times \Phi: \quad u^B\big(x; \varphi\big) \leq U^B(\varphi), \quad \text{with equality if, and only if, } \theta(x; \varphi) > 0$$

[19] Notation has been economized in two ways. First, since the value of inactivity is itself an equilibrium object, we write $\theta(x; \varphi)$ when in fact we mean $\theta(x; \varphi, U^B)$. Second, since market tightness is taken as given, $u^B\big(x; \varphi\big)$ is short for $u^B\big(x; \varphi, \theta\big)$, where $\theta : \mathcal{X} \times \Phi \to \mathbb{R}_+$ is a function. Similar concise notation will be used throughout the paper.

This says that a market either maximizes ex-ante payoffs for the inactive buyer, or it remains unvisited. In equilibrium, a non-zero measure of markets is active, and we let $\mathcal{X}^*(\varphi) \equiv \big\{x \in \mathcal{X} : \theta(x;\varphi) > 0\big\} \subseteq \mathcal{X}$ be the *equilibrium* set of markets in state $\varphi \in \Phi$. Hence, for any given $\varphi \in \Phi$:

$$\mu\Big(\theta(x;\varphi)\Big)\Big(x - U^B(\varphi)\Big) = \Gamma^B(\varphi) \tag{3}$$

for all $x \in \mathcal{X}^*(\varphi)$, where we have defined the opportunity cost of matching for the buyer in equilibrium market x as:

$$\Gamma^B(\varphi) \equiv c(\varphi) + rU^B(\varphi) - \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)\Big(U^B(\varphi') - U^B(\varphi)\Big) \tag{4}$$

By equation (3), $\theta(x;\varphi)$ is an increasing function of $x \in \mathcal{X}$. This result is intuitive: more ex-post profitable offers attract a larger mass of buyers per seller, while sellers offering less favorable contracts can expect to find a match sooner. In equilibrium, firms design contracts for which a low meeting rate for buyers is compensated with higher expected promised values. Further, the buyer-to-seller ratio is increasing in U^B: when the inactive buyers' outside option is higher, contracts must offer more attractive deals in order to compensate for the opportunity cost of matching.

Active Buyers

Consider now a buyer who is currently consuming the homogeneous good from a firm of size n and idiosyncratic productivity z, under contract $\boldsymbol{\omega} = \big\{p, \mathbf{x}'(n';\mathbf{s}')\big\}$. The contract specifies the current price and the new continuation promises to be delivered by the seller under each new possible state. The value for the buyer is given by the HJB equation:

$$\begin{aligned} rV^B(n,\boldsymbol{\omega};\mathbf{s}) = {} & v(\varphi) - p + (\delta_f + \delta_c)\Big(U^B(\varphi) - V^B(n,\boldsymbol{\omega};\mathbf{s})\Big) \\ & + (n-1)\delta_c\Big(x'(n-1;\mathbf{s}) - V^B(n,\boldsymbol{\omega};\mathbf{s})\Big) \\ & + \eta\Big(\theta\big(x'(n+1;\mathbf{s});\varphi\big)\Big)\Big(x'(n+1;\mathbf{s}) - V^B(n,\boldsymbol{\omega};\mathbf{s})\Big) \\ & + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z)\Big(x'(n;z',\varphi) - V^B(n,\boldsymbol{\omega};\mathbf{s})\Big) \\ & + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)\Big(x'(n;z,\varphi') - V^B(n,\boldsymbol{\omega};\mathbf{s})\Big) \end{aligned} \tag{5}$$

The right side of equation (5) has different additive terms. In the first line, the first term, $v - p$, shows flow surplus for the agreed-upon price, and the second term states the possibility of separation, due to either the destruction of the firm or the destruction of the match. The second line includes the event that a customer (other than the one in question) separates, in which case the firm becomes size $n-1$ and changes the promised value to all those customers that remain captive. The third line is the expected change in value due to the firm successfully attracting a customer, in which case the seller becomes size $n+1$ and implements value $x'(n+1;\mathbf{s}) \in \boldsymbol{\omega}$. The likelihood of the event depends upon how tight market $x'(n+1;\mathbf{s})$ is. Finally, the last two lines include the change in value due to an exogenous shock, whether idiosyncratic or aggregate.

Equation (5) shows that, when the buyer is captive and the seller is subject to size or productivity changes, the customer must internalize how the seller will optimally redesign the contract under the new state. This meaningful forward-looking aspect of demand arises because the seller is committing to its pricing contracts.

Incumbent Sellers

Consider a seller with idiosyncratic productivity $z \in \mathcal{Z}$ who is serving $n \in \mathbb{N}$ customers under the promised value of $x \in \mathcal{X}$. The seller must choose a new contract $\boldsymbol{\omega} = \big\{p, \mathbf{x}'(n';\mathbf{s}')\big\}$ to maximize the life-time value:

$$
\begin{aligned}
rV^S(n,x;\mathbf{s}) = \max_{\boldsymbol{\omega}\in\Omega} \Bigg\{ & pn - \mathcal{C}(n;\mathbf{s}) + \delta_f\Big(V_0^S(\varphi) - V^S(n,x;\mathbf{s})\Big) \\
& + n\delta_c\Big(V^S\big(n-1, x'(n-1;\mathbf{s});\mathbf{s}\big) - V^S(n,x;\mathbf{s})\Big) \\
& + \eta\Big(\theta\big(x'(n+1;\mathbf{s});\varphi\big)\Big)\Big(V^S\big(n+1, x'(n+1;\mathbf{s});\mathbf{s}\big) - V^S(n,x;\mathbf{s})\Big) \\
& + \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)\Big(V^S\big(n, x'(n;z',\varphi); z',\varphi\big) - V^S(n,x;\mathbf{s})\Big) \\
& + \sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi)\Big(V^S\big(n, x'(n;z,\varphi'); z,\varphi'\big) - V^S(n,x;\mathbf{s})\Big)\Bigg\}
\end{aligned} \tag{6}
$$

where $V_0^S(\varphi)$ denotes the value of having no customers (which we derive below).[20] The term $\big[pn - \mathcal{C}(n;\mathbf{s})\big]$ are the seller's flow profits, composed of revenue from selling to n customers, net of operating costs. The next term on the first line is the expected change in value if the seller goes bankrupt, in which case it instantly loses all customers and enters the pool of potential entrants. The third additive term includes the probability that one of the n customers separates, in which case the seller delivers the promised value $x'(n-1;\mathbf{s})$. The second line shows that, by posting a new offer $x'(n+1;\mathbf{s})$, the seller might attract the $(n+1)$-th buyer. When making a new offer, the seller understands the sorting behavior of buyers across states for different promised values through the equilibrium θ schedule. Finally, the value of the firm could change due to an exogenous state transition.[21] When choosing $\boldsymbol{\omega}$, the seller is constrained by:

$$
V^B(n,\boldsymbol{\omega};\mathbf{s}) \geq x \tag{7}
$$

Equation (7) is a *promise-keeping (PK) constraint* guaranteeing that, with its choice of the contract, the seller honors the promises that were made in the past: the value that each buyer of

[20] The object $\Omega \equiv \mathbb{R} \times [\underline{x}, \overline{x}]^{\overline{k}}$ denotes the set of admissible contracts, where $\overline{k} \equiv 3k_z k_\varphi - 1$ is the number of new promises made. For $n = 1$, we note that $x'(n-1;\mathbf{s}) = \emptyset$.

[21] Figure A.2 in the Appendix provides a graphical depiction of all possible seller transitions.

the firm obtains under the contract must be weakly greater than the value x that was promised to her. This prevents customers from separating in equilibrium, rendering long-lasting relationships.

Potential Entrants

Inactive firms 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. Prior to selling the good, they must also draw an initial productivity level z_0 from the π_z distribution. For each possible realization $z_0 \in \mathcal{Z}$, their proposed contract is a promised value to their first customer.

Formally, the ex-ante value of the potential entrant in aggregate state φ solves:

$$rV_0^S(\varphi) = -\kappa(\varphi) + \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) v_0^S(z_0, \varphi) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)\Big(V_0^S(\varphi') - V_0^S(\varphi)\Big) \tag{8}$$

This value is composed of the set-up flow cost κ (first additive term), the expected value of posting a contract under productivity draw z_0 (second term), and the expected change in the ex-ante value of entry for a change in the aggregate state (third term). We have defined the expected value of entry under a draw z_0 by:

$$v_0^S(z_0, \varphi) \equiv \max_{x' \in \mathcal{X}} \; \eta\big(\theta(x'; \varphi)\big)\Big(V^S\big(1, x'; z_0, \varphi\big) - V_0^S(\varphi)\Big) \tag{9}$$

The seller understands how inactive buyers sort across markets, in that the $\theta(\cdot; \varphi)$ schedule is taken as given. There is no PK constraint in this case as the firm does not yet have any customers.

We assume free entry into the product market for the first customer. Since the total mass of sellers adjusts freely, firms will flood into the economy so long as the expected value of posting a contract exceeds the set-up cost $\kappa(\varphi) > 0$. Therefore, in an equilibrium with positive entry in all aggregate states, it must be the case that:

$$\forall \varphi \in \Phi : \qquad V_0^S(\varphi) = 0$$

The free-entry condition thus pins down the average market tightness among single-customer firms in the cross-section of initial productivity draws.

2.3 Optimal Pricing Contract

In this section, we derive and describe the properties of the optimal contract for a typical firm. Our main result is that, since contracts are complete, and sellers and buyers can engage in revenue-neutral transfers schemes, the seller's and the joint surplus problems are equivalent. As we shall see, this simplifies the state space and renders the equilibrium computationally tractable.[22]

Consider a typical firm whose state vector is $(n, x; \mathbf{s})$. As seen in the last section, the optimal contract $\boldsymbol{\omega} = \big\{p, \mathbf{x}'(n'; \mathbf{s}')\big\}$ can be obtained as the solution to the problem of the seller, described

[22] Kaas and Kircher (2015) and Schaal (2017) use similar insights for tractability, though in the context of firm-dynamics search models of the labor market.

in (6). A standard monotonicity argument reveals that sellers will offer the lowest values to their buyers such that the seller's promises are still honored, and so the PK constraint (7) will hold with equality. To economize on notation, for the remainder of the paper we write the value of a buyer of firm $(n, x; \mathbf{s})$ as simply x instead of $V^B(n, \boldsymbol{\omega}; \mathbf{s})$.

Next, define the *joint surplus* in a typical 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 firm:

$$W(n, x; \mathbf{s}) \equiv V^S(n, x; \mathbf{s}) + nx$$

In Appendix B.1 we show that the joint surplus can be written in the following HJB form:

$$\begin{aligned}(r+\delta_f)W(n,x;\mathbf{s}) = \max_{\mathbf{x}'(n';\mathbf{s}')} \Bigg\{ & n\Big(v(\varphi) + (\delta_f+\delta_c)U^B(\varphi)\Big) - \Big(\mathcal{C}(n;\mathbf{s}) + \eta\Big(\theta\big(x'(n+1;\mathbf{s});\varphi\big)\Big)x'(n+1;\mathbf{s})\Big) \\ & + \eta\Big(\theta\big(x'(n+1;\mathbf{s});\varphi\big)\Big)\Big(W\big(n+1, x'(n+1;\mathbf{s});\mathbf{s}\big) - W(n,x;\mathbf{s})\Big) \\ & + n\delta_c\Big(W\big(n-1, x'(n-1;\mathbf{s});\mathbf{s}\big) - W(n,x;\mathbf{s})\Big) \\ & + \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)\Big(W\big(n, x(z',\varphi); z', \varphi\big) - W(n,x;\mathbf{s})\Big) \\ & + \sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi)\Big(W\big(n, x(z,\varphi'); z, \varphi'\big) - W(n,x;\mathbf{s})\Big)\Bigg\} \end{aligned} \tag{10}$$

Intuitively, equation (10) represents the joint surplus as the present discounted value of the buyers' total surplus, net of the seller's total costs. On the first line, the term $n\big(v + (\delta_f + \delta_c)U^B\big)$ represents the aggregate flow surplus for all the n customers of the firm, composed of the sum of the per-customer utility from consumption and the expected per-customer gains from separation. The second line collects (inside the parentheses) the total costs of the match for the seller, equal to total operating costs (first term) plus the expected cost of the new contract (second term). The latter is composed of the product of the promised value, $x'(n+1)$, and the endogenous probability of attraction, $\eta\big(\theta(x'(n+1))\big)$. The remaining terms in equation (10) take into account the usual transitions (i.e. separation, growth, and exogenous shocks). Then, we can show:

Proposition 1 (Joint Surplus Problem)

i. The seller's and the joint surplus problems are equivalent. *In particular:*

(a) If some $\boldsymbol{\omega} = \{p, \mathbf{x}'(n';\mathbf{s}')\}$ solves (6)-(7), then $\mathbf{x}'(n';\mathbf{s}')$ is a solution to (10).

(b) Conversely, for every vector $\mathbf{x}'(n';\mathbf{s}')$ that solves (10), there exists a unique p for which $\{p, \mathbf{x}'(n';\mathbf{s}')\}$ is a solution to (6)-(7).

ii. The joint surplus function $W(n, x; \mathbf{s})$ is constant in x.

For the proof, see Appendix B.1. Part *i.* of Proposition 1 establishes that the contract that maximizes the seller's profits can be found by solving an alternative problem, given by (10). In this problem, the contract maximizes the profits of all the parties involved in a utilitarian fashion, provided that the seller extracts rents from each buyer up to the limit established by promise-keeping. Since the contract space is complete (that is, it specifies continuation promises for each and every possible future state), and both agents have linear preferences, there always exists a menu of price and promised utility that, for any future state, redistributes rents among the seller and its customers in a payoff-maximizing way.

Part *ii.* of the proposition follows immediately from this logic, and clarifies why the joint-surplus problem is much simpler to solve than the seller's problem. Since price and continuation promises map one-for-one, the maximized surplus is invariant to the rent-sharing components of the contract (x and p). Conveniently, this means that the problem can be split in two stages. In *Stage 1*, the firm sets the vector of continuation promises $\mathbf{x}'(n';\mathbf{s}')$ that maximizes the size of the surplus under every possible combination of future states. In *Stage 2*, the price level implements such an allocation by ensuring that PK binds in every single state. Further, the surplus is also constant in the firm's outstanding promise, x, for this is a predetermined state that was chosen optimally previously by the firm. Thus, given $\mathbf{s}$, there exists a sequence $\{W_n(\mathbf{s})\}_{n=1}^{+\infty}$ such that:

$$\forall(n,x)\in\mathbb{N}\times\mathcal{X}: \qquad W_n(\mathbf{s}) = W(n,x;\mathbf{s})$$

As a result, the policy that solves problem (10) is not a function of x, and neither is the optimal price level, simplifying the characterization of the equilibrium. By ex-ante indifference, the option value of matching for the buyer is constant across markets and given by $\Gamma^B(\varphi)$ (equation (4)). Therefore, the tightness of market x is:

$$\theta(x;\varphi) = \mu^{-1}\left(\frac{\Gamma^B(\varphi)}{x-U^B(\varphi)}\right) \tag{11}$$

By Assumption 2.i and continuity of θ on x, equation (10) describes the maximization of a continuous function over a compact support, so there exist promises $\big\{x'(n+1;\mathbf{s}), x'(n-1;\mathbf{s}), \mathbf{x}'(n;\mathbf{s}')\big\}$ and a price level $p(n;\mathbf{s})$ that solve the joint surplus problem.

Stage 1. Continuation promises We begin with the choice of $x'(n+1;\mathbf{s})$. First, by equation (11) and differentiability of η, the following first-order condition is sufficient:[23]

$$\underbrace{\frac{\partial\eta\big(\theta(x';\varphi)\big)}{\partial x'}}_{\substack{\text{Change in}\\ \text{attraction rate}}}\cdot\underbrace{\Big(W_{n+1}(\mathbf{s})-W_n(\mathbf{s})\Big)}_{\substack{\text{Gain in joint surplus}\\ \text{from additional customer}}} = \underbrace{\frac{\partial\eta\big(\theta(x';\varphi)\big)}{\partial x'}\cdot x'}_{\substack{\text{Expected cost of}\\ \text{paying } x' \text{ utils}}} + \underbrace{\eta\big(\theta(x';\varphi)\big)}_{\substack{\text{Price adjustment}\\ \text{by PK}}} \tag{12}$$

This expression summarizes all the relevant dynamic forces at play in a single equation. Intuitively, the optimal continuation value x' equates the expected marginal benefit of upgrading the size of the firm by one customer (left-hand side), to the expected marginal costs of such a transition (right-hand side). On the left-hand side: an increase by one util in the promised value increases the joint surplus in case the seller makes a size transition (second term). These gains must then

[23] Sufficiency obtains because the second-order condition follows from Assumption 2*(iii)* if, in the language of Assumption 2*(iii)*, we pick $h(x)\equiv\frac{\Gamma^B}{x-U^B}$ and $\widehat{x}\equiv W_{n+1}-W_n$.

be weighted by the marginal effect of the raised promised value on the likelihood that the firm meets a new customer (first term). On the right-hand side: for every util spent on the continuation promise, the seller incurs in two associated costs. First, the direct cost of delivering the new value to the additional customer, weighted by the change in the meeting rate. Second, the decrease in the price level of all currently captive buyers of the firms, by $\eta\big(\theta(x'; \varphi)\big)$ utils, which is required by promise-keeping.

The objects $\mathbf{x}'(n; \mathbf{s}')$, $\mathbf{s}' \neq \mathbf{s}$, and $x'(n-1; \mathbf{s})$, cannot be determined by a surplus-maximizing condition similar to (12) because the joint surplus is constant in these objects (Proposition 1, part *ii.*). Instead, these values are purely redistributive: they affect only the way in which the total surplus is split ex-post between buyers and seller. Specifically, the firm's choices for these objects must be consistent with the sorting behavior of inactive buyers (equation (11)). By symmetry, as decisions are history-independent, the optimal *downsizing* choice for a size-n seller must be consistent with the optimal *upsizing* choice for a firm of size $n-2$. Similar considerations apply when transitioning to a different (idiosyncratic or aggregate) exogenous state. Finally, for $n = 1$, the free entry condition must be satisfied, implying:

$$\kappa(\varphi) = \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \eta\Big(\theta\big(x'(1; z_0, \varphi); \varphi\big)\Big)\Big(W_1(z_0, \varphi) - x'(1; z_0, \varphi)\Big) \tag{13}$$

Summing up, the set of equilibrium markets is $\mathcal{X}^* \equiv \big\{x'(n; z, \varphi) : (n, z, \varphi) \in \mathbb{N} \times \mathcal{Z} \times \Phi\big\}$, where $x'(n; z, \varphi)$ solves (12) for each $n \geq 2$, and (13) for $n = 1$. Market tightness levels, $\theta_n(z, \varphi) \equiv \theta\big(x'(n; z, \varphi); \varphi\big)$, are then found via equation (11). Since $\theta(x; \varphi)$ is an increasing and continuous function of x, θ_n inherits the size-dependence of x'. In turn, θ_n determines firm growth rates in equilibrium. Formally, the law of motion for firm size is:

$$n_{t+\Delta} - n_t = \begin{cases} 1 & \text{w/prob. } \eta\big(\theta_{n_t+1}(z, \varphi)\big)\Delta + o(\Delta) \\ -1 & \text{w/prob. } n_t \delta_c \Delta + o(\Delta) \\ -n_t & \text{w/prob. } \delta_f \Delta + o(\Delta) \\ 0 & \text{else} \end{cases} \tag{14}$$

for a small time lapse of length $\Delta > 0$, where $\lim_{\Delta \searrow 0} \frac{o(\Delta)}{\Delta} = 0$. For instance, when x' is decreasing in n, smaller firms attract more buyers per unit of time by offering higher ex-post values, so the buyer-to-seller ratio is higher in those markets, and these firms grow relatively faster compared to other firms. Figure A.1 in the Appendix depicts the different markets in equilibrium for this case. All equilibrium markets are distributed on the θ schedule defined by buyer's ex-ante revenue equalization. To grow, the seller makes a state-contingent promise that is strictly below the current valuation of her buyers (depicted on the horizontal axis), but always above U^B to prevent voluntary separating. In equilibrium, the resulting collection of markets makes inactive buyers indifferent.

Stage 2. Prices Finally, the equilibrium price can be backed out of the PK constraint, which is binding. First, we replace $V^B(n, \boldsymbol{\omega}; z, \varphi) = x'(n; z, \varphi)$ and the components of $\boldsymbol{\omega}$ in equation (5). Then, solving for p in the resulting equation, we obtain:

$$\begin{aligned} p_n(z,\varphi) = &\underbrace{v(\varphi) - rx'(n;z,\varphi)}_{\text{Baseline component}} + \underbrace{\delta_f\Big(U^B(\varphi) - x'(n;z,\varphi)\Big)}_{\text{Exit component}} + \underbrace{\eta\big(\theta_{n+1}(z,\varphi)\big)\Big(x'(n+1;z,\varphi) - x'(n;z,\varphi)\Big)}_{\text{Growth component}} \\ &+ \underbrace{n\delta_c\left(\frac{U^B(\varphi) + (n-1)x'(n-1;z,\varphi)}{n} - x'(n;z,\varphi)\right)}_{\text{Separation component}} \\ &+ \underbrace{\sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)\Big(x'(n;z',\varphi) - x'(n;z,\varphi)\Big)}_{\text{Idiosyncratic-shock component}} + \underbrace{\sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi)\Big(x'(n;z,\varphi') - x'(n;z,\varphi)\Big)}_{\text{Aggregate-shock component}} \end{aligned} \tag{15}$$

The optimal price level can be decomposed into the following additive parts. The first one is the price level that would prevail if, in the absence of any exogenous shock, each customer were to stay matched forever with its seller and size did not change going forward.[24] As this component is driven by fundamentals, we call it the *baseline price level*. The remaining terms in (15) introduce the necessary adjustments for possible changes in firm states. To provide intuition, consider again the parametrization under which x' decreases in n (as in Figure A.1). Over and above the baseline price, the firm first offers a price reduction of $\delta_f\big(U^B - x'(n)\big) \leq 0$ to compensate the customer for the expected loss in value in the event that the firm exits the market. We label this the *exit component*. Second, the term $\eta(\theta_{n+1})\big(x'(n+1) - x'(n)\big) \leq 0$ is a compensation for the eventuality that the firm grows. This compensation is thus labeled the *growth component*. Third, the firm adjusts the price for the event of customer separation: a reduction in size lowers the seller's value and has a pecuniary externality on all those customers that remain matched, so the price must again be adjusted to remain compatible with the seller's commitment. We call this term the *separation component*. If a separation occurs, then the separating customer obtains U^B, whereas the remaining $n-1$ customers each obtain $x'(n-1)$. This amounts to an average value of $\frac{U^B+(n-1)x'(n-1)}{n}$ per customer. Finally, the last two terms in equation (15) adjust the price level for expected changes in the exogenous states. In Section 3 we will discuss the different price effects that may be present in equilibrium and provide some intuition for the direction of the dependence on size.

2.4 Distribution Dynamics

To close the equilibrium, we need to describe the dynamics of the distribution of agents. The equilibrium of the economy described above features heterogeneous agents making forward-looking decisions and sorting into distinct product markets in the presence of both idiosyncratic and aggregate shocks. The distribution of agents across markets in turn depends on the aggregation of such decisions. Yet, our equilibrium characterization has been silent on the composition of buyers and

[24] Indeed, in that case we would have $p = v - rx$, that is $x = \int_0^{+\infty} e^{-rt}(v-p)\mathrm{d}t$, the PDV of perpetually obtaining the fixed surplus $(v-p)$.

sellers across market segments, or the evolution of this distribution. This is because the equilibrium is *block-recursive*.

In our model, block recursivity arises from two key ingredients. First, we assume that search is directed, and thus sellers' offers are not contingent on the identity, and in particular the outside option, of the buyer. As a result, market tightness, which embodies agents' distributions, serves as a sufficient statistic for both sellers and buyers when making decisions, and allows them to make decisions without having to forecast the evolution of aggregates over future states of the economy. Second, the ex-ante revenue-equalization condition across all markets among inactive buyers (equation (3)) implies that the equilibrium tightness on each market adjusts to be consistent with agents' beliefs.[25] Because market tightness is a sufficient statistic to evaluate payoffs in this economy, the model allows for the description of the distribution's transitional dynamics by means of simple flow equations (below).

Let $S_{n,t}(z) \geq 0$ be the total measure of firms of size n with idiosyncratic productivity $z \in \mathcal{Z}$ at time $t \geq 0$. Recall that all such firms are seeking new customers in market $x'(n+1; z, \varphi)$. Let $B^I_{n+1}(z,\varphi)$ be the measure of (inactive and searching) buyers within market $x'(n+1; z, \varphi)$. Then, market tightness must adjust so that:

$$B^I_{n+1,t}(z,\varphi) = \theta_{n+1}(z,\varphi) S_{n,t}(z) \tag{16}$$

at every $t \geq 0$ for all $n \in \mathbb{N}$. Using that $\eta(\theta) = \theta\mu(\theta)$, equation (16) can alternatively be written as $\mu\big(\theta_n(z,\varphi)\big) B^I_{n,t}(z,\varphi) = \eta\big(\theta_n(z,\varphi)\big) S_{n-1,t}(z)$, stating that the measure of inactive buyers who become customers of a (n,z)-type firm is equal to the measure of sellers of productivity z and size $n-1$ who acquire their n-th customer.

Similarly, let $B^A_{n,t}(z)$ be the measure of customers that are matched with firms of type (n,z) at time t. By construction, we have:

$$B^A_{n,t}(z) = n S_{n,t}(z) \tag{17}$$

at any $t \geq 0$. The measures of inactive and active buyers must add up to the total mass of buyers in the economy at all times, and thus:

$$\forall \varphi \in \Phi,\ \forall t \geq 0: \quad \underbrace{\sum_{n=1}^{+\infty} \sum_{z\in\mathcal{Z}} B^A_{n,t}(z)}_{=B^A_t} + \underbrace{\sum_{n=1}^{+\infty} \sum_{z\in\mathcal{Z}} B^I_{n,t}(z,\varphi)}_{=B^I_t} = 1 \tag{18}$$

This equation establishes an aggregate feasibility constraint: the unit mass of buyers must be either matched with a firm and consuming, or looking for a seller.

Market tightness jumps instantaneously in response to an aggregate shock.[26] This is because the

[25] Kaas and Kircher (2015) exploit similar insights to obtain tractability. An alternative approach would be to assume free entry of firms across all product markets, as in Menzio and Shi (2010, 2011) and Schaal (2017).
[26] Our notation convention is the following: (i) explicit dependence on φ denotes that a variable jumps with φ; (ii) a subscript t denotes time-dependence for fixed φ.

mass of inactive buyers adjusts immediately to guarantee that the indifference condition among unmatched buyers (equation (3)) is met in all states of nature. However, by block recursivity, tightness remains constant along each aggregate state. The mass of potential entrants, denoted $S_{0,t}(\varphi)$, jumps following a φ-shock, and otherwise evolves smoothly due to sellers flowing in and out of inactivity in the transition. The distribution of firms, $\{S_{n,t}(z)\}$, 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 $n = 1$ and some $z \in \mathcal{Z}$ are:

$$\begin{aligned}\boldsymbol{\partial}_t S_{1,t}(z) \quad &= \quad \pi_z(z)\eta\big(\theta_1(z,\varphi)\big)S_{0,t}(\varphi) + 2\delta_c S_{2,t}(z) + \sum_{\tilde{z}\neq z}\lambda_z(z|\tilde{z})S_{1,t}(\tilde{z}) \\ &\quad - \Big(\delta_f + \delta_c + \eta\big(\theta_2(z,\varphi)\big) + \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)\Big)S_{1,t}(z) \end{aligned} \tag{19}$$

where $\boldsymbol{\partial}_t$ is the partial derivative operator with respect to time. Inflows (first line) are given by successful entrants drawing productivity z upon entry, and by the share of incumbents that are either losing a customer, or transitioning into z from some $\tilde{z} \neq z$. Outflows (second line) are given by firms that either die, lose their only customer, gain a second customer, or transition to a distinct productivity state, $\tilde{z} \neq z$. The aggregate state φ enters the law of motion implicitly through its influence on the jump dynamics of S_0 and θ_1. For $n \geq 2$:

$$\begin{aligned}\boldsymbol{\partial}_t S_{n,t}(z) \quad &= \quad \eta\big(\theta_n(z,\varphi)\big)S_{n-1,t}(z) + (n+1)\delta_c S_{n+1,t}(z) + \sum_{\tilde{z}\neq z}\lambda_z(z|\tilde{z})S_{n,t}(\tilde{z}) \\ &\quad - \Big(\delta_f + n\delta_c + \eta\big(\theta_{n+1}(z,\varphi)\big) + \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)\Big)S_{n,t}(z) \end{aligned} \tag{20}$$

The interpretation of equation (20) is almost identical to that of the previous equation. Finally, the measure of potential entrants, $S_{0,t}(\varphi)$, obeys the following ODE:

$$\boldsymbol{\partial}_t S_{0,t}(\varphi) = \delta_f \mathcal{S}_t + \delta_c \sum_{z\in\mathcal{Z}} S_{1,t}(z) - \sum_{z_0\in\mathcal{Z}} \pi_z(z_0)\eta\big(\theta_{1,t}(z_0,\varphi)\big)S_{0,t}(\varphi) \tag{21}$$

where $\mathcal{S}_t \equiv \sum_{n=1}^{+\infty}\sum_{z\in\mathcal{Z}} S_{n,t}(z)$ is the total measure of *incumbent* firms (i.e. firms with one or more customers). The usual interpretation applies, with the particularity that entering firms must now draw an initial productivity level from the π_z distribution.

Equations (19)-(21) offer a full characterization of the model's dynamics. We may now equate flows in and out of every possible state to obtain the *stationary distribution* of firms: $\boldsymbol{\partial}_t S_{n,t}(z) = 0$, $\forall n, z$.[27] As the last step to close the equilibrium, we must find the aggregate measures of agents in the stationary solution: $\big(\mathcal{S}, S_0, B^A, B^I\big)$. For the derivations, see Appendix G.1.

[27] A solution for the stationary distribution may not exist unless (i) there is an upper-bound on seller size, or (ii) there are no (z,φ)-shocks and $\delta_f = 0$. For (i), see Proposition 4. For (ii), see Appendix G.2.

2.5 Equilibrium Definition and Efficiency

We are now ready to define an equilibrium:

Definition 1 *A Recursive Equilibrium is, for each aggregate state $\varphi \in \Phi$ and time $t \in \mathbb{R}_+$, a set of value functions $V^S(\cdot, \varphi) : \mathbb{N} \times \mathcal{X} \times \mathcal{Z} \to \mathbb{R}$ and $V^B(\cdot, \varphi) : \mathbb{N} \times \Omega \times \mathcal{Z} \to \mathbb{R}$; a value of inactivity $U^B(\varphi) \in \mathbb{R}$; joint surplus and prices $\{W_n(z, \varphi), p_n(z, \varphi) : (n, z) \in \mathbb{N} \times \mathcal{Z}\}$, and continuation promises $\{x'(n+1; z, \varphi), x'(n-1; z, \varphi), \{x'(n; z', \varphi) : z' \in \mathcal{Z}\}, \{x'(n; z, \varphi') : \varphi' \in \Phi\}\}$; a decision rule $\widehat{x}(\varphi)$ for inactive buyers; a market tightness function $\theta(\cdot, \varphi) : \mathcal{X} \to \mathbb{R}_+$; aggregate measures of agents $\{S_{0,t}(\varphi), \mathcal{S}_t, B_t^A, B_t^I\}$; and a distribution of sellers and buyers $\{S_{n,t}(z), B_{n,t}^A(z), B_{n,t}^I(z, \varphi) : (n, z) \in \mathbb{N} \times \mathcal{Z}\}$; such that: [i] the value functions solve (5) and (6), $U^B(\varphi)$ satisfies the free-entry condition (13), and the joint surplus $W_n(z, \varphi)$ solves (10); [ii] continuation promises and prices satisfy (12) and (15), respectively; [iii] $\widehat{x}(\varphi)$ solves the inactive buyer's problem, (1)-(2); [iv] market tightness $\theta(x; \varphi)$ is consistent with the sorting behavior of inactive buyers, (3); and [v] aggregates and the distribution of agents satisfy the flow equations in Section 2.4.*

Proposition 2 below states that, given the search frictions, a Recursive Equilibrium is efficient. In our environment, the planner chooses distributions of buyers and sellers, as well as market tightness levels, in order to maximize aggregate welfare:

$$\mathbb{E}_0 \left\{ \int_0^{+\infty} e^{-rt} \mathbb{W}_t(\varphi_t) \mathrm{d}t \right\} \tag{22}$$

where

$$\mathbb{W}_t(\varphi_t) \equiv -\kappa(\varphi_t) S_{0,t}(\varphi_t) + \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} \Big(v(\varphi_t) B_{n_t,t}^A(z_t) - \mathcal{C}(n_t; z_t, \varphi_t) S_{n_t,t}(z_t) - c(\varphi_t) B_{n_t,t}^I(z_t, \varphi_t) \Big)$$

subject to the cross-sectional and dynamic properties of the distribution of agents described in Section 2.4. Aggregate welfare in this economy equals the present discounted sum of consumption gains by active buyers, net of search costs by inactive buyers, and production and entry costs by firms. Using this definition, we then establish:

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

For the proof, see Appendix B.2. This result implies that, given the search frictions, our model features efficient firm dynamics and efficient pricing behavior. In particular, markup dispersion is necessary to optimally distribute trade gains among buyers and sellers, as prices in our environment serve to efficiently direct buyer search toward specific product markets. The result is in contrast to models explaining dispersion in firm-level revenue through (inefficient) resource misallocation (e.g. Hsieh and Klenow (2009)). While we do not rule out other interpretations, our setting demonstrates that this type of dispersion may also be generated through efficient pricing in the context of frictional product market.

3 Discussion

This section presents the qualitative properties of the equilibrium, with an emphasis on how product market frictions lead firms of different sizes to set different combinations of price and promised utilities, and to experience different subsequent growth paths. Then, we comment on the key assumptions and describe the role that they play in the model.

Qualitative Features

We begin with a useful result. Under a standard matching function, the model admits a partial analytical representation in the following sense: for each given value of inactivity U^B, the equilibrium joint surplus W_{n+1} can be found as the solution of a *second-order difference* equation, i.e. as a function of n, W_n, W_{n-1}, and parameters. Assume $\mu(\theta) = \theta^{\gamma-1}$, with $\eta(\theta) = \theta\mu(\theta)$, where $\gamma \in (0,1)$ is the matching elasticity.[28] Then:

Proposition 3 (Solution of the Joint Surplus) *For each $(z,\varphi) \in \mathcal{Z} \times \Phi$:*

(a) The joint surplus $W_n(z,\varphi)$ solves a second-order difference equation in n:

$$W_{n+1}(z,\varphi) = W_n(z,\varphi) + U^B(\varphi) + \left(\frac{\Gamma^B(\varphi)}{\gamma}\right)^{\gamma} \left(\frac{\Gamma^S_n(z,\varphi)}{1-\gamma}\right)^{1-\gamma} \tag{23}$$

where $\Gamma^S_n(z,\varphi)$ is a function of n, W_n, W_{n-1}, and parameters (in Appendix B.3).

(b) Promised utility equals: $x_{n+1}(z,\varphi) = \gamma\Big(W_{n+1}(z,\varphi) - W_n(z,\varphi)\Big) + (1-\gamma)U^B(\varphi)$.

For the proof, see Appendix B.3. Proposition 3 shows that, in spite of the rich dynamics of the model, the joint surplus can be expressed in a very tractable form. The formulas have intuitive interpretations. Let us suppress the exogenous state dependence everywhere to alleviate notation and ease the intuition. First, in the Appendix we show that:

$$\Gamma^B = \mu(\theta_{n+1}) \underbrace{\Big(x_{n+1} - U^B\Big)}_{\substack{\text{Ex-post} \\ \text{buyer net gains} \\ \text{from new match}}} \qquad \text{and} \qquad \Gamma^S_n = \eta(\theta_{n+1}) \underbrace{\Big(W_{n+1} - W_n - x_{n+1}\Big)}_{\substack{\text{Ex-post} \\ \text{seller net gains} \\ \text{from new match}}}$$

For each equation, the right hand-side of the equality gives the product of the *ex-post* net gain from a *new* match, times the probability that a match occurs from the perspective of buyer and seller, respectively. Thus, in equilibrium Γ^B and Γ^S_n are the *ex-ante* net gains from matching for the buyer and the seller, respectively. Γ^B is constant in n (and z) because of ex-ante buyer indifference (equation (3)). Γ^S_n depends on the seller's size: the seller extracts 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.

[28] This arises from a Cobb-Doublas matching function, $M(B,S) = B^{\gamma}S^{1-\gamma}$. However, similar results hold for the more general CES function $\mu(\theta) = (1+\theta^{\xi})^{-1/\xi}$, of which Cobb-Douglas is a special case when $\xi \to 0$.

Therefore, 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) states that the matching elasticity parameter γ governs how the surplus is shared *ex-post* between the seller and the new customer. The customer's ex-post net gains are given by:

$$x_{n+1} - U^B = \gamma\Big(W_{n+1} - W_n - U^B\Big)$$

showing that a fraction γ of the total gains in joint surplus are absorbed by the new incoming buyer. Interestingly, the ex-post gains for the seller must incorporate rents shared with the new buyer, as well as all those shared with the pre-existing buyers:

$$\underbrace{V_{n+1}^S - V_n^S}_{\substack{\text{Seller's } \textit{overall} \text{ ex-post} \\ \text{net gain from growing}}} = \underbrace{(1-\gamma)\Big(W_{n+1} - W_n - U^B\Big)}_{\substack{\text{[A] Surplus extracted} \\ \text{directly out of } \textit{new} \text{ customer}}} + \underbrace{n\Big(x_n - x_{n+1}\Big)}_{\substack{\text{[B] Surplus transferred between} \\ \text{seller and } \textit{pre-existing} \text{ customers}}} \tag{24}$$

In words, after a $n \to (n+1)$ transition, the seller absorbs a share $(1-\gamma)$ of the total net gain in joint surplus from the *new* customer (term [A]). In addition, some surplus is transferred between the seller and the n pre-existing customers (term [B]).

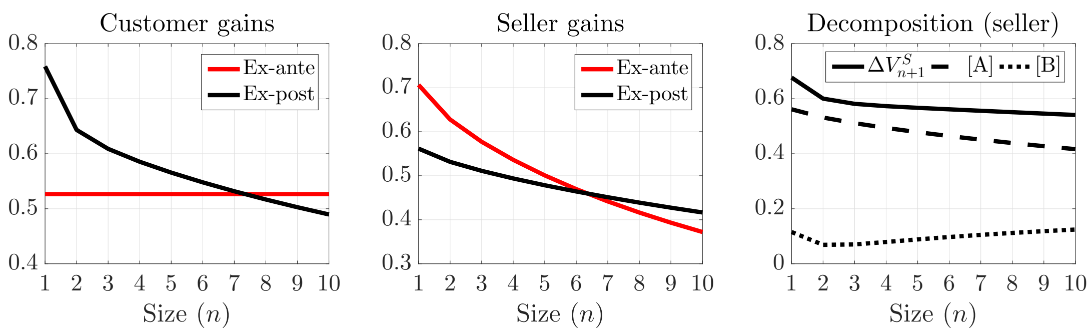

Figure 1: Ex-ante and ex-post gains for the new buyer and the seller (left and middle panels), and decomposition of seller gains, across seller size n (right panel), for a numerical example. The red lines are Γ^B and Γ^S_n. The black lines are $(x_{n+1} - U^B)$ and $(W_{n+1} - W_n - x_{n+1})$. The right panel is the decomposition in equation (24).

For example, when buyer valuation is decreasing in n (i.e. $x_n > x_{n+1}$), term [B] is a positive transfer from all n pre-existing customers to their seller. Figure 1 shows, in this case, how net gains for buyers and sellers change with size (from both ex-ante and ex-post perspectives). We see in the right-most panel that, as the seller grows, it increasingly prefers to extract rents from the current customer base (term [B] increases), and is less concerned with extracting surplus from the new consumer (term [A] decreases).

Finally, we describe the relationship between promises (x_n) and sizes (n) in equilibrium. To emphasize that size-dependence may arise even with constant returns to scale, we assume in this example that marginal costs are constant in n (i.e. $\mathcal{C}(n) \propto n$). Figure 2 shows a numerical example in which promised utilities are strictly decreasing with size. This will be the case in our estimated model. The reason for this relation is that the seller needs to raise resources from future customers

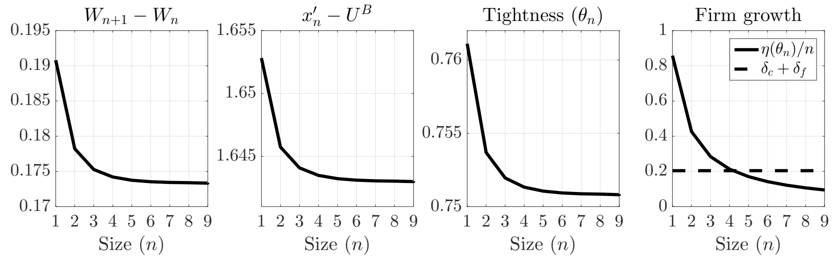

Figure 2: Change in joint surplus, net customer ex-post gains, market tightness, and firm growth, as a function of size, for the case with constant marginal cost (i.e. $\mathcal{C}(n) \propto n$). Firm growth (equation (14)) has been decomposed into the rate of customer attraction (solid line) and customer attrition (dashed line). The intersection of these two lines marks the stationary size. On the left (right) of this point, firms are growing (shrinking) on average.

quickly in order to overcome the initial market penetration cost, κ. At the same time, it must entice customers to remain matched. Thus, it lowers the promise as more customers join, but always so that $x_n > U^B$, $\forall n$. As a result, market tightness is higher for smaller firms, and these firms experience faster growth on average.

In contrast, when the costs to market penetration are sufficiently low, the need to raise resources from the future is weaker. In this case, the profile of continuation values may *increase*, and the price *fall*, as the seller grows (for an example, see Figure A.3 in Appendix A). In sum, the model can generate different price-size correlations, depending on parameters. However, in both cases, we obtain size-dependent firm growth rates (right-most panel of Figures 2 and A.3), even though these examples feature constant returns to scale in technology.

Discussion of the Model's Main Assumptions

Our model is somewhat stylized, particularly on the contractual environment. Next, we comment on the most relevant assumptions:

Price discrimination Perhaps most importantly, we have assumed that there can be no price dispersion *within* firms, in that sellers cannot price discriminate across customers of different tenures. While this assumption is realistic for many sectors of the economy, especially those in which sellers face a large number of potential buyers (such as retail, our application in the quantitative part of the paper), it may not be apt for some cases. An example are industries in which personalized buyer-seller relationships may explicitly develop (newspapers, cell phone and Internet services, commercial banks, etc.). Gourio and Rudanko (2014b) propose a model for such type of relationships, whereby sellers attract buyers by offering price discounts below valuation, and then charging the valuation thereafter. Though the focus of our paper is altogether different, it is still worth clarifying that our pricing policy does *not* collapse to this pricing strategy when we allow prices to be dispersed within firms. In Appendix C, we extend the baseline model to allow for price discrimination. We show that the equilibrium can still be solved for with a joint surplus problem (Proposition C.1), and that the model still yields *unique* predictions for firm dynamics (Corollary C.1). However, this version of the model also gives rise to a new feature: equilibrium price indeterminacy (Proposition C.2).

Endogenous separations Another important assumption is that customers cannot bypass the costly inactive state when they separate (either voluntarily or due to a shock) from their seller. Allowing for endogenous seller-to-seller transitions would incorporate an additional dimension into the firm's pricing decisions. Besides the dynamic rent-extraction trade-off between incoming customers and the current base, the firm would now have to solve an attraction-attrition trade-off: a more ex-post profitable contract for inactive buyers may enhance the chance of a customer match, but also increase the likelihood of a voluntary separation. We propose how to endogenize this margin in Appendix F.2, and discuss the technical challenges it presents.

Commitment We have also assumed perfect commitment on the seller's side. Intuitively, long-term contracts are a stand-in for a reputational concern on the side of the firm. By promising to deliver a utility level, the seller can balance the price with the continuation value to lure customers into remaining matched. In turn, customers understand that the firm will not price gouge ex-post, and they remain loyal to their seller in order to avoid going through the costly inactive state. If we were to dispense of the commitment assumption, we would lose block recursivity, and thereby the attractive analytical features of the equilibrium. The reason for this is that, due to a time-inconsistency problem, firms could engage in various forms of pricing strategies, all of which could be sustained under "implicit contracts", paired with trigger strategies on the buyer side (e.g. Nakamura and Steinsson (2011)). These type of contracts might exhibit history-dependence, which in our framework would break the recursive structure. Further, sellers would need to keep track of the buyer distribution, as ex-ante revenue equalization would no longer hold. For these reasons, seller's commitment is a key aspect for our set-up.

Buyer valuation Finally, we have assumed that consumers are homogeneous in their valuation of the good, which we have treated as a parameter for each aggregate shock, $v(\varphi)$. In Appendix F.1 we present a micro-foundation for marginal utility, based on more general CES preferences over the continuum of products sold by firms.

4 Quantitative Analysis

4.1 Data

To estimate the model, we use micro-pricing data on the U.S. retail sector. Although the model is general, the retail sector adheres well with its basic features. In this interpretation of the theory, sellers are stores, and buyers are private consumers. We use data of unique granularity which allows us to focus on narrowly-defined (physically) homogenous products sold by sellers of different sizes within the same market, in accord with the environment of the model. The types of goods in the data are non-durable consumption products that, as in the model, are likely to engage customers and sellers into repeated purchases. The fact that retail stores face a potentially large number of customers implies that the customer anonymity assumption likely provides a good approximation for the bulk of the observed store transactions. Indeed, the majority of supermarket prices are not likely to target specific consumers.[29] Finally, under this framing of the model, we interpret the buyer's search cost as a proxy for the transport, information, and/or utility costs associated to finding and/or switching away from trusted suppliers.

We use micro-level data from the Information Resources, Inc. (IRI) scanner data set (2001-2007).[30] We use weekly average prices for products sold over 5,000 U.S. retail (drug and grocery) stores on 50 different (geographic) markets defined at the MSA level. Products are identified by their Universal Product Code (UPC).[31] The details of data selection, goods considered, and variable construction can be found in Appendix D. To capture the degree to which the same good is sold at different prices by different stores, we follow the literature (e.g. Kaplan *et al.* (2018)) and focus on *relative* prices. The relative price of good u in store s of market m and week t is the log-deviation from the average price of such good across stores:

$$\widehat{p}_{usm,t} = \log P_{usm,t} - \frac{1}{N_{um,t}} \sum_{s=1}^{N_{um,t}} \log P_{usm,t} \tag{25}$$

where $N_{um,t}$ is the number of stores selling the good in that market and week. To measure price dispersion in the data we take the average of the standard deviation of $\widehat{p}_{usm,t}$ across products, stores, markets, and time. In our selected sub-sample, dispersion is high (10.55%), in line with previous studies (e.g. Kaplan and Menzio (2015)).

4.2 Estimation

To estimate the model, the first step is to parametrize the cost function of firms and establish the structure of the exogenous shocks. For the former, we choose:

$$\mathcal{C}(n; z, \varphi) = \tilde{c}(z, \varphi) \cdot n^{\psi} \tag{26}$$

where $\tilde{c}(z, \varphi) > 0$ is a size-invariant scale parameter, and $\psi \geq 1$ is a curvature parameter controlling for the degree of returns to scale.[32] When marginal costs are increasing in size ($\psi > 1$), there is a natural upper bound on seller size for each state, given by $n^*(z, \varphi) \equiv \big(\psi\tilde{c}(z, \varphi)/v\big)^{\frac{1}{1-\psi}}$, beyond which the static flow surplus $\pi_n(z, \varphi) \equiv nv - \mathcal{C}(n; z, \varphi)$ is strictly decreasing and the seller does not want to grow further. The scale parameter in the cost function changes across sellers and aggregate states, with $\tilde{c}(z, \varphi) = we^{z+\varphi}$, where $w > 0$ controls the optimal scale. This specification of shocks is isomorphic to multiplicative idiosyncratic and aggregate TFP shocks in the production

[29] A possible exception are promotions, which we filter out of the data.
[30] The data are available for request at `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), and Paciello *et al.* (2017).
[31] The UPC is an array of numerical digits that is uniquely assigned to a given item. The description of products is very detailed, including information about the brand, flavor, and several packaging attributes.
[32] As seen in Section 3, curvature in the cost function is not necessary for the existence of price dispersion in the model. However, the parameter ψ will help us pin down the size dependence in prices.

function, standard in the literature (e.g. Kaas and Kircher (2015)).

Finally, we must parametrize the states. The size grid is $\mathcal{N} \equiv \{1, 2, \ldots, \overline{n}\}$, for some upper bound $\overline{n} < +\infty$ that is large enough.[33] One can then show that the model's dynamics are convergent:

Proposition 4 (Stability) *The dynamical system represented by the flow equations (19)-(20)-(21) over the grid $(n, z) \in \mathcal{N} \times \mathcal{Z}$ is stable, and converges to an invariant distribution for each aggregate state $\varphi \in \Phi$. The invariant distribution can be analytically characterized.*

For the proof, see Appendix B.4. Further, we must specify the structure of the exogenous shocks, (z, φ). In principle, we should estimate $k_s(k_s - 1)$ transition rates for each shock $s \in \{z, \varphi\}$, a potentially large number. To reduce dimensionality, in practice we assume that each shock follows an Ornstein-Uhlenbeck process in logs:[34]

$$\mathrm{d}\log z_t = -\rho_z \log z_t \mathrm{d}t + \sigma_z \mathrm{d}B_t^z$$
$$\mathrm{d}\log \varphi_t = -\rho_\varphi \log \varphi_t \mathrm{d}t + \sigma_\varphi \mathrm{d}B_t^\varphi$$

where (B_t^z, B_t^φ) are standard Brownian motions, and $(\rho_z, \rho_\varphi, \sigma_z, \sigma_\varphi) > 0$ are parameters. This reduces the estimation to only two parameters per shock: a persistence ρ, and a volatility σ. For details of this estimation procedure, see Appendix E.1.

Calibration Strategy

We seek to match aggregate moments related to store dynamics in the U.S. retail sector as well as average moments across all years of our sample of micro-pricing data.

The model is parsimonious, with 11 free parameters to be identified. Of these, 9 are deep parameters: $(v, r, \delta_f, \delta_c, w, \psi, \kappa, \gamma, c)$, corresponding to the value of consumption, the time discount rate, the separation rates of sellers and consumers, the scale and curvature parameters of the operating cost function, the entry cost for new sellers, the matching elasticity, and the search cost for inactive buyers, respectively. In addition, we must set values for the persistence and dispersion parameters of the exogenous productivity state process: (ρ_z, σ_z). We assume that z can take up to $k_z = 25$ different values, and normalize its mean to unity. We do not estimate the aggregate shocks φ because the spirit of our calibration exercise is to estimate an economy in its long-run equilibrium. These will be re-introduced in Section 5, when we study the response of markups to aggregate supply and demand shocks.

External identification The parameters (v, r, δ_c) 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

[33] In particular, this upper bound is never binding in the sense that $\overline{n} \gg n^*(\underline{z}, \underline{\varphi})$. In our estimation, we take $\overline{n} = 50$.
[34] Ornstein-Uhlenbeck processes are continuous-time Markov chains that can be viewed as the analogue of AR(1) processes in continuous time.

95% annually. The exogenous separation probability is set to match a 0.044% weekly customer turnover rate (corresponding to $\delta_c \approx 0.2041$ at our yearly frequency), which implies that customer relationships last a bit more than 4 and a half years on average. We borrow this value from Paciello *et al.* (2017), who estimated it on the same IRI data we use here. The number falls within the range of values reported by Gourio and Rudanko (2014b) (between 10% and 25% annually).

Internal identification Because of the high non-linearity of the model, identifying each of the remaining parameters separately is unfeasible. However, we can provide some intuition for how each one is informative about specific moments. We estimate the parameters jointly by matching a combination of aggregate and seller-level moments via simulated method of moments (SMM). To implement this procedure, we use an algorithm that randomly searches in the parameter space, and then employs an unweighted minimum-distance criterion function that compares empirical moments to model-implied moments from both the stationary solution as well as simulated data.

For the stationary solution, we solve a fixed-point algorithm that uses value function iteration on W and a bisection procedure to solve for the value of inactivity, U^B. Appendix E.2 outlines the details of this method.[35] To obtain moments from simulated data, we generate histories for many distinct sellers over $T = 100$ years of data which we discretize with time steps of equidistant length $\Delta = 0.01$ each. All sellers are drawn from the stationary distribution at time $t = 0$ and evolve endogenously through simulated Markov chains that replicate the dynamics described in Section 2.4. We drop the first half of the time series when computing average simulated moments to make sure we draw from the stationary distribution. For the entrants' productivity distribution π_z, we use the ergodic distribution that is implied by the calibrated Markov chain for z.

The set of targeted moments can be grouped into two categories: (i) aggregate moments, and (ii) store-level moments related to the distribution of sales and prices. At the aggregate level, we target an average annual entry rate of 8.9%, which we compute as the average ratio of stores aged 52 weeks or less to the total number of existing stores within each year. We define the entry rate in the model as the ratio of *actual* entrants to the total mass of incumbents. The exit rate is the measure of sellers who either die or lose their last remaining customer. By equation (21):

$$EntryRate = \frac{S_0}{\sum_{n,z} S_n(z)} \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0)\eta\big(\theta_1(z_0)\big); \quad ExitRate = \delta_f + \delta_c \frac{\sum_z S_1(z)}{\sum_{n,z} S_n(z)} \tag{27}$$

These formulas hold in and out of steady state, but are equal to each other in the absence of φ-shocks, so the entry rate in the data helps us identify the exogenous exit rate δ_f in the model. At the aggregate level, we also target the cross-sectional average markup. 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 sample. Using firm-level data, typical estimates range from about 10% to as much as 50% or more (e.g. DeLoecker and Warzynski (2012), DeLoecker *et al.* (2016), Christopoulou and Vermeulen (2008)). We choose to target a markup of 39%, a number that we

[35] Further, in Appendix G.3 we show that, for a given U^B, the value function exists and is unique.

impute from the average ratio over our sample period (2001-2007) of gross margins to sales in the retail sector.[36] To be consistent with the empirical target, in the model we compute *measured* markups as the sales-weighted average of the ratio of price to marginal cost:

$$\overline{m} = \sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} m_n(z); \qquad \text{with } m_n(z) \equiv s_n(z)\frac{p_n(z)}{mc_n(z)} \tag{28}$$

where $s_n(z) = \frac{np_n(z)}{\sum_{n,z} np_n(z)}$ is the sales share of type (n, z) firms, and $mc_n(z) \equiv \mathcal{C}(n; z)-\mathcal{C}(n-1; z)$ is the marginal cost of this type of firms. Though many parameters affect the average markup, γ is the most relevant one, as it governs how the gains from trade are shared between the customers and their seller.

At the store level, we target moments from the distribution of prices and sales that we obtain from our IRI sample. The cost parameters (ψ, w) determine firm profitability across sizes, so they play an important role in determining the degree by which firms of similar productivity choose to set different prices for the same product. We choose to target two moments that relate to this dimension of heterogeneity. First, we target price dispersion (10.55% in the data). Second, we target the inter-decile range in the distribution of relative prices between the tenth percentile and the median relative price, equal to 1.1215 (see Table D.2 in the Appendix). We target this measure of left-tail dispersion because the size distribution is right-skewed, so targeting the lower end of the distribution ensures that we capture the pricing behavior of the bulk of population of sellers.

Next, we need to discipline the parameters of the exogenous productivity process, z. We target the yearly autocorrelation in normalized store-level sales (pinning down the persistence ρ_z), and the dispersion in the distribution of normalized sales (pinning down the volatility σ_z). Finally, we need to calibrate the search cost for buyers, c, and the market penetration cost for sellers, κ. As we discussed in Section 3, these parameters are important to pin down the dependence between seller size and seller price, which determines two key aspects of firm dynamics: (i) the growth rate of sellers across sizes; and (ii) the stationary size. For the former, we target the correlation between store product-level sales growth rates and the relative price of those products. For the latter, we target the stationary size of sellers in the data, so that the average size in the model and the data are comparable. In the model, we measure average size as the mean number of units sold per firm. Since each customer consumes only one unit, the average size is (see e.g. Luttmer (2006)):

$$\overline{L} = \left(\sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} \frac{1}{n} L_n(z)\right)^{-1}; \qquad \text{with } L_n(z) \equiv \frac{nS_n(z)}{\sum_{n,z} nS_n(z)} \tag{29}$$

where $L_n(z)$ is the fraction of active buyers that are customers of sellers of type (n, z). In our sample, the average number of units sold of each product within a store is 12.4 in volume-equivalent

[36] We obtain this number 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 28%, implying an average markup of $.28/(1 - .28) \approx .39$. For comparison, Hottman (2017) estimates average markups in the U.S. retail sector and finds slightly lower numbers, in the range 29-33%.

terms,[37] so we target this number in the estimation.

Estimation Results

Table A.1 presents the full set of calibrated parameters. Table A.2 shows the result of the calibration exercise in terms of moment-matching. We are able to match both aggregate entry rates and average markups accurately, as well as the average and the left-tail dispersion in relative prices.[38] The model slightly under-predicts dispersion in normalized sales, probably as the result of outliers in the data. On the other hand, the correlation between sales growth and relative prices in the model is slightly strong relative to its empirical counterpart. This likely reflects factors attenuating the relationship between prices and sales that are not captured by the model.

Figure A.4 in Appendix A plots the joint surplus, pricing policy function, measured markups, and promised utility, in the space of seller sizes (n) and productivities (z), for the calibrated set of parameter values. We find that matches with more customers and higher productivity levels (i.e. lower values for z) earn a larger surplus. Moreover, the pricing policy is increasing in the size of the customer base, and decreasing in productivity. Even though marginal costs are higher for larger firms (as $\psi > 1$), measured markups are still increasing in size. In Figure A.5 we plot the distribution of normalized sales and that of the seller's customer base that result from the simulation of the estimated model. The figure demonstrates that our model can generate an invariant size distribution with a fat right-tail in both seller revenues and output that resembles its empirical counterpart (see Figure D.1 in the Appendix).

Validation To validate the results of our calibration, we assess the model's performance on non-targeted moments. We look at two sets of moments. First, we check the model's performance on other measures of relative price dispersion, namely inter-decile ranges between the first and ninth deciles, and fifth and ninth deciles. The results are in Panel A of Table A.3. The model's predictions regarding these measures are in line with the microdata. Then, we look at the model's ability to generate a quantitatively correct behavior of price *changes*. For this exercise, in the model we compute micro-price statistics along the stationary solution using the formulas derived in Appendix G.4. In the data, we define the weekly frequency of price changes within a store and market of interest as the share of goods sold by that store in that week that experience a price change.[39] For the distribution of price changes, we look at the absolute value of log differences.

[37] The IRI sample provides a conversion system whereby units of different product categories can be made comparable. We use this standardization for this calculation.

[38] We also considered a sales-weighted *harmonic* average markup, $\overline{m}^h \equiv \left(\sum_{n,z} s_n(z) \left(\frac{p_n(z)}{mc_n(z)}\right)^{-1}\right)^{-1}$, instead of the arithmetic average (equation (28)), as the latter can sometimes overpredict the aggregate markup (see e.g. Baqaee and Farhi (2018)). In the baseline calibration, the arithmetic markup is indeed larger than the harmonic one, though the difference is not large: $\overline{m}^h = 1.3508$.

[39] We focus only on *regular* price changes, which we define (following Coibion *et al.* (2015)) as changes in prices that are larger than 1% or US$0.01 in absolute value for products that are neither entering nor coming out of promotion, and whose initial price is less than, or equal to, $5. For non-promotional goods with initial prices higher than $5, this threshold 0.5%. These criteria eliminate small price changes that may be due to rounding or reporting errors.

Panel B in Table A.3 reports the simulated moments and their empirical counterparts. The calibrated model does a good job in predicting the empirical frequency of price changes, even though these moments were not targeted. The model predicts relatively well the median price duration, though the average duration is not accurately predicted because the distribution of price durations in the model is not sufficiently skewed.[40] The model also predicts several moments of the distribution of expected price changes, especially the average price change and the median. Moreover, about one third of the dispersion in the size of price changes can be explained. Figure A.6 in Appendix A shows how these measures of price changes vary across firm size: in the estimated model, smaller firms experience more frequent and larger price changes, as these are the sellers most eager to grow.

5 Aggregate Fluctuations with Customer Capital

In this section, we analyze the impact of aggregate supply and demand fluctuations at both the macroeconomic level as well as in the cross-section of firms using the estimated model.[41] We seek to understand how sellers' incentives to accumulate customers can generate amplification and persistence through both level and distributional effects, generate incomplete pass-through on prices, and imply procyclicality in markups. Relative to the literature, we emphasize two new dimensions: (i) heterogeneities in the response in the cross-section of firms, both on impact and in the transition (Section 5.1); and (ii) using the decomposition of equation (15), we identify the main channels behind the adjustment in prices in response to each type of shock (Section 5.2).

5.1 Level and Cross-Sectional Effects

Supply Shocks Starting from the stationary equilibrium of the calibrated economy, we study the response to a temporary negative 1% supply shock to the marginal cost. The aggregate state φ follows a mean-reverting process in logs (details in Appendix E.1). The shock hits at some time t_0, and the process continues without further shocks for $t > t_0$. Figure A.9 presents the results. The response of the economy to the aggregate shock combines both level and compositional effects. Let us describe each of these in detail.

First, due to an increase in the cost of serving each customer (panel (a)), the flow payoff in joint surplus falls on impact (panel (b)). To mitigate the effects on their own profits, sellers lower the continuation utility that they promise to deliver to each customer going forward (panel (c)). Thus, active buyers are hit harder by the shock than sellers. As a result, firms attract less inactive

Promotional goods are flagged by the IRI directly. In order to filter out temporary price reductions that may not have been flagged, we also exclude changes that return to their initial level within 3 weeks after the initial change.

[40] To transform frequency f to duration d, we use $d = -\frac{1}{\log(1-f)}$. See details in Appendix G.4. For medians, we apply the formula directly on the median frequency to obtain the median duration. For means, we first use the formula to compute the implied duration for each store and price, and then take the mean.

[41] Because the equilibrium is block-recursive, aggregate dynamics in the model are internalized by agents and anticipated fully, making these "shocks" akin to business cycles fluctuations rather than unexpected perturbations.

buyers, for ex-ante values from matching have lowered. The average tightness in the market falls (panel (d)), and with it firm growth. Interestingly, while prices increase in response to the shock, the pass-through is incomplete (panel (e)). The increase in prices is due to the fact that, when faced with an adverse shock to their costs, sellers choose to re-balance their contracts by front-loading payments from their buyers. They implement this by choosing to exploit their customers more today (through a higher price). Yet, to honor their promises, they must choose an increasing path of buyer utilities in the transition. As the shock is smoothed out inter-temporally via these two contracting instruments, the immediate price response is muted. In the calibrated economy, in particular, the price response is only about 12% the size of the shock. Since prices react less than one-for-one, the average measured markup (panel (l), solid line) falls with output, i.e. it is procyclical. Note, moreover, that this dynamic re-balancing of payments increases flow sales in spite of the decrease in the extensive margin of demand, though this increase is only temporary (panel (f), solid line). Flow profits decrease, in contrast, as the rise in sales is overwhelmed by the increase in costs (panel (f), dashed line).

The model is well-equipped to study the distributional features of these responses. Looking at the response by seller size, we find that smaller sellers (dashed line) respond stronger on impact, while the largest sellers' response (dotted line) is weaker than the average. Similar features have been documented in the data. For instance, Hong (2017) has found differential responses of markups across firm sizes, with smaller firms displaying more elastic responses to output shocks. In the model, this occurs because smaller sellers experience more frequent and larger price changes per unit of time, since the optimal pricing policy is concave in seller size.

To explain the behavior of measured markups in the transition and in the cross section, we must first understand the distributional consequences of the shock. First, the rate of inactive buyers gradually increases (panel (g)), so firms start to shrink on average (panel (h)). These trends continue for a few periods, until they are eventually reversed by the continued increase in promised utilities and seller growth rates. In the first phase of the transition, therefore, the size distribution is gradually shifting to the left. Because measured markups and size are positively correlated in the calibrated economy, the increase in the relative measure of small firms means that the contribution of low-markup firms to the aggregate markup is now relatively more important. Therefore, the persistence of the shock on markups is higher for smaller sellers, as reflected in the fact that markups take longer to mean-revert for these type of sellers relative to larger ones.[42]

To recap, sellers smooth out the effects of the adverse shock inter-temporally by raising prices imperfectly and depressing promised utilities for the future. For further illustration of this result, Figure A.7 shows how the responses discussed above vary with the average duration of customer

[42] We can also explain the behavior of the entry and exit rates. First, since the risk of exiting has become higher for smaller firms, the aggregate exit rate goes up (panel (i)). Interestingly, the entry rate goes up as well (panel (j)), so much so that the number of inactive sellers decreases (panel (k)). Indeed, in the *entry market*, and unlike other markets, the tightness has increased. This is because, in order to enter into the economy, potential sellers must raise resources to pay for the fixed market penetration cost, κ. Since inactive buyers' ex-ante match value has worsen, the seller must now set the initial promised compensation sufficiently high to guarantee that the costs of entry are still being recouped in expectation.

relationships, as measured by δ_c. In particular, we compare the baseline economy (solid line), with an economy in which the duration of customer relationships is one-third shorter (dashed line) and the remaining parameters remain fixed at their original calibrated values. In line with our intuition, Figure A.7 shows that the response is dampened when customer relationships are shorter (that is, when δ_c is higher). This is because, when sellers expect their customers to remain captive for a shorter time, sellers care more about their immediate profits, so promised utility needs to decrease less on impact (panel (a)) as these captive buyers are not expected to remain matched for as long. Accordingly, the seller can charge higher prices on impact (panel (c)), thereby alleviating the adverse effects of the shock on her own valuation (panel (b)). The effects of the shock on prices and continuation utilities become naturally less persistent. Finally, as the price pass-through is less incomplete, the absolute response of the average markup (panel (d)) is weaker. In the limit as δ_c gets very large, markups would tend to full acyclicality, as prices would respond one-for-one and promised utilities would remain unaffected by the shock.

Demand Shocks Recent research has emphasized the relevance that consumer shopping behavior may have on macroeconomic dynamics (e.g. Bai *et al.* (2017), Petrosky-Nadeau and Wasmer (2015), Michaillat and Saez (2015), Paciello *et al.* (2017)). In these papers, a shock in demand can have a strong impact on search incentives, and lead to persistent price responses. In this section, we argue that the underlying size heterogeneity and the forces of firm entry of our model can provide additional insights into the response of markups to aggregate demand shocks.

Hence, we consider a shock to the instantaneous utility v. We implement a 1% negative shock to v at time t_0, and let the process mean-revert without any further shocks for all $t > t_0$. We choose an autocorrelation for the φ process implying a half life of about three years, following Paciello *et al.* (2017) and in line with estimates by Bai *et al.* (2017). Figure A.10 presents the results. A negative shock to the utility from consumption leads to a decrease in the number of buyers looking for a seller, since consumption is worth less. Because the buyers' outside option has relatively improved, firms lower the promised utility in an attempt to smooth the effects of the shock. Once again, the burden of the shock is passed almost entirely onto the customer: the seller's value decreases only slightly (panel (c), dashed line), and it is the decrease in the value of the buyer (panel (c), solid line) which accounts for the bulk of the drop in joint surplus (panel (b)). Market tightness falls on impact (panel (d)), leading to a drop in the matching rate which causes sellers to progressively shrink in size (panel (h)), and more buyers to remain inactive (panel (g)). In this respect, the demand shock is akin to a productivity shock (Figure A.9), in line with the intuition in Bai *et al.* (2017). The compositional effects are similar to those in our previous analysis: a left-ward shift in the firm size distribution, which accounts for the increase in the exit rate and in the relative contribution of low-markup firms to the average response. However, the behavior of prices in response to the demand shock is different than before.

First, the incomplete pass-through that we observed for supply shocks, and which was due to firms optimally tilting their pricing contract toward more immediate payoffs through higher prices today, is no longer present here. A shock to the marginal propensity to consume has a near one-for-one impact on the extensive margin of demand because of linearity in consumer preferences. This means that virtually all the adjustment has to be made along promised utilities, which respond a lot more to the demand shock compared to the case of supply shocks. Sellers understand the temporary nature of the shock ex-ante, so in the transition they increase their promises, and seller growth picks up. As before, the size distribution shifts left in a first phase of the transition. Thus, the level effects of the demand shock are again accompanied by interesting compositional effects.

As a result of the demand shock, the relative attractiveness of small firms improves, as markups decrease relatively more for these firms (see panel (l)). This induces a short-lived spike in the entry of small seller, and a further downward pressure on prices. In addition, the entry of (small) sellers causes a surge in the contribution of low-price firms to the aggregate price level. At the other end of the distribution, the shock decreases the relative mass of large sellers, which attenuates the markup response for these type of sellers. Thus, the cyclicality in the response of measured markups is partly explained by compositional shifts in the firm distribution, whereby the entry of new firms with low prices amplifies the response of the economy to aggregate shocks.

5.2 Firms' Motives for Price Adjustment

Taking stock, markups are more volatile among small sellers after aggregate fluctuations, due to incompleteness in size-dependent pass-through after supply shocks, and to stronger procyclical price variation after demand shocks. To shed light on the channels through which these price adjustments occur, in this section we decompose the price response into the response of the different *price components* ("baseline", "exit", "growth", and "separation"), which we identified in equation (15). As we explained following equation (15), 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 firm exit, growth, or separation.

Figure A.8 shows the response of each component, for the same supply and demand shocks introduced above, overlapped with the overall price response (in red). For supply shocks (left panel), the key observation is that the incomplete pass-through in prices (of about 12%) is mostly driven by a strongly procyclical response in the "growth" component. Intuitively, there is a strong force pushing for a reduction in prices today coming from the fact that firms must compensate their current customers for the acceleration of firm growth in the future, which the customers dislike. Conversely, the "separation" and "exit" channels push for price countercyclicality. Intuitively, after the shock, these channels push for an increase in prices on impact and a decline in the transition because firms shrink in size and become more prone to separations, for which the customers require compensation. For demand shocks (right panel), the main driver of the cyclicality in prices is the "baseline" component. Intuitively, prices adjust to a large extent simply because consumers do not value the produce as much. Again, the "exit" and "separation" components push for countercyclicality, for reasons similar to those given above, but they are not enough to overturn the overall procyclicality of prices.

In sum, in response to adverse supply shocks, sellers grow more concerned with expanding their base (as they offer large price compensation through this margin to generate high rates of subsequent growth). While this phenomenon is to some also extent present after adverse demand shocks, in that case prices change primarily because the seller's current buyers depress their valuation for consumption. This illustrates a clear asymmetry in the motives for price adjustment of firms, depending on the nature of shocks. An interesting avenue of future research is then to check empirically whether firms report similar motives for changing prices along the business cycle in reality.

6 Conclusion

Empirical evidence suggests that a major source of variation in the performance of businesses originates in heterogeneity along idiosyncratic demand components across firms. Further, price differences seem to be important drivers of revenue differences for firms that operate within the same product markets and which have similar productivity levels. Taken at face value, these observations may alter our understanding of the aggregate behavior of markups.

To address these questions, we have presented a dynamic search model of demand accumulation through pricing with aggregate and idiosyncratic shocks and a relevant scope for firm dynamics in order to study the connection between customer capital at the microeconomic level and the macroeconomic dynamics of prices. In the model, sellers of different sizes strategically use menus of prices and continuation promises in order to trade off two conflicting concerns: attracting new customers to increase future market share, and extracting surplus from incumbent customers to increase current profits. The model exhibits cross-sectional price dispersion, and offers a micro-founded interpretation for sluggish firm- and aggregate- dynamics.

Then, we have analyzed a number of predictions on both pricing and firm dynamics dimensions. Using product-level data from the U.S. retail sector, we have estimated the model and conducted experiments on the response of the economy to aggregate shocks to productivity and demand. We have found both level and compositional effects. In response to adverse and mean-reverting aggregate shocks, sellers inter-temporally smooth out the effects of the shock on prices by transferring the burden onto their future buyers, giving rise to an incomplete price response and markup procyclicality. Moreover, we have also shown that smaller sellers experience stronger and more persistent responses, and have identified the main channels of price variation: growth-related concerns after shocks to marginal costs, and variation in fundamentals after shocks to marginal utility.

Overall, these results suggest that incorporating micro-founded pricing behavior into quantitative macroeconomic models can shed new light on certain patterns of macroeconomic dynamics and firm heterogeneity. Further investigating how customer markets may help explain these and other dimensions remains an exciting avenue for future work.

References

Acemoglu, D. and Akcigit, U. (2012). Intellectual property rights policy, competition and innovation. *Journal of the European Economic Association*, **10** (1), 1 – 42.

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*, **Forthcoming**.

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.

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. (2018). Productivity and misallocation in general equilibrium. *Working Paper, Harvard University.*

Berman, N., Rebeyrol, V. and Vicard, V. (2017). Demand learning and firm dynamics: Evidence from exporters. *Review of Economics and Statistics*, **Forthcoming**.

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

—, Klenow, P. J. and Malin, B. A. (2018). Resurrecting the role of the product market wedge in recessions. *American Economic Review*, **108** (4-5), 1118 – 1146.

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.

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**.

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.

Elsby, M. W. L. and Michaels, R. (2013). Marginal jobs, heterogeneous firms, and unemployment flows. *American Economic Journal: Macroeconomics*, **5** (1), 1 – 48.

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. (2004). The pricing behavior of firms in the Euro Area: New survey evidence. In *Inflation Persistence in the Euro Area (European Central Bank).*

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.*

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.*

GALEOTTI, M. and SCHIANTARELLI, F. (1998). The cyclicality of markups in a model with adjustment costs: Econometric evidence for U.S. industry. *Oxford Bulleting of Economics and Statistics*, **60** (2), 121 – 142.

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, 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 cyclicality, 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.

Jovanovic, 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. (2018). Relative price dispersion: Evidence and theory. *American Economic Journal: Microeconomics*, **Forthcoming**.

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.

Nakamura, E. and Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, **123** (4), 1415 – 1464.

— and — (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. (2017). Price dynamics with customer markets. *International Economic Review*, **Forthcoming**.

Perla, J. (2017). Product awareness, industry life cycles, and aggregate profits. *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.

ROSS, S. (1996). *Stochastic Processes*. John Wiley & Sons, Inc., New York.

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.

STOKEY, N. L. and LUCAS, R. E. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

TAUCHEN, G. (1986). Finite state Markov-Chain approximations to univariate and vector autoregressions. *Economics Letters*, **20** (2), 177 – 181.

VARIAN, H. R. (1980). A model of sales. *American Economic Review*, **70** (4), 651 – 659.

Appendix

A Tables and Figures

Param.	Value	Description	Source / Target
		Calibrated externally	
v	1	Value of consumption	Normalization
r	0.05	Discount rate	5% annual risk-free rate
δ_c	0.2041	Separation rate	Paciello *et al.* (2017)
		Estimated internally	
δ_f	0.0738	Firm exit rate	Annual store entry rate
γ	0.5339	Matching elasticity	Average markup
ψ	1.4044	Cost curvature	Standard deviation of relative prices
w	0.1510	Cost scale	p50-p10 inter-decile range in relative prices
c	0.5457	Buyer search cost	Average store size
κ	1.6214	Firm entry cost	Sales growth and relative price correlation
ρ_z	0.0751	Persistence of z	Autocorrelation in normalized store sales
σ_z	0.1034	Volatility of z	Dispersion in normalized store sales
Frequency	Annual		

Table A.1: Full set of calibrated parameters in the baseline estimation. Notes: (ρ_z, σ_z) correspond to the Euler-Maruyama equation (E.1) of the Ornstein-Uhlenbeck process for z (see Appendix E.1).

Moment	Model	Data	Data Source
A. Aggregate moments			
Annual entry rate	0.087	0.089	IRI
Average markup (2001-07)	1.388	1.383	U.S. Census
B. Store-level moments			
sd(Relative prices)	0.1072	0.1055	IRI (Table D.1)
p50-p10 IDR relative prices	1.1224	1.1215	IRI (Table D.1)
Average store size	10.73	12.44	IRI
corr(Sales growth, Relative price)	–0.023	–0.007	IRI
ac(Normalizes sales)	0.854	0.828	IRI
sd(Normalized sales)	0.6	0.474	IRI

Table A.2: Targeted moments: model versus data. Notes: Average markup is weighted by sales shares. *IDR* means inter-decile range. *sd*, *corr*, and *ac* mean "standard deviation", "correlation", and "autocorrelation", respectively.

Moment	Model	Data	
A. Distribution of Relative Prices			
p90-p10 range	1.1994	1.2504	
p90-p50 range	1.0508	1.1149	
B. Distribution of Price Changes			
Average frequency	0.9639	0.9609	annualized
Median frequency	0.9814	0.9264	annualized
Average implied duration	0.2788	0.7817	years
Median implied duration	0.2503	0.3568	years
Average absolute change	0.0305	0.0313	log-points
Median absolute change	0.0312	0.0197	log-points
St. Dev. absolute change	0.0505	0.1415	%-points

Table A.3: Non-targeted moments: model vs. data. Notes: Data moments are taken from our IRI sample. To annualize a weekly rate x, we use $1-(1+x)^{-52}$. See Appendix G.4 for the calculation and aggregation of firm-level price statistics in the model.

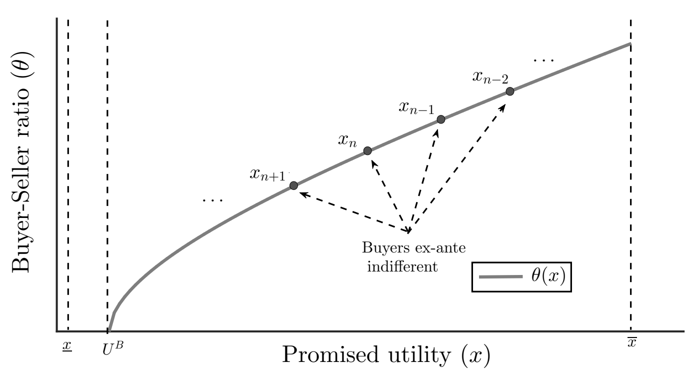

Figure A.1: Equilibrium tightness $\theta : x \mapsto \mu^{-1}\left(\frac{\Gamma^B}{x-U^B}\right)$, and set of equilibrium markets.

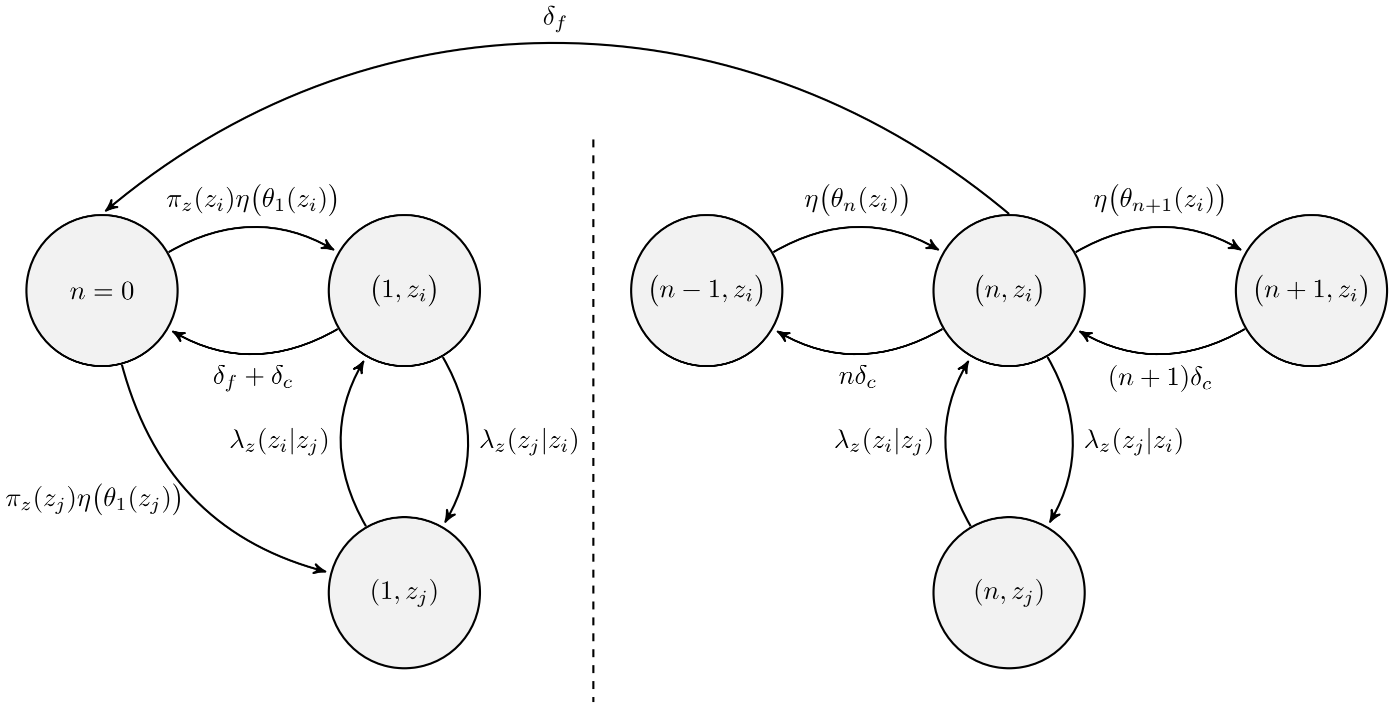

Figure A.2: Seller transitional dynamics for a typical incumbent (right-hand side block) and for entrants (left-hand side block), where φ is fixed for expositional ease. Labels on arrows indicate flow rates.

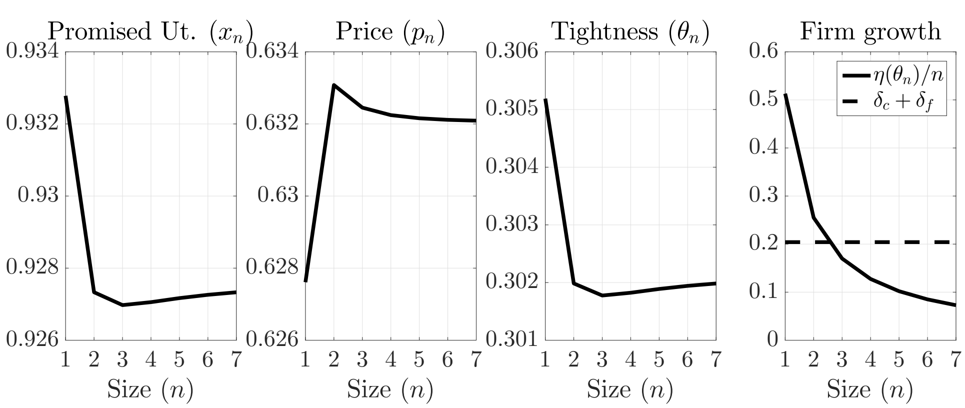

Figure A.3: Same as Figure 2, but with a lower value for κ.

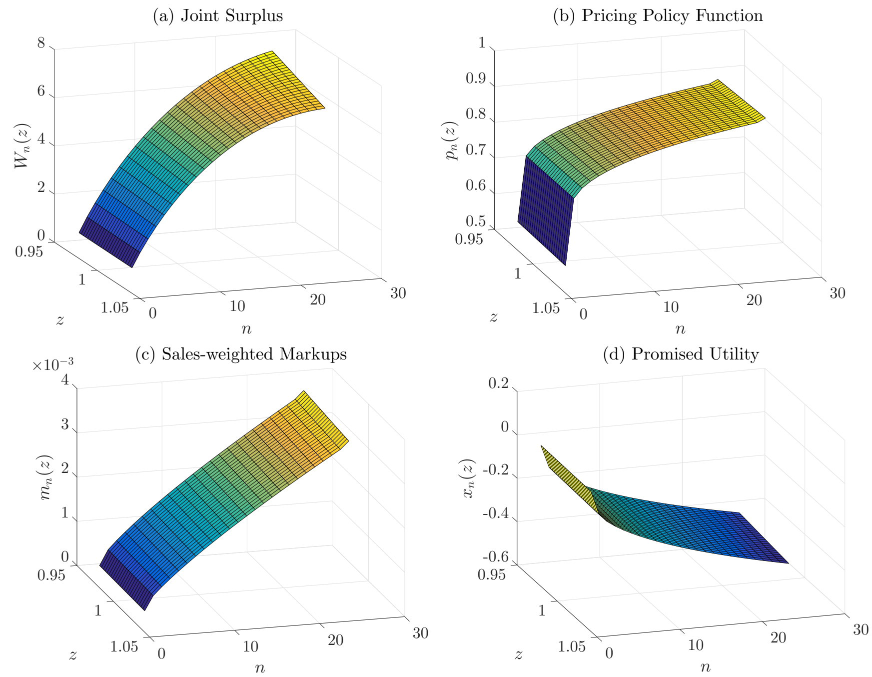

Figure A.4: Joint surplus function, pricing policy function, sales-weighted markups, and promised utility, in the (n, z) space, for the calibrated set of parameters. Higher values of z mean higher costs per customer (i.e. lower productivity).

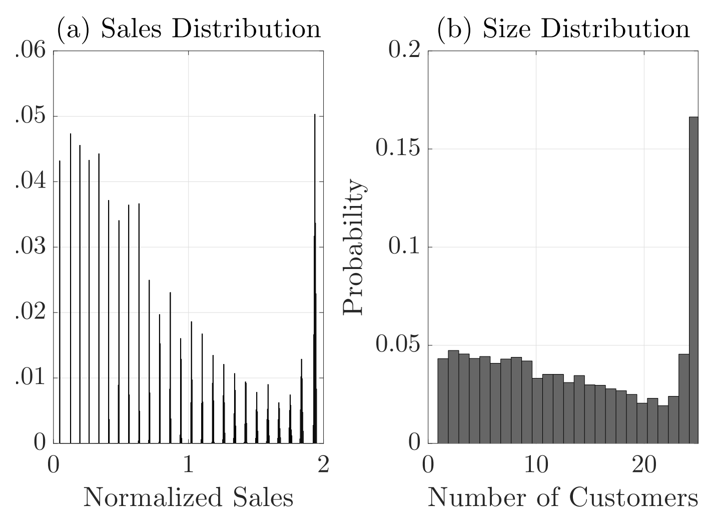

Figure A.5: Distribution of sales and number of customers, for the simulated economy under the calibrated set of parameters. Sales have been normalized by their mean.

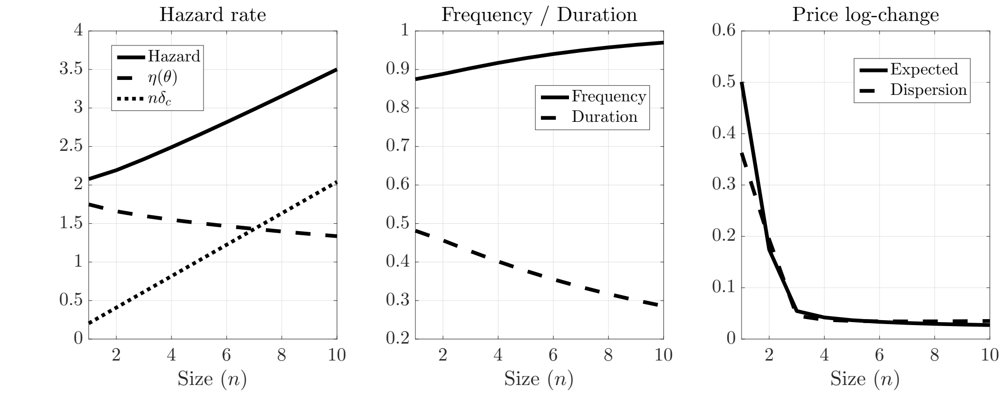

Figure A.6: Hazard rate, frequency, duration, expected price change, and price change dispersion, as a function of size (n), in the calibrated economy.

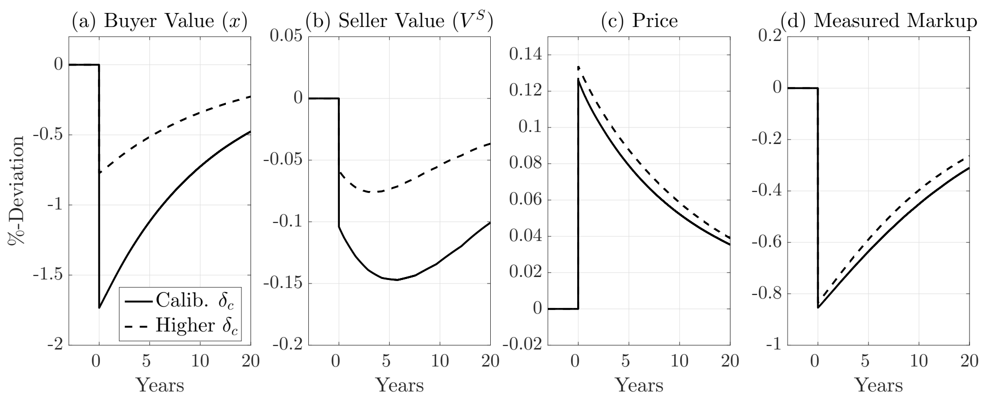

Figure A.7: Impulse responses of selected variables to an aggregate negative and temporary 1% shock to the marginal cost, for different values of δ_c. Notes: See Figure A.9.

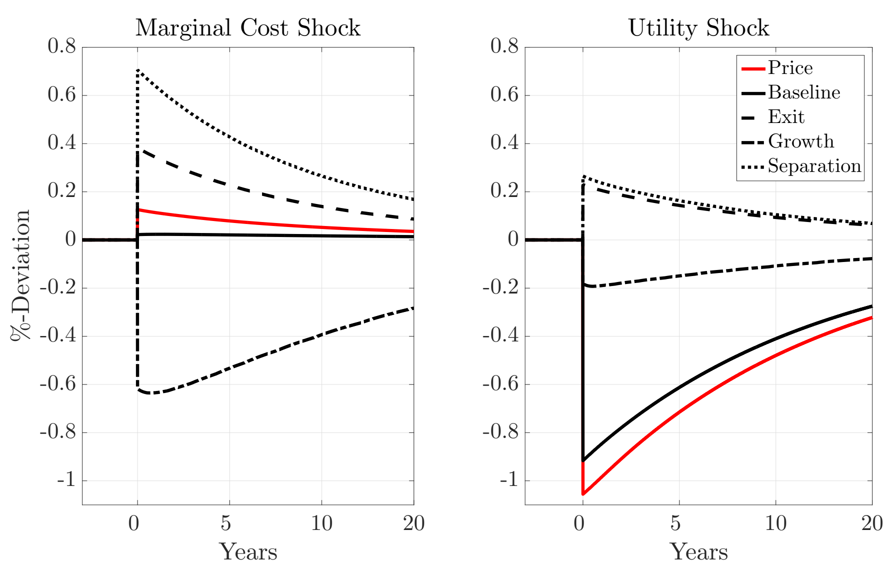

Figure A.8: Impulse responses to negative and temporary 1% supply and demand shocks (same as Figures A.9-A.10). Decomposition of the average price response by the price components identified in equation (15). Each component is averaged using the theoretical distribution of sellers across states. The overall price response is in red. The exogenous shock adjustments (i.e. the last two terms of equation (15)) are not being plotted.

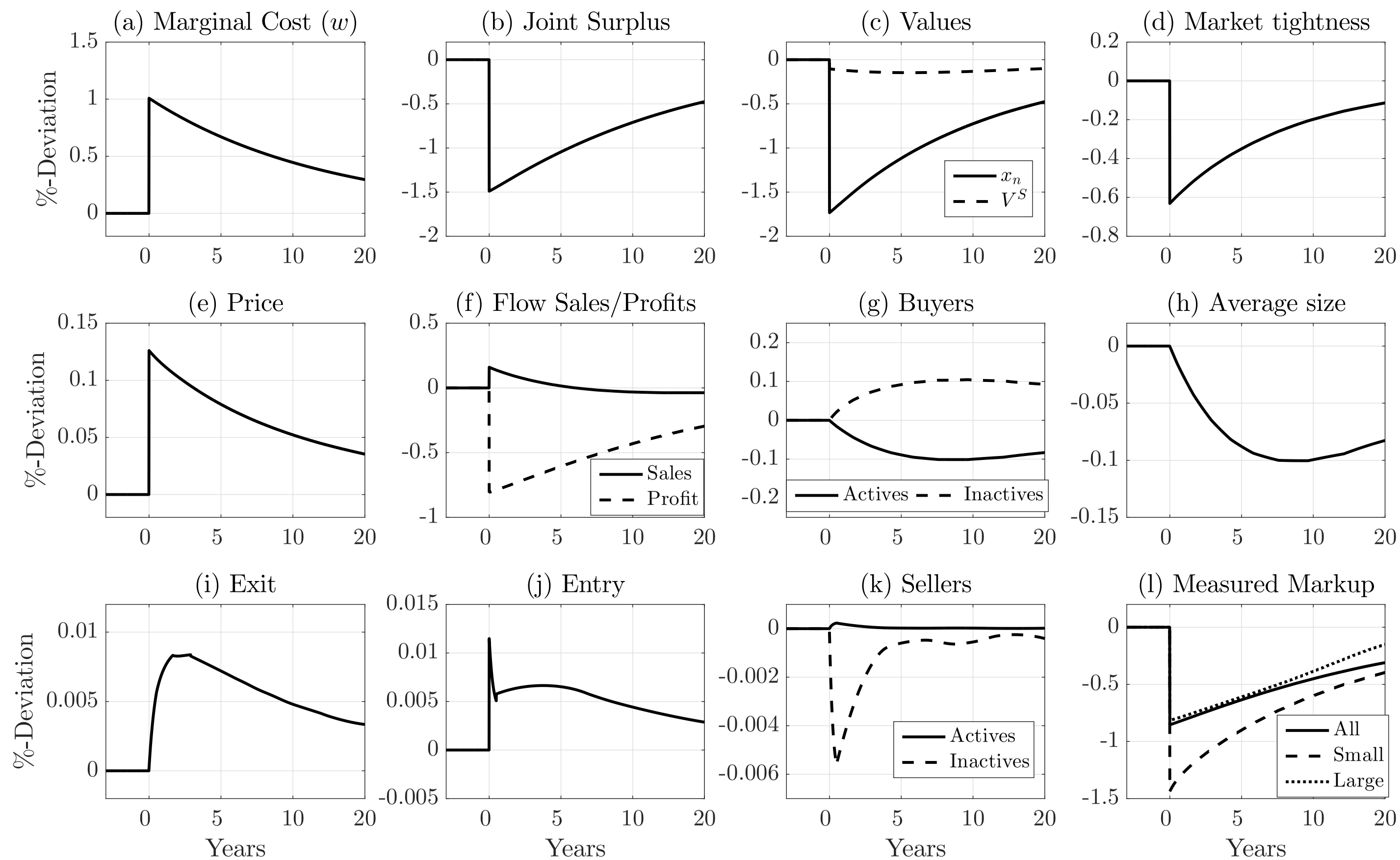

Figure A.9: Impulse responses of selected variables to a negative and temporary 1% shock to aggregate productivity (i.e. an increase in the marginal cost).

Notes: All responses expressed in %-deviations from steady-state. The shock hits at date $t_0 = 0$. The shock is implemented with $k_\varphi = 25$ grid nodes, and paths are smoothed out with cubic splines. Panel (a) depicts the path of the exogenous state. Panels (b) to (f) depict cross-sectional averages using the simulated distribution of firms over idiosyncratic states. That is, for any policy or value function $f(n,z)$, we plot the %-deviation of $\frac{1}{N_t}\sum_{n,z} f(n,z)m_t(n,z)$, where $m_t(n,z)$ is the count of firms of type (n,z) at time t, and $N_t \equiv \sum_{n,z} m_t(n,z)$ is the total count of incumbent firms. The average number of customers per firm in panel (h) is computed using equation (29). Panels (i) and (j) are computed using equation (27). Panel (l) is computed using equation (28).

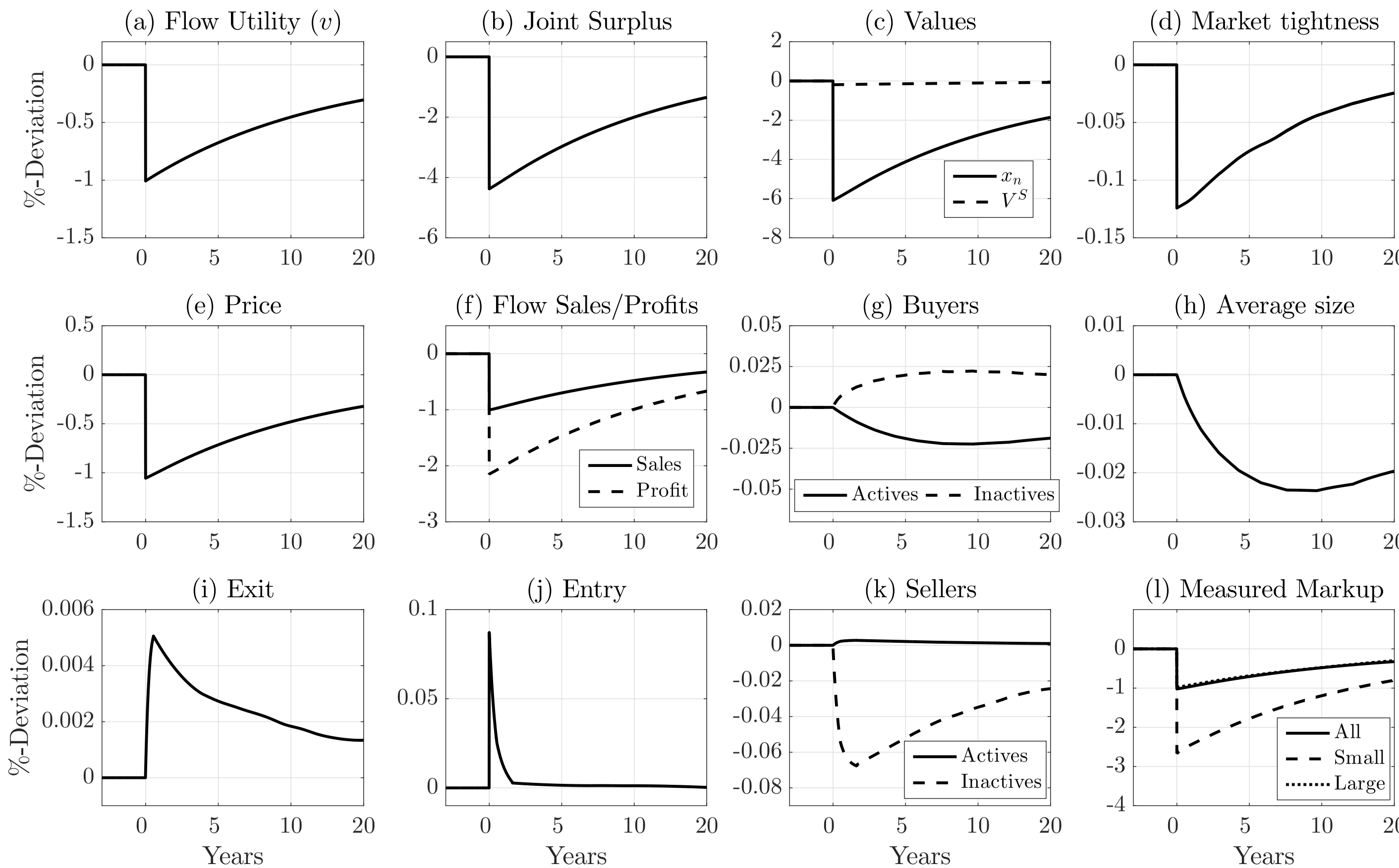

Figure A.10: Impulse responses of selected variables to a negative and temporary 1% shock to aggregate demand (i.e. a decrease in the utility of consumption v), expressed in %-deviations from steady-state. Notes: See Figure A.9.

B Proofs

B.1 Proof of Proposition 1: Joint Surplus Problem

Proof. Denote by $\overline{\boldsymbol{\omega}} = \left\{\overline{p}, \overline{\mathbf{x}}'(n';\mathbf{s}')\right\}$ a generic policy of the typical seller in state $(n;z,\varphi)$, where $\overline{p}$ is the price level,

$$\overline{\mathbf{x}}'(n';\mathbf{s}') = \left\{\overline{x}'(n+1;z,\varphi), \overline{x}'(n-1;z,\varphi), \left\{\overline{x}'(n;z',\varphi) : z' \in \mathcal{Z}\right\}, \left\{\overline{x}'(n;z,\varphi') : \varphi' \in \Phi\right\}\right\}$$

is the set of promised utilities, and $\overline{x}'(n+1;z,\varphi)$ and $\overline{x}'(n-1;z,\varphi)$ are the upsizing and downsizing choices, respectively. Recall that $\overline{x}'(n;z,\varphi) = x$ by stationarity.

The value of the seller in equilibrium, $V^S(n,x;z,\varphi)$, can be written as the maximand on the right-hand side of (6), evaluated at $\overline{\boldsymbol{\omega}}$. That is:

$$V^S(n,x;z,\varphi) = \max_{\overline{\boldsymbol{\omega}}\in\Omega} \quad \widetilde{V}^S(n;z,\varphi|\overline{\boldsymbol{\omega}}) \qquad \text{s.t.} \quad x \leq V^B(n,\overline{\boldsymbol{\omega}};z,\varphi)$$

where $\widetilde{V}^S(n;z,\varphi|\overline{\boldsymbol{\omega}})$ is given by:[43]

$$\begin{aligned}\widetilde{V}^S(n;z,\varphi|\overline{\boldsymbol{\omega}}) \equiv & \frac{1}{\rho(n;z,\varphi)}\Bigg[\overline{p}n - \mathcal{C}(n;z,\varphi) + \eta\Big(\theta\big(\overline{x}'(n+1;z,\varphi);\varphi\big)\Big)V^S\Big(n+1,\overline{x}'(n+1;z,\varphi);z,\varphi\Big) \\ & + n\delta_c V^S\Big(n-1,\overline{x}'(n-1;z,\varphi);z,\varphi\Big) \\ & + \sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)V^S\Big(n,\overline{x}'(n;z',\varphi);z',\varphi\Big) + \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)V^S\Big(n,\overline{x}'(n;z,\varphi');z,\varphi'\Big)\Bigg]\end{aligned} \tag{B.1.1}$$

and we have defined $\rho(n;z,\varphi) \equiv r + \delta_f + n\delta_c + \eta\big(\theta(\overline{x}'(n+1;z,\varphi);\varphi)\big)$ as the *effective* discount rate of the firm.

From (B.1.1), it is optimal to offer the highest possible price that is consistent with promise-keeping, for any given policy $\overline{\boldsymbol{\omega}}$. Indeed, the price has no bearing on the agents' incentives within the search market. Therefore, the PK constraint must bind with equality, and we can solve for the price $\overline{p}$ such that $x = V^B(n,\overline{\boldsymbol{\omega}};z,\varphi)$ using equation (5):

$$\begin{aligned} p^{PK} : \overline{\mathbf{x}}'(n';\mathbf{s}') \mapsto \ & \Big\{v(\varphi) - \rho(n;z,\varphi)x + \delta_f U^B(\varphi) + \eta\Big(\theta\big(\overline{x}'(n+1;z,\varphi);\varphi\big)\Big)\overline{x}'(n+1;z,\varphi) \\ & + \delta_c\Big(U^B(\varphi) + (n-1)\overline{x}'(n-1;z,\varphi)\Big) + \sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)\overline{x}'(n;z',\varphi) + \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\overline{x}'(n;z,\varphi')\Big\}\end{aligned} \tag{B.1.2}$$

Using the above notation, we can now substitute the price level $p^{PK}\big(\overline{\mathbf{x}}'(n';\mathbf{s}')\big)$ from (B.1.2) into the seller's value (B.1.1). After some straightforward algebra, we obtain:

$$\widetilde{W}(n,x;z,\varphi|\overline{\boldsymbol{\omega}}) = \frac{1}{\rho(n;z,\varphi)}\Bigg[n\Bigg(v(\varphi) + (\delta_f+\delta_c)U^B(\varphi)\Bigg)$$

[43] Here we are arguing by free-entry that $\widetilde{V}^S(0;\emptyset,\varphi) = 0$, $\forall\varphi\in\Phi$, to simplify the expression for $\widetilde{V}^S$. Moreover, we use that $\sum_{z'\in\mathcal{Z}}\lambda_z(z'|z) = \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi) = 0$.

$$
\begin{aligned}
&- \left(\mathcal{C}(n;z,\varphi) + \eta\Big(\theta\big(\overline{x}'(n+1;z,\varphi);\varphi\big)\Big)\overline{x}'(n+1;z,\varphi) \right) \\
&+ \eta\Big(\theta\big(\overline{x}'(n+1;z,\varphi);\varphi\big)\Big) W\Big(n+1, \overline{x}'(n+1;z,\varphi);z,\varphi\Big) + n\delta_c W\Big(n-1, \overline{x}'(n-1;z,\varphi);z,\varphi\Big) \\
&+ \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z) W\Big(n, \overline{x}'(n;z',\varphi);z',\varphi\Big) + \sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi) W\Big(n, \overline{x}'(n;z,\varphi');z,\varphi'\Big)\Bigg]
\end{aligned} \tag{B.1.3}
$$

where we have defined:

$$
\widetilde{W}(n,x;z,\varphi|\overline{\boldsymbol{\omega}}) \equiv \widetilde{V}^S(n;z,\varphi|\overline{\boldsymbol{\omega}}) + nx \qquad \text{and} \qquad W(n,x;z,\varphi) \equiv \max_{\overline{\boldsymbol{\omega}}\in\Omega} \widetilde{W}(n,x;z,\varphi|\overline{\boldsymbol{\omega}})
$$

as the joint surplus under contract $\overline{\boldsymbol{\omega}}$, and the *maximized* joint surplus, respectively. Next, note that the right-hand side of equation (B.1.3) does not depend on x nor p, and so we can write the joint surplus under a given policy as:

$$
\widetilde{W}(n,x;z,\varphi|\overline{\boldsymbol{\omega}}) = \widetilde{W}_n\Big(\overline{\mathbf{x}}'(n';\mathbf{s}');z,\varphi\Big)
$$

This proves Part 2 of the proposition. Part 1 now readily follows. Since the joint surplus is invariant to the price level by construction, the optimal contract can be found by splitting the program into two separate stages. In the first stage, the seller chooses the vector of continuation values $\overline{\mathbf{x}}'(n';\mathbf{s}')$ that maximizes (B.1.3). In the second stage, once the surplus has been maximized, the seller chooses the promise-compatible price level via equation (B.1.2).

Formally, the optimal contract is $\boldsymbol{\omega}^* = \Big\{p^*, \mathbf{x}'^*(n';\mathbf{s}')\Big\}$, where:

$$
\mathbf{x}'^*(n';\mathbf{s}') \equiv \arg\max_{\mathbf{x}} \ \widetilde{W}_n\big(\mathbf{x};z,\varphi\big) \tag{B.1.4a}
$$

$$
p^* \equiv \ p^{PK}\Big(\overline{\mathbf{x}}'^*(n';\mathbf{s}')\Big) \tag{B.1.4b}
$$

By expressing the problem of the seller in terms of $\widetilde{W}$, we have just shown that the contract that is optimally chosen by the firm, $\boldsymbol{\omega}^*$, must maximize the joint surplus. Conversely, for any vector $\overline{\mathbf{x}}'(n';\mathbf{s}')$ of continuation values that maximizes the joint surplus, there is a price level, given by $p^* = p^{PK}\big(\overline{\mathbf{x}}'^*(n';\mathbf{s}')\big)$, that maximizes the seller's value subject to the PK constraint.

Therefore, the seller's problem (equation (6)) and the joint surplus problem (equations (B.1.4a)-(B.1.4b)) are equivalent. □

B.2 Proof of Proposition 2: Efficiency

Proof. Consider a benevolent planner that is constrained by the search frictions of the economy and seeks to maximize aggregate welfare subject to the resource constraints of the economy. The planner can allocate resources freely, so the problem does not feature contracts or prices. Instead, we label each market segment directly by its tightness, θ. To simplify notation, it is understood that time subscripts embody the entire history of aggregate shocks, which is taken to be some arbitrary path $\varphi^t = (\varphi_j : j \leq t) \subseteq \Phi$.

The planner chooses:

- The tightness in each market segment, $\boldsymbol{\Theta}_t \equiv \big\{\theta_{n_t,t}(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\big\}$;
- Distributions of inactive and active buyers across markets, $\mathbf{B}^I_t \equiv \big\{B^I_{n_t,t}(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\big\}$ and $\mathbf{B}^A_t \equiv \big\{B^A_{n_t,t}(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\big\}$;

- A measure of potential entrants $S_{0,t}$;
- A distribution of firms across states, $\mathbf{S}_t \equiv \{S_{n_t,t}(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\}$.

The planner's objective is:

$$\max_{\substack{\mathbf{\Theta}_t, \mathbf{B}_t^I, \mathbf{B}_t^A \\ S_{0,t}, \mathbf{S}_t}} \quad \mathbb{E}_0 \int_0^{+\infty} e^{-rt} \mathbb{W}_t(\varphi_t) \mathrm{d}t \tag{B.2.1}$$

where

$$\mathbb{W}_t(\varphi_t) = -\kappa(\varphi_t) S_{0,t} + \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} \Big[v(\varphi_t) B^A_{n_t,t}(z_t) - \mathcal{C}(n_t; z_t, \varphi_t) S_{n_t,t}(z_t) - c(\varphi_t) B^I_{n_t,t}(z_t) \Big]$$

The planner is subject to three sets of constraints. The first one concerns the evolution of the firm distribution:

$$\boldsymbol{\partial}_t S_{0,t} = \delta_f \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} S_{n_t,t}(z_t) + \delta_c \sum_{z_t \in \mathcal{Z}} S_{1,t}(z_t) - \sum_{z^e \in \mathcal{Z}} \pi_z(z^e) \eta\big(\theta_{1,t}(z^e)\big) S_{0,t} \tag{B.2.2a}$$

$$\begin{aligned} \boldsymbol{\partial}_t S_{1,t}(z_t) = {} & \pi_z(z_t) \eta\big(\theta_{1,t}(z_t)\big) S_{0,t} + 2\delta_c S_{2,t}(z_t) + \sum_{\tilde{z} \neq z_t} \lambda_z(z_t|\tilde{z}) S_{1,t}(\tilde{z}) \\ & - \Big(\delta_f + \delta_c + \eta\big(\theta_{2,t}(z_t)\big) + \sum_{\tilde{z} \neq z_t} \lambda_z(\tilde{z}|z_t) \Big) S_{1,t}(z_t) \end{aligned} \tag{B.2.2b}$$

$$\begin{aligned} \forall n_t \geq 2: \quad \boldsymbol{\partial}_t S_{n_t,t}(z_t) = {} & \eta\big(\theta_{n_t,t}(z_t)\big) S_{n_t-1,t}(z_t) + (n_t+1)\delta_c S_{n_t+1,t}(z_t) + \sum_{\tilde{z} \neq z_t} \lambda_z(z_t|\tilde{z}) S_{n_t,t}(\tilde{z}) \\ & - \Big(\delta_f + n_t \delta_c + \eta\big(\theta_{n_t+1,t}(z_t)\big) + \sum_{\tilde{z} \neq z_t} \lambda_z(\tilde{z}|z_t) \Big) S_{n_t,t}(z_t); \end{aligned} \tag{B.2.2c}$$

for all $z_t \in \mathcal{Z}$, where z^e denotes the productivity draw upon entry. The second set of constraints describes the distribution of buyers across firms at any given time:

$$\forall (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}: \quad B^A_{n_t,t}(z_t) = n_t S_{n_t,t}(z_t) \tag{B.2.3a}$$

$$\forall (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}: \quad B^I_{n_t,t}(z_t) = \theta_{n_t,t}(z_t) S_{n_t-1,t}(z_t) \tag{B.2.3b}$$

$$1 = \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} \Big(B^A_{n_t,t}(z_t) + B^I_{n_t,t}(z_t) \Big) \tag{B.2.3c}$$

Equation (B.2.3a) states that each customer consumes a single unit; equation (B.2.3b) states that each market segment is in equilibrium, in the sense that the measure of buyers who find a firm in any given market equals the measure of firms within that market who find a new customer; equation (B.2.3c) says that every buyer in the economy is in either the active or the inactive state.

Finally, the mass of potential entering firms must be non-negative in any aggregate state:

$$S_{0,t} \geq 0 \tag{B.2.4}$$

To solve, first we use constraints (B.2.3a) and (B.2.3b) to rewrite (B.2.3c) as:

$$\sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} n_t S_{n_t,t}(z_t) + \sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} \theta_{n_t+1,t}(z_t) S_{n_t,t}(z_t) + S_{0,t}\sum_{z_t\in\mathcal{Z}}\theta_{1,t}(z_t) = 1 \tag{B.2.5}$$

Substituting constraints (B.2.3a) and (B.2.3b) into the objective function:

$$\begin{aligned}\max_{\boldsymbol{\Theta}_t, S_{0,t}, \mathbf{S}_t}\ \mathbb{E}_0\int_0^{+\infty} e^{-rt}\Bigg\{ -\Bigg(\kappa(\varphi_t) + c(\varphi_t)\sum_{z_t\in\mathcal{Z}}\theta_1(z_t)\Bigg) S_{0,t} + v(\varphi_t)\sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} n_t S_{n_t,t}(z_t) \\ - \sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} \mathcal{C}(n_t; z_t, \varphi_t) S_{n_t,t}(z_t) - c(\varphi_t)\sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}}\theta_{n_t+1,t}(z_t) S_{n_t,t}(z_t)\Bigg\}\mathrm{d}t\end{aligned}$$

subject to (B.2.2a), (B.2.2b), (B.2.2c), (B.2.4), and (B.2.5). Conveniently, the variables $\mathbf{B}_t^I$ and $\mathbf{B}_t^A$ have disappeared from the problem. The state vector now only includes measures of firms: $\mathbb{S}_t \equiv [S_{0,t}, \mathbf{S}_t]$. The current-value Hamiltonian of the simplified planning problem is:

$$\begin{aligned}\mathcal{H}_t(\boldsymbol{\Theta_t}; \mathbb{S}_t) \equiv\ & -\Bigg(\kappa(\varphi_t) + c(\varphi_t)\sum_{z_t\in\mathcal{Z}}\theta_1(z_t)\Bigg) S_{0,t} + v(\varphi_t)\sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} n_t S_{n_t,t}(z_t) \\ & - \sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}}\mathcal{C}(n_t; z_t, \varphi_t) S_{n_t,t}(z_t) - c(\varphi_t)\sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}}\theta_{n_t+1,t}(z_t) S_{n_t,t}(z_t) \\ & + \phi_t\Bigg[1 - \sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} n_t S_{n_t,t}(z_t) - \sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}}\theta_{n_t+1,t}(z_t) S_{n_t,t}(z_t) - S_{0,t}\sum_{z_t\in\mathcal{Z}}\theta_{1,t}(z_t)\Bigg] \\ & + \psi_{0,t}\Bigg[\delta_f\sum_{n_t=1}^{+\infty}\sum_{z_t\in\mathcal{Z}} S_{n_t,t}(z_t) + \delta_c\sum_{z_t\in\mathcal{Z}} S_{1,t}(z_t) - \sum_{z^e\in\mathcal{Z}}\pi_z(z^e)\eta\big(\theta_{1,t}(z^e)\big) S_{0,t}\Bigg] \\ & + \sum_{z_t\in\mathcal{Z}}\Bigg\{\psi_{1,t}(z_t)\Bigg[\pi_z(z_t)\eta\big(\theta_{1,t}(z_t)\big)S_{0,t} + 2\delta_c S_{2,t}(z_t) + \sum_{\tilde z\neq z_t}\lambda_z(z_t|\tilde z) S_{1,t}(\tilde z) \\ & \qquad - \Bigg(\delta_f + \delta_c + \eta\big(\theta_{2,t}(z_t)\big) + \sum_{\tilde z\neq z_t}\lambda_z(\tilde z|z_t)\Bigg) S_{1,t}(z_t)\Bigg] \\ & \quad + \sum_{n_t=2}^{+\infty}\psi_{n_t,t}(z_t)\Bigg[\eta\big(\theta_{n_t,t}(z_t)\big) S_{n_t-1,t}(z_t) + (n_t+1)\delta_c S_{n_t+1,t}(z_t) + \sum_{\tilde z\neq z_t}\lambda_z(z_t|\tilde z) S_{n_t,t}(\tilde z) \\ & \qquad - \Bigg(\delta_f + n_t\delta_c + \eta\big(\theta_{n_t+1,t}(z_t)\big) + \sum_{\tilde z\neq z_t}\lambda_z(\tilde z|z_t)\Bigg) S_{n_t,t}(z_t)\Bigg]\Bigg\} + \vartheta_t S_{0,t}\end{aligned}$$

where $\psi_{n,t}(z)\geq 0$, $n\geq 1$ (respectively, $\psi_{0,t}\geq 0$) is the co-state variable on the flow equation for $S_{n,t}(z)$ (respectively, $S_{0,t}$); $\phi_t \geq 0$ is the multiplier on (B.2.5); and $\vartheta_t \geq 0$ is the multiplier on the non-negative entry condition, where the corresponding complementary slackness hold. In vector notation, the necessary conditions for optimality are:

$$\begin{aligned}\nabla_{\boldsymbol{\Theta}}\mathcal{H}_t(\boldsymbol{\Theta_t}; \mathbb{S}_t) &= \mathbf{0} \\ \nabla_{\mathbb{S}}\mathcal{H}_t(\boldsymbol{\Theta_t}; \mathbb{S}_t) &= -\nabla_t\boldsymbol{\psi}_t + r\boldsymbol{\psi}_t\end{aligned}$$

where ∇ denotes the gradient operator, and $\boldsymbol{\psi}_t$ is a stacked vector of co-state variables. These conditions are also sufficient because the Hamiltonian is quasi-concave. Indeed, the objective function is linear in both control and state variables, and because of Assumption 2 establishing concavity of η, all the constraints are concave in the control and linear in the states.

Regarding the first set of optimality conditions, for given $z_t \in \mathcal{Z}$ we have:

$$[\theta_1]: \qquad \phi_t + c(\varphi_t) = \Big(\psi_{1,t}(z_t) - \psi_{0,t}\Big)\pi_z(z_t)\frac{\partial \eta(\theta)}{\partial \theta}\Big|_{\theta=\theta_{1,t}(z_t)} \tag{B.2.6a}$$

$$[\theta_n : n \geq 2]: \qquad \phi_t + c(\varphi_t) = \Big(\psi_{n_t,t}(z_t) - \psi_{n_t-1,t}(z_t)\Big)\frac{\partial \eta(\theta)}{\partial \theta}\Big|_{\theta=\theta_{n_t,t}(z_t)} \tag{B.2.6b}$$

As for the second set of conditions, we have:

$$\begin{aligned}
[S_0]: \quad -\boldsymbol{\partial}_t \psi_{0,t} + r\psi_{0,t} \;&=\; -\kappa(\varphi_t) - \big(\phi_t + c(\varphi_t)\big)\sum_{z_t\in\mathcal{Z}}\theta_1(z_t) && \text{(B.2.7a)}\\
&+ \sum_{z^e\in\mathcal{Z}}\pi_z(z^e)\eta\big(\theta_{1,t}(z^e)\big)\psi_{1,t}(z^e) - \psi_{0,t}\sum_{z^e\in\mathcal{Z}}\pi_z(z^e)\eta\big(\theta_{1,t}(z^e)\big) + \vartheta_t\\
[S_{n_t}(z_t)]: \quad -\boldsymbol{\partial}_t\psi_{n_t,t}(z_t) + r\psi_{n_t,t}(z_t) \;&=\; n_t\big(v(\varphi_t) - \phi_t\big) - \big(\phi_t + c(\varphi_t)\big)\theta_{n_t+1,t}(z_t) - \mathcal{C}(n_t, z_t; \varphi_t)\\
&+ \delta_f\Big(\psi_{0,t} - \psi_{n_t,t}(z_t)\Big) + n_t\delta_c\Big(\psi_{n_t-1,t}(z_t) - \psi_{n_t,t}(z_t)\Big) && \text{(B.2.7b)}\\
&+ \eta\big(\theta_{n_t+1}(z_t)\big)\Big(\psi_{n_t+1,t}(z_t) - \psi_{n_t,t}(z_t)\Big) + \sum_{\tilde{z}\in\mathcal{Z}}\lambda_z(\tilde{z}|z_t)\Big(\psi_{n_t,t}(\tilde{z}) - \psi_{n_t,t}(z_t)\Big)
\end{aligned}$$

for given $z_t \in \mathcal{Z}$, where in the last line we have used that $\lambda_z(z|z) = -\sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)$ for all $z \in \mathcal{Z}$, by the properties of the Markov chain.

We now show that a block-recursive equilibrium with non-negative entry of firms satisfies the optimality conditions of the planner by appropriately choosing the co-state variables of the planning problem. By equations (B.2.7a)-(B.2.7b), the co-state variables can be represented as HJB equations. Equations (B.2.6a)-(B.2.6b) are the corresponding first order conditions of those equations. Therefore, it suffices to find the values of the multipliers for which the HJB equations of the planner coincide with the joint surplus problem of the decentralized allocation.

Pick a decentralized equilibrium allocation $\big\{W_n(z,\varphi), x_n(z,\varphi), \theta_n(z,\varphi), U^B(\varphi) : (n,z,\varphi) \in \mathbb{N}\times\mathcal{Z}\times\Phi\big\}$, and consider the following realization for the multipliers:

$$\begin{aligned}
\phi_t(\varphi^t) \;&=\; rU^B(\varphi_t)\\
\psi_{0,t}(\varphi^t) \;&=\; 0\\
\forall n_t, z_t: \quad \psi_{n_t,t}(z_t,\varphi^t) \;&=\; W_{n_t}(z_t,\varphi_t) - n_tU^B(\varphi_t)
\end{aligned}$$

Under this guess, notice that $\boldsymbol{\partial}_t\psi_{0,t} = \boldsymbol{\partial}_t\psi_{n,t}(z_t) = 0$, $\forall n \geq 1$. Moreover, the multipliers depend only on the current realization of the aggregate state, and not the entire history. Further, for a sufficiently low value of κ, we can impose strictly positive entry and therefore $\vartheta_t = 0$, $\forall t$.

Plugging these guesses into (B.2.7b), after some simple algebra we obtain:

$$(r+\delta_f)W_{n_t}(z_t,\varphi_t) = n_t\Big(v(\varphi_t) + (\delta_f+\delta_c)U^B(\varphi_t)\Big) - \mathcal{C}(n_t, z_t; \varphi_t)$$

$$
\begin{aligned}
&- \Big[\big(rU^B(\varphi_t) + c(\varphi_t)\big)\theta_{n_t+1,t}(z_t) + \eta\big(\theta_{n_t+1}(z_t)\big)U^B(\varphi_t) \Big] \\
&+ n_t\delta_c \Big(W_{n_t-1}(z_t,\varphi_t) - W_{n_t}(z_t,\varphi_t) \Big) \\
&+ \eta\big(\theta_{n_t+1}(z_t)\big) \Big(W_{n_t+1}(z_t,\varphi_t) - W_{n_t}(z_t,\varphi_t) \Big) \\
&+ \sum_{\tilde{z}\in\mathcal{Z}} \lambda_z(\tilde{z}|z_t) \Big(W_{n_t}(\tilde{z},\varphi_t) - W_{n_t}(z_t,\varphi_t) \Big)
\end{aligned}
$$

The last equation resembles the maximized HJB equation for the joint surplus (equation (10)) except for the second line in square brackets. Using that $\eta(\theta) = \theta\mu(\theta)$ and $x_{n+1}(z,\varphi) = U^B(\varphi) + \frac{rU^B(\varphi)+c(\varphi)}{\mu(\theta_{n+1}(z,\varphi))}$ by inactive buyers' indifference, we obtain:

$$
\big(rU^B(\varphi_t) + c(\varphi_t)\big)\theta_{n_t+1,t}(z_t) + \eta\big(\theta_{n_t+1}(z_t)\big)U^B(\varphi_t) = \eta\left(\theta_{n_t+1,t}(z_t,\varphi_t)\right) x_{n_t+1,t}(z_t,\varphi_t) \tag{B.2.8}
$$

Using this into the above equation and grouping terms, we will then recognize the value of the joint surplus in the decentralized solution, equation (10).

Similarly, plugging the guess for the multipliers into (B.2.7a), we obtain:

$$
\kappa(\varphi_t) = -\big(rU^B(\varphi_t) + c(\varphi_t)\big) \sum_{z_t\in\mathcal{Z}} \theta_1(z_t) + \sum_{z^e\in\mathcal{Z}} \pi_z(z^e)\eta\big(\theta_{1,t}(z^e)\big) \Big(W_1(z^e) - U^B(\varphi_t) \Big)
$$

A final manipulation using (B.2.8) again then allows us to obtain the free entry condition in the decentralized allocation, equation (13).

Summing up, under an appropriate choice of the co-states, the planner's solution is equivalent to the problem of the decentralized economy. Hence, the equilibrium is constrained-efficient. □

B.3 Proof of Proposition 3: Joint Surplus Solution

Proof. The equilibrium allocation is composed of sequences:

$$
\big\{ W_n(z,\varphi), x_n(z,\varphi), \theta_n(z,\varphi), p_n(z,\varphi) : (n,z,\varphi) \in \mathbb{N} \times \mathcal{Z} \times \Phi \big\}
$$

satisfying equations (10), (11), and (15), where the free entry condition (13) pins down x_1 and the first-order condition (12) pins down x_n given x_{n-1}, $n \geq 2$, for any $\varphi \in \Phi$.

For $\mu(\theta) = \theta^{\gamma-1}$, $\gamma \in (0,1)$, equation (11) defines the following equilibrium mapping:

$$
\theta : (x;\varphi) \mapsto \left(\frac{x - U^B(\varphi)}{\Gamma^B(\varphi)} \right)^{\frac{1}{1-\gamma}} \tag{B.3.1}
$$

Some algebra shows that equation (12) can be written as:

$$
W_{n+1}(z,\varphi) - W_n(z,\varphi) - x_{n+1}(z,\varphi) = \frac{1-\gamma}{\gamma} \Big(x_{n+1}(z,\varphi) - U^B(\varphi) \Big) \tag{B.3.2}
$$

One can also write this condition as:

$$
x_{n+1}(z,\varphi) - U^B(\varphi) = \gamma \underbrace{\Big(W_{n+1}(z,\varphi) - W_n(z,\varphi) - U^B(\varphi) \Big)}_{\equiv \Gamma_{n+1}(z,\varphi)}
$$

showing that the buyer absorbs a fraction γ of the marginal gains from matching, $\Gamma_{n+1}(z,\varphi)$. Next, define:

$$\Gamma_n^S(z,\varphi) \equiv (r+\delta_f)W_n(z,\varphi) - \pi_n(z,\varphi) + n\delta_c\Big(W_n(z,\varphi) - W_{n-1}(z,\varphi)\Big) - n(\delta_c+\delta_f)U^B(\varphi) - \Xi_n(z,\varphi) \quad \text{(B.3.3)}$$

where $\pi_n(z,\varphi) \equiv nv(\varphi) - \mathcal{C}(n;z,\varphi)$ is the flow joint surplus, and:

$$\Xi_n(z,\varphi) \equiv \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)W_n(z',\varphi) + \sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi)W_n(z,\varphi')$$

is the expected value of the joint surplus across exogenous states. Next, note:

$$\begin{aligned}
\Gamma_n^S(z,\varphi) &= \eta\big(\theta_{n+1}(z,\varphi)\big)\Big(W_{n+1}(z,\varphi) - W_n(z,\varphi) - x_{n+1}(z,\varphi)\Big) \\
&= \left(\frac{1-\gamma}{\gamma}\right)\eta\big(\theta_{n+1}(z,\varphi)\big)\Big(x_{n+1}(z,\varphi) - U^B(\varphi)\Big) \\
&= \left(\frac{1-\gamma}{\gamma}\right)\theta_{n+1}(z,\varphi)\Gamma^B(\varphi)
\end{aligned}$$

where the first line uses the HJB equation for the joint surplus (equation (10)), the second line uses (B.3.2), and the third line uses (B.3.1) and $\eta(\theta) = \theta\mu(\theta)$. The right-hand side of the first equality allows us to interpret Γ^S as the expected match surplus for the seller (see main text).

Using the last equality, we have found the market tightness:

$$\theta_{n+1}(z,\varphi) = \left(\frac{\gamma}{1-\gamma}\right)\frac{\Gamma_n^S(z,\varphi)}{\Gamma^B(\varphi)} \quad \text{(B.3.4)}$$

for all $n \geq 1$. Finally, we can write (B.3.2) as:

$$W_{n+1}(z,\varphi) - W_n(z,\varphi) = U^B(\varphi) + \frac{1}{\gamma}\big(x_{n+1}(z,\varphi) - U^B(\varphi)\big) = U^B(\varphi) + \frac{1}{\gamma}\Gamma^B(\varphi)\theta_{n+1}(z,\varphi)^{1-\gamma}$$

Using (B.3.4) and rearranging terms, we obtain our desired result:

$$W_{n+1}(z,\varphi) = W_n(z,\varphi) + U^B(\varphi) + \left(\frac{\Gamma^B(\varphi)}{\gamma}\right)^{\gamma}\left(\frac{\Gamma_n^S(z,\varphi)}{1-\gamma}\right)^{1-\gamma} \quad \text{(B.3.5)}$$

This is a second-order difference equation in n. The boundary conditions are $W_0 = 0$ (as the joint value is nil when the seller has no customers), and W_1 set to satisfy the free entry condition (13). By (B.3.2), we know that $W_1 - x_1 = (1-\gamma)(W_1 - U^B)$ and $x_1 - U^B = \gamma(W_1 - U^B)$, and thus we can write (13) as:

$$\kappa(\varphi) = (1-\gamma)\left(\frac{\Gamma^B(\varphi)}{\gamma}\right)^{\frac{\gamma}{\gamma-1}} \sum_{z_0\in\mathcal{Z}} \pi_z(z_0)\Big(W_1(z_0,\varphi) - U^B(\varphi)\Big)^{\frac{1}{1-\gamma}}$$

our desired result. □

B.4 Proof of Proposition 4: Invariant Distribution

Proof. Let $\big\{\theta_n(z,\varphi) : (n,z,\varphi) \in \mathcal{N}\times\mathcal{Z}\times\Phi\big\}$ be an equilibrium collection of market tightness levels, where $\mathcal{N} = \{1,\ldots,\overline{n}\}$, and $\overline{n} < +\infty$ is a large integer. In matrix notation, for each aggregate state $\varphi \in \Phi$, equations (19)-(20)-(21) can be written succinctly as:

$$\boldsymbol{\partial}_t \mathbf{S}_t(\varphi) = \mathbf{T}_\varphi \mathbf{S}_t(\varphi) \tag{B.4.1}$$

where $\mathbf{S}_t(\varphi) \equiv \left(S_{0,t}(\varphi), \mathbf{S}_{1,t}^\top, \ldots, \mathbf{S}_{\overline{n},t}^\top\right)^\top$, with $\mathbf{S}_{n,t} \equiv \left(S_{n,t}(z_1), \ldots, S_{n,t}(z_{k_z})\right)^\top$, and $\mathbf{T}_\varphi$ is the partitioned matrix:

$$\mathbf{T}_\varphi \equiv \begin{pmatrix} t_{11} & \boldsymbol{\delta}_f^e + \boldsymbol{\delta}_c^e & \boldsymbol{\delta}_f^e & \boldsymbol{\delta}_f^e & \cdots & \boldsymbol{\delta}_f^e & \boldsymbol{\delta}_f^e & \boldsymbol{\delta}_f^e \\ \boldsymbol{\eta}_1^e(\varphi)^\top & \mathbf{D}_1(\varphi) & \boldsymbol{\delta}_{2,c} & \mathbf{0}_{k_z:k_z} & \cdots & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} \\ \mathbf{0}_{k_z:1} & \boldsymbol{\eta}_2(\varphi) & \mathbf{D}_2(\varphi) & \boldsymbol{\delta}_{3,c} & \cdots & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0}_{k_z:1} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \cdots & \mathbf{D}_{\overline{n}-2}(\varphi) & \boldsymbol{\delta}_{\overline{n}-1,c} & \mathbf{0}_{k_z:k_z} \\ \mathbf{0}_{k_z:1} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \cdots & \boldsymbol{\eta}_{\overline{n}-1}(\varphi) & \mathbf{D}_{\overline{n}-1}(\varphi) & \boldsymbol{\delta}_{\overline{n},c} \\ \mathbf{0}_{k_z:1} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \mathbf{0}_{k_z:k_z} & \cdots & \mathbf{0}_{k_z:k_z} & \boldsymbol{\eta}_{\overline{n}}(\varphi) & \mathbf{D}_{\overline{n}}(\varphi) \end{pmatrix}$$

where $t_{11} \equiv -\sum_z \pi_z(z)\eta(\theta_1(z,\varphi))$ is a scalar, $\mathbf{0}_{p:q}$ denotes a $p \times q$ matrix of zeros, and $\mathbf{T}_\varphi$ is a $K \times K$ square matrix, where $K = 1 + \overline{n}k_z$. Further, we have defined the $1 \times k_z$ row vectors:

$$\boldsymbol{\delta}_f^e \equiv \Big(\delta_f, \ldots, \delta_f\Big); \qquad \boldsymbol{\delta}_c^e \equiv \Big(\delta_c, \ldots, \delta_c\Big); \qquad \boldsymbol{\eta}_1^e(\varphi) \equiv \Big(\pi_z(z_1)\eta\big(\theta_1(z_1,\varphi)\big), \ldots, \pi_z(z_{k_z})\eta\big(\theta_1(z_{k_z},\varphi)\big)\Big);$$

and the $k_z \times k_z$ matrices:

$$\begin{aligned} \forall n = 2, \ldots, \overline{n}: \qquad & \boldsymbol{\delta}_{n,c} \equiv \mathrm{diag}\Big(n\delta_c, \ldots, n\delta_c\Big); \\ & \boldsymbol{\eta}_n(\varphi) \equiv \mathrm{diag}\Big(\eta\big(\theta_n(z_1,\varphi)\big), \ldots, \eta\big(\theta_n(z_{k_z},\varphi)\big)\Big); \\ \forall n = 1, \ldots, \overline{n}: \quad & \mathbf{D}_n(\varphi) \equiv \begin{pmatrix} d_n(z_1,\varphi) & \lambda_z(z_1|z_2) & \ldots & \lambda_z(z_1|z_{k_z}) \\ \lambda_z(z_2|z_1) & d_n(z_2,\varphi) & \ldots & \lambda_z(z_2|z_{k_z}) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_z(z_{k_z}|z_1) & \lambda_z(z_{k_z}|z_2) & \ldots & d_n(z_{k_z},\varphi) \end{pmatrix} \end{aligned}$$

where diag(.) denotes a diagonal matrix, and the diagonal elements of $\mathbf{D}_n(\varphi)$ are given by:

$$d_n(z_j,\varphi) \equiv \begin{cases} -\Big(\delta_f + n\delta_c + \eta\big(\theta_{n+1}(z_j,\varphi)\big) + \sum_{\ell \neq j} \lambda_z(z_\ell|z_j)\Big) & \text{for } n = 1, \ldots, \overline{n}-1 \\ -\Big(\delta_f + \overline{n}\delta_c + \sum_{\ell \neq j} \lambda_z(z_\ell|z_j)\Big) & \text{for } n = \overline{n} \end{cases}$$

System (B.4.1) describes an *irreducible* Markov chain, as any state $(n', z') \in \mathcal{N} \times \mathcal{Z}$ can be reached almost surely from some $(n, z) \neq (n', z')$. Moreover, the Markov chain is *aperiodic*. These properties, plus the fact that the state space is finite, guarantee that the Markov chain is *ergodic*. Therefore, by Theorem 11.2 of Stokey and Lucas (1989), the system converges to a unique steady-state distribution $\mathbf{S}^*(\varphi)$, for each $\varphi \in \Phi$.

For the analytical characterization, note that the transition matrix $\mathbf{T}_\varphi$ is constant, so we can solve the differential equation (B.4.1) directly. The solution is:

$$\mathbf{S}_t(\varphi) = e^{\mathbf{T}_\varphi t}\mathbf{S}_0(\varphi)$$

where the initial distribution $\mathbf{S}_0(\varphi) \in \mathbb{R}_+^K$ is given. To compute $e^{\mathbf{T}_\varphi t}$, consider the eigenvalue decomposition $\mathbf{T}_\varphi = \mathbf{E}_\varphi \mathbf{\Lambda}_\varphi \mathbf{E}_\varphi^{-1}$, where $\mathbf{\Lambda}_\varphi \equiv \big(\lambda_1(\varphi), \ldots, \lambda_K(\varphi)\big)$ is the diagonal matrix of eigenvalues, and $\mathbf{E}_\varphi$ collects the corresponding eigenvectors. Defining $\mathbf{Z}_t(\varphi) \equiv \mathbf{E}_\varphi^{-1}\mathbf{S}_t(\varphi)$, then $\boldsymbol{\partial}_t \mathbf{Z}_t(\varphi) = \mathbf{\Lambda}_\varphi \mathbf{Z}_t(\varphi)$, and because $\mathbf{\Lambda}_\varphi$ is a diagonal matrix, we can solve this differential equation element-by-element, i.e. $\boldsymbol{\partial}_t Z_{i,t}(\varphi) = \lambda_i(\varphi) Z_{i,t}(\varphi)$ for each $i = 1, \ldots, K$. This is a simple system of ODEs with solution:

$$Z_{i,t}(\varphi) = c_i e^{\lambda_i(\varphi)t}, \qquad i = 1, \ldots, K$$

where $c_i \in \mathbb{R}$ is the constant of integration. Since $\mathbf{S}_t(\varphi) = \mathbf{E}_\varphi \mathbf{Z}_t(\varphi)$, we have obtained:

$$\mathbf{S}_t(\varphi) = \sum_{i=1}^{K} c_i e^{\lambda_i(\varphi)t} \boldsymbol{v}_i \tag{B.4.2}$$

where $\boldsymbol{v}_i$ is the $K \times 1$ eigenvector associated to the i-th eigenvalue. Therefore, the stability of system (B.4.2) as $t \to +\infty$ depends on the sign of the eigenvalues of $\mathbf{T}_\varphi$. The trace of $\mathbf{T}_\varphi$ is:

$$tr(\mathbf{T}_\varphi) = \sum_{i=1}^{K} \lambda_i(\varphi) = -\sum_{j=1}^{k_z} \pi_z(z_j)\eta(\theta_1(z_j, \varphi)) + \sum_{n=1}^{\overline{n}} \sum_{j=1}^{k_z} d_n(z_j) < 0$$

The trace being unambiguously negative means that there is at least one negative eigenvalue, if not more. Letting $1 \leq \ell \leq K$ denote the number of negative eigenvalues, and re-ordering the eigenvalues from small to large with no loss of generality, we can then impose $c_j = 0$, $\forall j \in \{\ell+1, \ell+2, \ldots, K\}$, on equation (B.4.2), and let $t \to +\infty$ to find the stable solution. That is:

$$\mathbf{S}^*(\varphi) = \lim_{t \to +\infty} \sum_{j=1}^{\ell} c_j e^{\lambda_j(\varphi)t} \boldsymbol{v}_j \in \mathbb{R}_+^K$$

is the unique invariant distribution of sellers in state $\varphi \in \Phi$. □

C Price Discrimination

The assumption of no price discrimination across different customers is not key to generate efficient firm dynamics. We argue that:

- So long as we maintain the assumption of dynamic contracts with commitment, our model still generates these dynamics as well as cross-sectional dispersion.
- Allowing for discrimination leads to price indeterminacy.

If sellers were to use prices as their only instrument for customer attraction (instead of recursive contracts with price-utility pairs), an equilibrium with price discrimination across customers of different tenures would look similar to that of Gourio and Rudanko (2014b): firms would attract customers by offering an instantaneous discount on the valuation v, and extract all surplus by charging v immediately after the customer joins the seller, and until separation. However, imposing price discrimination in our model does not lead to this prediction. This is because sellers must still trade off static payoffs coming from the current price with dynamic ones coming from the promised utilities.

Importantly, under price discrimination, tractability is preserved along several dimensions:

(i.) The seller's and the join surplus problems are equivalent;

(ii.) The joint surplus is constant in the distribution of contracts across customers;

(iii.) As a novelty, there is equilibrium price indeterminacy. Therefore, there exist equilibria featuring price dispersion both within and across firms.

Let us discuss these results more formally. For this, we must extend our baseline framework to allow for discrimination across buyers. Let $\boldsymbol{\omega}_i = \{p_i, \mathbf{x}'_i(n';\mathbf{s}')\}$ be the contract offered to the typical customer $i = 1,\ldots,n$, which is composed of an individual-specific price level p_i, and a personalized menu of continuation utilities $\mathbf{x}'_i(n';\mathbf{s}')$, one for each $n' \in \{n-1, n, n+1\}$ and $\mathbf{s}' \in \{(z',\varphi),(z,\varphi')\}$. A seller is characterized by the collection $\{x_i\}_{i=1}^n$ of outstanding promises, and must choose: (i) a menu of contracts $\{\boldsymbol{\omega}_i\}_{i=1}^n$ for the n current customers; and (ii) a starting promised utility $x'_0 \in \mathbb{R}$ for the new $(n+1)$-th incoming customer (if there is any). The HJB equation for the seller now reads:

$$
\begin{aligned}
rV^S\Big(n, \{x_i\}_{i=1}^n; z,\varphi\Big) = \max_{x'_0, \{\boldsymbol{\omega}_i\}_{i=1}^n} \Bigg\{ & \sum_{i=1}^n p_i - \mathcal{C}(n;z,\varphi) + \delta_f\Big(V_0^S(\varphi) - V^S\Big(n, \{x_i\}_{i=1}^n; z,\varphi\Big)\Big) \\
& + \delta_c \sum_{j=1}^n \Big(V^S\Big(n-1, \{x'_i(n-1;z,\varphi)\}_{i=1}^n \setminus_- \{x'_j(n-1;z,\varphi)\}; z,\varphi\Big) - V^S\Big(n, \{x_i\}_{i=1}^n; z,\varphi\Big)\Big) \\
& + \eta\big(\theta(x'_0;\varphi)\big)\Big(V^S\Big(n+1, \{x'_i(n+1;z,\varphi)\}_{i=1}^n \cup_+ \{x'_0\}; z,\varphi\Big) - V^S\Big(n, \{x_i\}_{i=1}^n; z,\varphi\Big)\Big) \\
& + \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)\Big(V^S\Big(n, \{x'_i(n;z',\varphi)\}_{i=1}^n; z',\varphi\Big) - V^S\Big(n, \{x_i\}_{i=1}^n; z,\varphi\Big)\Big) \\
& + \sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi)\Big(V^S\Big(n, \{x'_i(n;z,\varphi')\}_{i=1}^n; z,\varphi'\Big) - V^S\Big(n, \{x_i\}_{i=1}^n; z,\varphi\Big)\Big)\Bigg\}
\end{aligned}
$$

where $\setminus_-$ and $\cup_+$ are multiset difference and union operators.[44] The most important differences relative to the baseline model (equation (6)) have been highlighted in blue. Now, when a customer $i = 1,\ldots,n$ separates, the vector of promised utilities shrinks in cardinality and the customers that remain matched obtain the new promise $x'_i(n-1;z,\varphi)$. The firm attracts new buyers by offering a starting utility x'_0 to the entering customer. The promise-keeping constraint now reads:

$$\forall i = 1,\ldots,n: \quad x_i \le V^B(n, \boldsymbol{\omega}_i; z,\varphi)$$

for all $(z,\varphi) \in \mathcal{Z}\times\Phi$, establishing that the firm commits to each and every customer. We then solve for the optimal menu of contracts by solving for the joint surplus problem:

Proposition C.1 (Joint Surplus Problem with Price Discrimination) *In the economy with price discrimination, the seller's and the joint surplus problems are* equivalent*, in that:*

(i) Given a menu of contracts $\boldsymbol{\omega}_i = \{p_i, \mathbf{x}'_i(n';\mathbf{s}')\}$ for $i = 1,\ldots,n$ that maximize the seller's value subject to the promise-keeping constraint, $\{\mathbf{x}'_i(n';\mathbf{s}')\}_{i=1}^n$ maximizes:

$$W\big(n, \{x_i\}_{i=1}^n; z,\varphi\big) \equiv V^S\big(n, \{x_i\}_{i=1}^n; z,\varphi\big) + \sum_{i=1}^n x_i;$$

[44] These operators are defined by $\{a,b,b\}\setminus_-\{b\} = \{a,b\}$ and $\{a,b\}\cup_+\{b\} = \{a,b,b\}$, and they are needed here because different customers may be promised the same valuation.

(ii) Conversely, for every $\big\{\mathbf{x}_i'(n';\mathbf{s}')\big\}_{i=1}^n$ that maximizes $W\big(n,\{x_i\}_{i=1}^n;z,\varphi\big)$, there exists a menu of personalized price levels $\big\{p_i\big\}_{i=1}^n$ such that the collection $\big\{p_i,\mathbf{x}_i'(n';\mathbf{s}')\big\}_{i=1}^n$ constitutes a solution to the seller's problem.

Proof. The argument is conceptually similar to that of the baseline model (see Appendix B.1). Let $\big\{\overline{x}_0',\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n\big\}$ be a generic policy for the firm, with $\overline{\boldsymbol{\omega}}_i\equiv\big\{\overline{p}_i,\overline{\mathbf{x}}_i'(n';\mathbf{s}')\big\}$ and $\overline{\mathbf{x}}_i'(n';\mathbf{s}')=\big\{\overline{x}_i'(n+1;z,\varphi),\overline{x}_i'(n-1;z,\varphi),\{\overline{x}_i'(n;z',\varphi):z'\in\mathcal{Z}\},\{\overline{x}_i'(n;z,\varphi'):\varphi'\in\Phi\}\big\}$, for $i=1,\dots,n$. The firm's problem can be written as:

$$V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\equiv\max_{\overline{x}_0',\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n}\widetilde{V}^S\Big(n;\overline{x}_0',\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big)\ \text{ s.t. } x_i\le V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi),\forall i=1,\dots,n$$

where:

$$\begin{aligned}\widetilde{V}^S\Big(n;\overline{x}_0',\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big)\equiv\frac{1}{\rho(n;z,\varphi)}\Bigg[&\sum_{i=1}^n\overline{p}_i-\mathcal{C}(n;z,\varphi)\\&+\delta_c\sum_{j=1}^n V^S\Big(n-1,\big\{\overline{x}_i'(n-1;z,\varphi)\big\}_{i=1}^n\backslash_-\big\{\overline{x}_j'(n-1;z,\varphi)\big\};z,\varphi\Big)\\&+\eta\big(\theta(\overline{x}_0';\varphi)\big)V^S\Big(n+1,\big\{\overline{x}_i'(n+1;z,\varphi)\big\}_{i=1}^n\cup_+\big\{\overline{x}_0'\big\};z,\varphi\Big)\\&+\sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)V^S\Big(n,\big\{\overline{x}_i'(n;z',\varphi)\big\}_{i=1}^n;z',\varphi\Big)+\sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)V^S\Big(n,\big\{\overline{x}_i'(n;z,\varphi')\big\}_{i=1}^n;z,\varphi'\Big)\Bigg]\end{aligned}\tag{C.1}$$

is the value of the seller, with $\rho(n;z,\varphi)\equiv r+\delta_f+n\delta_c+\eta\big(\theta(\overline{x}_0';\varphi)\big)$ being the effective discount rate. The value of buyer $i=1,\dots,n$ under this policy is:

$$\begin{aligned}rV^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi)=v(\varphi)-p_i&+(\delta_f+\delta_c)\Big(U^B(\varphi)-V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi)\Big)\\&+(n-1)\delta_c\Big(\overline{x}_i'(n-1;z,\varphi)-V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi)\Big)+\eta\big(\theta(\overline{x}_0';\varphi)\big)\Big(\overline{x}_i'(n+1;z,\varphi)-V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi)\Big)\\&+\sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)\Big(\overline{x}_i'(n;z',\varphi)-V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi)\Big)+\sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\Big(\overline{x}_i'(n;z,\varphi')-V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi)\Big)\end{aligned}$$

Notice that the firm is re-optimizing after changing size. By monotonicity of preferences, the promise-keeping constraint will bind for each customer:

$$x_i=V^B(n,\overline{\boldsymbol{\omega}}_i;z,\varphi),\qquad\forall i=1,\dots,n$$

From this equation, we can solve for the promise-compatible price level to be charged to each customer under the policy $\Big\{\overline{x}_0',\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n\Big\}$:

$$\begin{aligned}p_i^{PK}\Big(\big\{\overline{x}_0',\{\mathbf{x}_j'(n';\mathbf{s}')\}_{j=1}^n\big\}\Big)=v(\varphi)&-\rho(n;z,\varphi)x_i+\delta_fU^B(\varphi)\\&+\delta_c\Big(U^B(\varphi)+(n-1)\overline{x}_i'(n-1;z,\varphi)\Big)+\eta\big(\theta(\overline{x}_0');\varphi\big)\overline{x}_i'(n+1;z,\varphi)\\&+\sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)\overline{x}_i'(n;z',\varphi)+\sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\overline{x}_i'(n;z,\varphi')\end{aligned}\tag{C.2}$$

Importantly, note that the price level for a specific customer i is independent of the distribution of utilities for all the *other* customers, that is:

$$p_i^{PK}\Big(\big\{\overline{x}'_0, \{\mathbf{x}'_j(n';\mathbf{s}')\}_{j\neq i}\big\} \cup_+ \{\mathbf{x}'_i(n';\mathbf{s}')\}\Big) = p_i^{PK}\Big(\big\{\overline{x}'_0, \{\mathbf{x}'_{\phi(j)}(n';\mathbf{s}')\}_{\phi(j)\neq i}\big\} \cup_+ \{\mathbf{x}'_i(n';\mathbf{s}')\}\Big)$$

for an arbitrary bisective function $\phi : \{1,\ldots,n\} \to \{1,\ldots,n\}$. Therefore, since the firm's problem internalizes the price level, the resulting maximization will be independent of the initial distribution of utilities. Indeed, plugging (C.2) into (C.1) and performing some straightforward algebra we obtain:

$$\begin{aligned}
\widetilde{W}\Big(n,\{x_i\}_{i=1}^n;\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big) \equiv \frac{1}{\rho(n;z,\varphi)}\Bigg[&n\Big(v(\varphi)+(\delta_f+\delta_c)U^B(\varphi)\Big) - \Big(\mathcal{C}(n;z,\varphi)+\eta\big(\theta(\overline{x}'_0;\varphi)\big)\overline{x}'_0\Big) \\
&+\delta_c\sum_{j=1}^n W\Big(n-1,\big\{\overline{x}'_i(n-1;z,\varphi)\big\}_{i=1}^n \backslash_- \big\{\overline{x}'_j(n-1;z,\varphi)\big\};z,\varphi\Big) \\
&+\eta\big(\theta(\overline{x}'_0;\varphi)\big)W\Big(n+1,\big\{\overline{x}'_i(n+1)\big\}_{i=1}^n \cup_+ \big\{\overline{x}'_0\big\};z,\varphi\Big) \\
+\sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)W\Big(n,\big\{\overline{x}'_i(n;z',\varphi)\big\}_{i=1}^n;z',\varphi\Big) &+ \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)W\Big(n,\big\{\overline{x}'_i(n;z,\varphi')\big\}_{i=1}^n;z,\varphi'\Big)\Bigg]
\end{aligned} \tag{C.3}$$

where we have defined:

$$\widetilde{W}\Big(n,\{x_i\}_{i=1}^n;\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big) \equiv \widetilde{V}^S\Big(n;\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big) + \sum_{i=1}^n x_i$$

as the joint surplus under policy $\big\{\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n\big\}$, and:

$$W\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big) \equiv \max_{\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n} \widetilde{W}\Big(n,\{x_i\}_{i=1}^n;\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big)$$

as the *maximized* joint surplus. Importantly, it is still the case that the right-hand side of (C.3) does not depend on the initial distribution of utilities $\{x_i\}_{i=1}^n$, nor the distribution of prices $\{p_i\}_{i=1}^n$, so we can write the joint surplus under a given policy as:

$$\widetilde{W}\Big(n,\{x_i\}_{i=1}^n;\overline{x}'_0,\{\overline{\boldsymbol{\omega}}_i\}_{i=1}^n;z,\varphi\Big) = \widetilde{W}_n\Big(\overline{x}'_0,\big\{\overline{\mathbf{x}}'_i(n';\mathbf{s}')\big\}_{i=1}^n;z,\varphi\Big)$$

This allows us to break up the optimal contracting problem into two separate stages. Where we denote an optimal contract by $\big\{x_0'^*,\{p_i^*,\mathbf{x}_i'^*(n';\mathbf{s}')\}_{i=1}^n\big\}$, we have:

$$\begin{aligned}
\big\{x_0'^*,\{\mathbf{x}_i'^*(n';\mathbf{s}')\}_{i=1}^n\big\} &= \arg\max\ \widetilde{W}_n\Big(\overline{x}'_0,\big\{\overline{\mathbf{x}}'_i(n';\mathbf{s}')\big\}_{i=1}^n;z,\varphi\Big) \\
p_i^* &= p_i^{PK}\Big(\Big\{\overline{x}_0'^*,\{\mathbf{x}_j'^*(n';\mathbf{s}')\}_{j=1}^n\Big\}\Big), \qquad \forall i=1,\ldots,n
\end{aligned}$$

Thus, the joint surplus and the seller's problems are equivalent. □

The characterization of the equilibrium is also similar to the baseline model. First, by utility-invariance of the joint surplus we can write $W_n(z,\varphi) = W\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)$, $\forall(n,z,\varphi)\in\mathbb{N}\times\mathcal{Z}\times\Phi$. Under the optimal policy, the joint surplus problem can then be written as follows:

$$
\begin{aligned}
(r+\delta_f)W_n(z,\varphi) = \max_{x_0', \{x_j'(n-1;z,\varphi)\}_{j=1}^n} \Bigg\{ & n\Big(v(\varphi)+(\delta_f+\delta_c)U^B(\varphi)\Big) - \Big(\mathcal{C}(n;z,\varphi)+\eta\big(\theta(x_0';\varphi)\big)x_0'\Big) \\
& + n\delta_c\Big(W_{n-1}(z,\varphi)-W_n(z,\varphi)\Big) + \eta\big(\theta(x_0';\varphi)\big)\Big(W_{n+1}(z,\varphi)-W_n(z,\varphi)\Big) \\
& + \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)\Big(W_n(z',\varphi)-W_n(z,\varphi)\Big) + \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\Big(W_n(z,\varphi')-W_n(z,\varphi)\Big)\Bigg\}
\end{aligned}
$$

The optimality condition for x_0' is:

$$
\frac{\partial\eta\big(\theta(x_0';\varphi)\big)}{\partial x_0'}\Big(W_{n+1}(z,\varphi)-W_n(z,\varphi)\Big) = \frac{\partial\eta\big(\theta(x_0';\varphi)\big)}{\partial x_0'}x_0' + \eta\big(\theta(x_0';\varphi)\big)
$$

similar to equation (12). By the usual arguments, the solution to this equation is *unique*. Denoting the solution by $x_{0,n}'(z,\varphi)$, the law of motion for the firm of size n_t at time t is then:

$$
n_{t+\Delta}-n_t = \begin{cases} 1 & \text{w/prob. } \eta\Big(\theta\big(x_{0,n_t}'(z,\varphi)\big)\Big)\Delta + o(\Delta) \\ -1 & \text{w/prob. } n_t\delta_c\Delta + o(\Delta) \\ -n_t & \text{w/prob. } \delta_f\Delta + o(\Delta) \\ 0 & \text{else} \end{cases}
$$

Therefore, we have found:

Corollary C.1 (Firm dynamics with price discrimination) *In the model with price discrimination, the equilibrium features unique predictions for firm dynamics.*

Although firm growth is pinned down uniquely, prices are not. Since any distribution of utilities (and thus prices) among incumbents is compatible with optimality, there is now a multiplicity of contracts that can be sustained in the optimal allocation.[45] This is stated formally in the following proposition:

Proposition C.2 (Price Indeterminacy) *There is a continuum of contracts* $\big\{p_i^*, \mathbf{x}_i'^*(n';\mathbf{s}')\big\}_{i=1}^n$ *that: (i) maximize the joint surplus, and (ii) leave both the buyers and the seller indifferent.*

Proof. Pick $\varepsilon\in\mathbb{R}$ arbitrarily. The goal of the proof is to show that there is some $\beta_n(\varphi)>0$ (possibly a function of size and the aggregate state) for which, if a given contract with $\boldsymbol{\omega}^b = \big\{p_i+\varepsilon\beta_n(\varphi), \mathbf{x}_i'(n';\mathbf{s}')+\varepsilon\big\}_{i=1}^n$ is optimal, then each customer and the seller maximize their value under contract $\boldsymbol{\omega}^a = \big\{p_i, \mathbf{x}_i'(n';\mathbf{s}')\big\}_{i=1}^n$. The value of contract $\boldsymbol{\omega}_i^b$ for customer $i=1,\ldots,n$ is:

$$
\begin{aligned}
rV^B(n,\boldsymbol{\omega}_i^b;z,\varphi) = \; & v(\varphi)-p_i-\varepsilon\beta_n(\varphi)+(\delta_f+\delta_c)\Big(U^B(\varphi)-V^B(n,\boldsymbol{\omega}_i;z,\varphi)\Big) \\
& + (n-1)\delta_c\Big(x_i'(n-1;z,\varphi)+\varepsilon-V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big) + \eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big(x_i'(n+1;z,\varphi)+\varepsilon-V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big) \\
& + \sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)\Big(x_i'(n;z',\varphi)+\varepsilon-V^B(n,\boldsymbol{\omega}_i;z,\varphi)\Big) + \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\Big(x_i'(n;z,\varphi')+\varepsilon-V^B(n,\boldsymbol{\omega}_i;z,\varphi)\Big) \\
= \; & v(\varphi)-p_i-\varepsilon\Big(\beta_n(\varphi)-(n-1)\delta_c-\eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big)+(\delta_f+\delta_c)\Big(U^B(\varphi)-V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big)
\end{aligned}
$$

[45] Additional equilibria may exist outside of the Markov-perfect equilibrium class. Here we only point out that equilibrium uniqueness is lost in Markov Perfect equilibria when sellers can price-discriminate.

$$
\begin{aligned}
&+ (n-1)\delta_c\Big(x_i'(n-1;z,\varphi) - V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big) + \eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big(x_i'(n+1;z,\varphi) - V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big)\\
&+ \sum_{z'\in\mathcal{Z}} \lambda_z(z'|z)\Big(x_i'(n;z',\varphi) - V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big) + \sum_{\varphi'\in\Phi} \lambda_\varphi(\varphi'|\varphi)\Big(x_i'(n;z,\varphi') - V^B(n,\boldsymbol{\omega}_i^b;z,\varphi)\Big)\\
= {}& rV^B(n,\boldsymbol{\omega}_i^a;z,\varphi) - \varepsilon\Big(\beta_n(\varphi) - (n-1)\delta_c - \eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big)
\end{aligned}
$$

where we have used $\sum_{z'} \lambda_z(z'|z) = \sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) = 0$. Thus, $V^B(n,\boldsymbol{\omega}_i^a) = V^B(n,\boldsymbol{\omega}_i^b)$ if, and only if:

$$
\beta_n(\varphi) = (n-1)\delta_c + \eta\big(\theta(x_0'(\varphi);\varphi)\big) \tag{C.4}
$$

As for the seller's value, note that:

$$
\begin{aligned}
rV^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big) = {}& \max_{x_0'(\varphi),\{\boldsymbol{\omega}_i\}_{i=1}^n} \left\{ \sum_{i=1}^n p_i + n\varepsilon\beta_n(\varphi) - \mathcal{C}(n;z,\varphi) + \delta_f\Big(V_0^S(\varphi) - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\right.\\
&+ \delta_c\sum_{j=1}^n\Big(V^S\Big(n-1,\{x_i'(n-1;z,\varphi)+\varepsilon\}_{i=1}^n \setminus_- \{x_j'(n-1;z,\varphi)+\varepsilon\};z,\varphi\Big) - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\\
&+ \eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big(V^S\Big(n+1,\{x_i'(n+1;z,\varphi)+\varepsilon\}_{i=1}^n \cup_+ \{x_0'(\varphi)\};z,\varphi\Big) - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\\
&+ \sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)\Big(V^S\Big(n,\{x_i'(n;z',\varphi)+\varepsilon\}_{i=1}^n;z',\varphi\Big) - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\\
&\left.+ \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\Big(V^S\Big(n,\{x_i'(n;z,\varphi')+\varepsilon\}_{i=1}^n;z,\varphi'\Big) - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\right\}\\
= {}& \max_{x_0'(\varphi),\{\boldsymbol{\omega}_i\}_{i=1}^n} \left\{ \sum_{i=1}^n p_i + n\varepsilon\beta_n(\varphi) - \mathcal{C}(n;z,\varphi) + \delta_f\Big(V_0^S(\varphi) - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\right.\\
&+ \delta_c\sum_{j=1}^n\Big(W_{n-1}(z,\varphi) - \sum_{i\neq j} x_i'(n-1;z,\varphi) - (n-1)\varepsilon - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\\
&+ \eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big(W_{n+1}(z,\varphi) - \sum_{i=1}^n x_i'(n+1;z,\varphi) - x_0' - n\varepsilon - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\\
&+ \sum_{z'\in\mathcal{Z}}\lambda_z(z'|z)\Big(W_n(z',\varphi) - \sum_{i=1}^n x_i'(n;z',\varphi) - n\varepsilon - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\\
&\left.+ \sum_{\varphi'\in\Phi}\lambda_\varphi(\varphi'|\varphi)\Big(W_n(z,\varphi') - \sum_{i=1}^n x_i'(n;z,\varphi') - n\varepsilon - V^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big)\Big)\right\}\\
= {}& rV^S\Big(n,\{x_i\}_{i=1}^n;z,\varphi\Big) + n\varepsilon\Big(\beta_n(\varphi) - (n-1)\delta_c - \eta\big(\theta(x_0'(\varphi);\varphi)\big)\Big)
\end{aligned}
$$

where we have used the definition of W in the second equality. Thus, by equation (C.4), the seller is also indifferent. In sum, contract $\{\boldsymbol{\omega}_i^a\}_{i=1}^n$ is optimal if, and only if, $\{\boldsymbol{\omega}_i^b\}_{i=1}^n$ is optimal. Generally, there is a continuum of optimal contracts, indexed by ε. □

In fact, the Gourio and Rudanko (2014b) pricing strategy can be seen as one such equilibrium, in which:

$$
x_i'(n;z,\varphi) = U^B(\varphi), \quad \forall i = 1,\dots,n, \ \forall (n,z,\varphi) \in \mathbb{N}\times\mathcal{Z}\times\Phi
$$

which, by equation (C.2), implies that $p_i = v$, $\forall i = 1,\dots,n$. In this case, x_0' plays the role of a price discount relative to valuation for the incoming customer, who will be charged a fixed price of v thereafter.

Firm Dynamics and Pricing under Customer Capital Accumulation

by

Pau Roldan-Blanco

Sonia Gilbukh

Online Appendix

D Data Description

The IRI scanner dataset spans a period of 12 years (from the first week of January 2001 to the last week of December 2012), and contains revenue and quantity information for over 5,000 retail (drug and grocery) stores over 50 Metropolitan Statistical Areas (MSA) in the U.S. The data are automatically generated by retailers through their point-of-sale systems. Products, at the UPC level, are grouped into broad categories. We narrow our attention to two large geographical markets (New York and Los Angeles) in the period 2001-2007, and consider 15 product categories.[46] We back out the weekly average price (P) from revenues (R) and quantities (Q):

$$P_{usm,t} = \frac{R_{usm,t}}{Q_{usm,t}}$$

for each UPC u within week t, in store s and (geographic) market m. Throughout, we restrict our sample to products that are commonly available across stores and not only sold in specific establishments. Given the overall number of stores in our sample, we drop goods sold in less than 10 stores on any week×market. To eliminate outliers, we drop stores with non-positive sales, transactions with prices above \$100 (accounting for the top .02% of the price distribution in the full sample), and cases with multiple observations at the store×market×week×UPC level, which we deem as mis-reported transactions. Finally, in the absence of a theory of price discounts, we focus only on *regular* prices by filtering out of the sample products on sale. A convenient feature of the data is that products are flagged whenever they go on promotion, which means that we need not employ a filtering algorithm as in Nakamura and Steinsson (2008) but we can rather exclude flagged products directly.[47]

	Full Sample	**Sub-Sample**
Number of product categories	15	15
Number of chains	64	64
Number of stores	278	278
Number of UPCs	19,721	11,483
Stores per chain (average)	27	26
Stores per product (average)	59	88
Products per store (average)	4,180	3,638
Average price (USD)	7.75	8.32
Price dispersion	15.73%	10.55%
Total sales (Billion USD)	2.86	1.60
Number of weeks	365	365
MSAs considered	NY, LA	NY, LA
Number of observations	89,112,170	59,813,217

Table D.1: Descriptive statistics before and after sample selection. Source: IRI data.

Table D.1 shows some descriptive statistics of our data before and after applying these restrictions. Price dispersion is measured as the average standard deviation of $\widehat{p}_{usm,t}$ (equation (25)) across products,

[46] The categories are: Beer, Blades, Carbonated Beverages, Cigarettes, Coffee, Cold Cereal, Deodorant, Diapers, Frozen Pizza, Frozen Dinners, Household Cleaners, Hotdogs, Laundry Detergent, Margarine and Butter, and Mayonnaise.

[47] A "promotion" is defined by the IRI as a temporary price reduction of 5% or greater. Sales are quite unresponsive to the business cycle, as documented by Coibion *et al.* (2015) for the IRI data, and therefore excluding should not change our life-cycle results significantly.

stores, markets, and time. In our full sample, dispersion at the barcode level is high (15.73%), in line with previous studies using similar micro pricing data from different sources (e.g. Kaplan and Menzio (2015)). The restricted sample has a lower dispersion (10.55%), as a result of having eliminated price outliers and uncommon goods. Figure D.1 shows the distribution of relative prices in our sample, alongside that of normalized sales (the ratio of store-level sales to its mean) and store sales growth rates. Table D.2 presents summary statistics for these distributions. We see that the store size distribution has a fat right tail, which accounts for the high dispersion in normalized sales.

	Relative prices	**Normalized sales**	**Sales growth**
Mean	0	1	−.0008
Median	.0009	.9087	−.0003
Percentiles			
1st	−.3257	.2729	−.1905
25^{th}	−.0415	.6656	−.0378
75^{th}	.0486	1.2281	.0373
90^{th}	.1097	1.7029	.0765
99^{th}	.2809	2.3819	.1782
Dispersion			
St. Dev.	.1055	.4744	.0694
p90-p10 range	1.2504	3.6163	1.1661
p90-p50 range	1.1149	1.8740	1.0798
p50-p10 range	1.1215	1.9297	1.0799

Table D.2: Distribution of relative prices, normalized store sales (i.e. the ratio of total dollar sales within the store to average sales across products sold in the store), and annualized sales growth rates, across the whole 2001-2007 sample. Source: IRI weekly data.

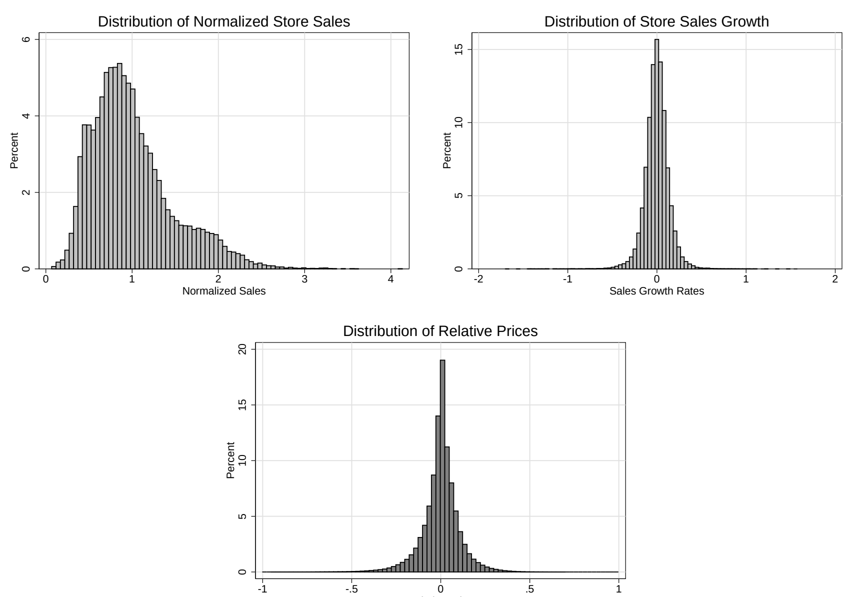

Figure D.1: Distribution of normalized store sales (top), store sales growth rates (middle), and relative prices at the UPC level (bottom) in our final sample. Source: IRI weekly data.

E Numerical Appendix

E.1 Numerical Implementation of the Exogenous Processes

This appendix shows how to parametrize and estimate (z, φ) as continuous-time Markov chain (CTMC) processes. The same identical structure applies to both shocks, so let us consider for instance the idiosyncratic shock (z).

The $k_z \times k_z$ infinitesimal generator matrix $\mathbf{\Lambda}_z$ to be estimated is:

$$\mathbf{\Lambda}_z = \begin{pmatrix} -\sum_{j \neq 1} \lambda_{1j} & \lambda_{12} & \dots & \lambda_{1k_z} \\ \lambda_{21} & -\sum_{j \neq 2} \lambda_{2j} & \dots & \lambda_{2k_z} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{k_z 1} & \lambda_{k_z 2} & \dots & -\sum_{j \neq k_z} \lambda_{k_z j} \end{pmatrix}$$

where λ_{ij} is short-hand for $\lambda_z(z_j|z_i)$, $z_i, z_j \in \mathcal{Z}$. Since this level of generality would require the estimation of a large number $k_z(k_z - 1)$ of transition rates, we reduce the parameter space by specializing the CTMC as follows:

- First, we assume z follows a driftless Ornstein-Uhlenbeck (OU) process in logs. An OU process is a type of mean-reverting and autoregressive CTMC which can be loosely viewed as the continuous-time analogue of an AR(1). Formally:

$$\mathrm{d} \log z_t = -\rho_z \log z_t \mathrm{d}t + \sigma_z \mathrm{d}B_t$$

  where B_t is a standard Brownian motion, and $\rho_z, \sigma_z > 0$ are parameters.

- Operationally, in the numerical version of the model in which time is partitioned and takes values in $\mathbb{T} = \{\Delta, 2\Delta, 3\Delta, \dots\}$, we use the *Euler-Maruyama method*, that is:

$$\log z_k = (1 - \rho_z \Delta) \log z_{k-1} + \sigma_z \sqrt{\Delta} \varepsilon_k^z, \qquad \varepsilon_k^z \sim iid\ \mathcal{N}(0,1) \tag{E.1}$$

  for each $k \in \mathbb{T}$. This is now an AR(1) processes with autocorrelation $\widetilde{\rho}_z \equiv 1 - \rho_z \Delta$ and variance $\dfrac{\sigma_z^2}{\rho_z(1 + \widetilde{\rho}_z)}$. Thus, $\rho_z > 0$ can be seen as a measure of mean-reversion, with *lower* values corresponding to *higher* persistence.

- The discrete-time process (E.1) is estimated using the Tauchen (1986) with a discrete-state Markov chain that we define on the theoretical grid, $\mathcal{Z}$. The outcome of this method are estimates for (ρ_z, σ_z), and a transition probability matrix $\mathbf{\Pi}_z = \big(\pi_{ij}\big)$, where π_{ij} denotes the probability of a z_i-to-z_j transition in the $\mathbb{T}$ space.

- For the mapping back into continuous time, we use the fact that, for small enough $\Delta > 0$, transition *probabilities* are well approximated by transition *rates* in the following sense:

$$\forall i = 1, \dots, k_z : \quad \pi_{ij} \approx \lambda_{ij} \Delta, \forall j \neq i \quad \text{and} \quad \pi_{ii} \approx 1 - \sum_{j \neq i} \lambda_{ij} \Delta$$

E.2 Stationary Solution Algorithm

To solve for the stationary equilibrium, we implement the following nested procedure:

- First, we solve the maximization of the joint surplus function using a value function iteration (VFI) algorithm, under a guess for U^B.
- To update U^B, we must check that the free entry condition is satisfied. Combining equations (3) and (13), we can write the equilibrium free entry condition as:

$$\kappa(\varphi) = \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \left\{ \eta \circ \mu^{-1} \left(\frac{\Gamma^B(\varphi)}{x_1'(z_0,\varphi) - U^B(\varphi)} \right) \Big(W_1(z_0,\varphi) - x_1'(z_0,\varphi) \Big) \right\}$$

  To find U^B, we use a bisection method: increase (or decrease) U^B if there are too many (or too few) entering firms.

Throughout, the state space grid is fixed at $\mathcal{N} \times \mathcal{Z} \times \Phi$, where $\mathcal{N} = \{1, \ldots, \bar{n}\}$, $\mathcal{Z} = \{z_i\}_{i=1}^{k_z}$, and $\Phi = \{\varphi_j\}_{j=1}^{k_\varphi}$. The following describes the steps of the algorithm:

Step 1. Set the counter to $k = 0$. Choose guesses $\underline{U}^{(0)}(\varphi)$ and $\overline{U}^{(0)}(\varphi) \gg \underline{U}^{(0)}(\varphi)$ for each $\varphi \in \Phi$. Set the value of inactivity to:

$$U^{B(0)}(\varphi) = \frac{1}{2}\left(\underline{U}^{(0)}(\varphi) + \overline{U}^{(0)}(\varphi) \right)$$

Step 2. For any given $k \in \mathbb{N}$ and $n \in \mathcal{N}$, use VFI to find the fixed point $W_n^{(k)}(z,\varphi)$ of:

$$\begin{aligned}
(r+\delta_f)W_n^{(k)}(z,\varphi) = \; & n\Big(v(\varphi) + (\delta_f + \delta_c)U^{B(k)}(\varphi)\Big) - \mathcal{C}(n; z, \varphi) \\
& + n\delta_c \Big(W_{n-1}^{(k)}(z,\varphi) - W_n^{(k)}(z,\varphi) \Big) \\
& + \max_{x'_{n+1}} \left\{ \eta \circ \mu^{-1} \left(\frac{\Gamma^{B(k)}(\varphi)}{x'_{n+1}(z,\varphi) - U^{B(k)}(\varphi)} \right) \Big(W_{n+1}^{(k)}(z,\varphi) - W_n^{(k)}(z,\varphi) - x'_{n+1} \Big) \right\} \\
& + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \Big(W_n^{(k)}(z',\varphi) - W_n^{(k)}(z,\varphi) \Big) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \Big(W_n^{(k)}(z,\varphi') - W_n^{(k)}(z,\varphi) \Big)
\end{aligned}$$

where $\Gamma^{B(k)} = c(\varphi) + rU^{B(k)}(\varphi) - \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \Big(U^{B(k)}(\varphi') - U^{B(k)}(\varphi) \Big)$. Store the corresponding policy functions: $\Big\{ x_{n+1}'^{(k)}(z,\varphi) : (n,z,\varphi) \in \mathcal{N} \times \mathcal{Z} \times \Phi \Big\}$.

Step 3. For each $\varphi \in \Phi$, compute the object:

$$\Delta^{(k)}(\varphi) \equiv \kappa(\varphi) - \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \left\{ \eta \circ \mu^{-1} \left(\frac{\Gamma^{B(k)}(\varphi)}{x_1'^{(k)}(z_0,\varphi) - U^{B(k)}(\varphi)} \right) \Big(W_1^{(k)}(z_0,\varphi) - x_1'^{(k)}(z_0,\varphi) \Big) \right\}$$

Stop if $\Delta^{(k)}(\varphi) \in [-\varepsilon, \varepsilon]$, $\forall \varphi \in \Phi$, for some small $\varepsilon > 0$. Otherwise, set:

$$U^{B(k+1)}(\varphi) = \frac{1}{2}\left(\underline{U}^{(k+1)}(\varphi) + \overline{U}^{(k+1)}(\varphi) \right)$$

for each $\varphi \in \Phi$, where:

(a) If $\Delta^{(k)}(\varphi) > \varepsilon$, then $\underline{U}^{(k+1)}(\varphi) = \underline{U}^{(k)}(\varphi)$ and $\overline{U}^{(k+1)}(\varphi) = U^{B(k)}(\varphi)$;

(b) If $\Delta^{(k)}(\varphi) < -\varepsilon$, then $\underline{U}^{(k+1)}(\varphi) = U^{B(k)}(\varphi)$ and $\overline{U}^{(k+1)}(\varphi) = \overline{U}^{(k)}(\varphi)$;

and go back to Step 2. with $[k] \leftarrow [k+1]$.

The VFI algorithm of Step 2 is guaranteed to converge because, given a U^B, the joint surplus is a contraction and therefore has a unique fixed point. (For a proof, see Appendix G.3).

F Model Extensions

F.1 Micro-Foundation for Marginal Utility (v)

Relative to the baseline model, we assume that there exists a representative household, comprised of a measure-one continuum of identical buyers, with quasilinear preferences:

$$U_t(\varphi_t) = \mathbb{E}_t \left\{ \int_t^{+\infty} e^{-r(t-\tau)} \Big(a_\tau + \ln C_\tau(\varphi_\tau) \Big) \mathrm{d}\tau \right\}$$

where $r > 0$ is the discount rate, and a_τ is consumption of a numeraire good at time $\tau \geq t$. Household consumption C is a bundle of output levels of a continuum of sellers via the CES aggregator:[48]

$$C(\varphi) = \int q(\varphi) c(z, \varphi) \mathrm{d}\lambda_f(z) \tag{F.1.1}$$

where $\lambda_f(z)$ is the distribution of active firms over the space of idiosyncratic states $z \sim \lambda_z(z'|z)$, and $c(z, \varphi)$ is consumption from firm z in aggregate state $\varphi \sim \lambda_\varphi(\varphi'|\varphi)$. Equation (F.1.1) assumes that the goods of the different firms are perfect substitutes, so we can interpret the continuum of sellers as effectively selling the same product. The shifter $q(\varphi)$ is the aggregate demand shock, acting as a shock to preferences. As the numeraire is separable, the marginal utility of buyers is simply the marginal rate of substitution between the numeraire good and aggregate consumption, $v(\varphi) \equiv \frac{1}{C(\varphi)}$.

Aggregate consumption C is obtained by the household at a price of P. The cost minimization problem of the household leads to the inverse demand function for type-z products:

$$c(z, \varphi) = \frac{q(\varphi)}{p(z, \varphi)} P(\varphi) C(\varphi), \qquad \text{with } P(\varphi) = \left(\int q(\varphi) \left(\frac{p(z, \varphi)}{q(\varphi)} \right)^{-1} \mathrm{d}\lambda_f(z) \right)^{-1}$$

where $p(z, \varphi)$ is the price that seller z sets for its product in aggregate state φ. Sellers are exactly as in the baseline model, i.e. they make instantaneous profits $\pi(n; z, \varphi) = p(n; z, \varphi)n - \mathcal{C}(n; z, \varphi)$, where $n \in \mathbb{Z}_+$ is the number of customers and $p(n; z, \varphi)$ is the price policy function. Therefore, the measure of firms in state z is $\lambda_f(z) = \sum_{n \in \mathbb{N}} S_n(z)$. By feasibility, we have $C(\varphi) = q(\varphi) \sum_{n \in \mathbb{N}} \sum_{z \in \mathcal{Z}} c_n(z, \varphi) S_n(z)$. As sellers sell an indivisible product, $c_n(z, \varphi) = n$, so:

$$C(\varphi) = q(\varphi) \sum_{n \in \mathbb{N}} \sum_{z \in \mathcal{Z}} n S_n(z, \varphi) = q(\varphi) B^A$$

Therefore, we have micro-founded the individual valuation of buyers to be $v(\varphi) = \frac{1}{q(\varphi) B^A}$. Intuitively, for a given aggregate shock, individual consumers value the product less when there are more active buyers, because that means that the product market is more congested.

[48] In keeping with the notation convention introduced in Section 2.4, if a variable depends explicitly on φ then it is a *jump* variable. Otherwise, the variable evolves smoothly.

F.2 Endogenous Customer Separations

To introduce customer seller-to-seller transitions, we can model customer search explicitly. While we assume that there is still an exogenous risk $\delta_c > 0$ of separation for each customer, additionally we now add the possibility that customers search, and potentially endogenously separate, while on the match. We assume that active buyers do not face a cost of search, as they do not discontinue their consumption when transitioning from one seller to the another.

Introducing this additional dimension into our full model with aggregate shocks is not at all straightforward. Endogenous buyer transitions across sellers would break the ex-ante indifference condition among inactive buyers, which in our baseline setting is key to pin down equilibrium market tightness. In order to preserve the block-recursive structure, one remedy would be to assume free entry across all markets on the seller side. This would change the environment substantially, so we leave it for future work.

Thus, suppose there are no aggregate shocks. The problem of an active buyer with value V^B is:

$$\max_{x\in[V^B,\overline{x}]} \; \mu(\theta(x))\big(x - V^B\big)$$

Note that the matched buyer only considers offers that deliver an expected value that weakly dominates the current perceived utility, V^B. Let $\widehat{x}(n,\boldsymbol{\omega};z)$ be the policy for a customer in a firm of type (n,z) under contract $\boldsymbol{\omega}$. The first-order condition reads:

$$\Big(\widehat{x}(n,\boldsymbol{\omega};z) - V^B(n,\boldsymbol{\omega};z)\Big)\frac{\partial\mu(\theta(x))}{\partial x}\bigg|_{x=\widehat{x}(n,\boldsymbol{\omega};z)} = -\mu\Big(\theta\big(\widehat{x}(n,\boldsymbol{\omega};z)\big)\Big) \tag{F.2.1}$$

Intuitively, the inactive buyer trades off the expected option value of transitioning (left-hand side) to the rate at which this offer can be obtained (right-hand side). Since we focus on equilibria in which market tightness is an increasing function of promised utilities, it is not difficult to show (e.g. Shi (2009)) that $\widehat{x}(n,\boldsymbol{\omega};z)$ is increasing in $V^B(n,\boldsymbol{\omega};z)$. In words, the more profitable a match is ex-post, the higher the offer for which the customer will apply next. Therefore, customers separate according to their initial state, and climb up the utility ladder. This effect tends to shift the mass of customers (and therefore sellers) toward higher promised utilities, and thus acts as a countervailing force to the equilibrium dynamics of the baseline model: when the sellers offering the worst terms of trade lose customers, they need to start setting up more favorable contracts.

The risk of endogenously losing customers must now be incorporated into the pricing decisions of sellers. The buyers' and seller's HJB equations are then identical to (5) and (6), respectively, except that we now must replace δ_c by an "effective" customer separation rate, given by:

$$\widehat{\delta}_c(n,\boldsymbol{\omega};z) \equiv \delta_c + \mu\Big(\theta\big(\widehat{x}(n,\boldsymbol{\omega};z)\big)\Big)$$

Likewise, the market tightness must incorporate that the pool of searching buyers is composed of both inactive as well as active buyers:

$$\theta_n(z) = \frac{1}{S_{n-1}(z)}\Big(B^I_n(z) + B^A_{\iota(n)}(z)\Big)$$

for any $n \geq 1$, where $\iota(n) \in \mathbb{N}$ is the size of the seller that a customer seeking to transition to a size-n seller is currently matched with, i.e. the solution to $x_n(z) = \widehat{x}(\iota(n),\boldsymbol{\omega};z)$.

G Additional Theoretical Results

G.1 Computing the Aggregate Stationary Measures of Agents

To derive aggregate measures, we first must derive the equilibrium *shares* of agent types. Throughout, we fix $\varphi \in \Phi$. Let $g_{n,t}(z) \equiv \frac{S_{n,t}(z)}{\mathcal{S}_t}$, where $\mathcal{S}_t \equiv \sum_{n\geq 1}\sum_z S_{n,t}(z)$ is the total measure of incumbents. After a period of size $\Delta > 0$, the share of firms of size $n \geq 2$ becomes:

$$\begin{aligned} g_{n,t+\Delta}(z) =& \Big[\eta\big(\theta_{n,t+\Delta}(z,\varphi)\big)\Delta + o(\Delta)\Big] g_{n-1,t}(z) + (n+1)\Big[\delta_c\Delta + o(\Delta)\Big] g_{n+1,t}(z) \\ &+ \sum_{\tilde{z}\neq z}\Big[\lambda_z(z|\tilde{z})\Delta + o(\Delta)\Big] g_{n,t}(\tilde{z}) + \Big[1 - \delta_f\Delta - n\delta_c\Delta - \eta\big(\theta_{n+1,t+\Delta}(z,\varphi)\big)\Delta - \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)\Delta + o(\Delta)\Big] g_{n,t}(z) \end{aligned} \tag{G.1.1}$$

Subtracting $g_{n,t}(z)$ from both sides of equation (G.1.1) and dividing by Δ gives:

$$\begin{aligned} \frac{g_{n,t+\Delta}(z) - g_{n,t}(z)}{\Delta} =& \Big[\eta\big(\theta_{n,t+\Delta}(z,\varphi)\big) + \frac{o(\Delta)}{\Delta}\Big] g_{n-1,t}(z) + (n+1)\Big[\delta_c + \frac{o(\Delta)}{\Delta}\Big] g_{n+1,t}(z) \\ &+ \sum_{\tilde{z}\neq z}\Big[\lambda_z(z|\tilde{z}) + \frac{o(\Delta)}{\Delta}\Big] g_{n,t}(\tilde{z}) - \Big[\delta_f + n\delta_c + \eta\big(\theta_{n+1,t+\Delta}(z,\varphi)\big) + \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z) + \frac{o(\Delta)}{\Delta}\Big] g_{n,t}(z) \end{aligned}$$

Taking the limit as $\Delta \to 0$:

$$\begin{aligned} \boldsymbol{\partial}_t g_{n,t}(z) =& \eta\big(\theta_{n,t}(z,\varphi)\big) g_{n-1,t}(z) + (n+1)\delta_c g_{n+1,t}(z) \\ &+ \sum_{\tilde{z}\neq z}\lambda_z(z|\tilde{z}) g_{n,t}(\tilde{z}) - \Big(\delta_f + n\delta_c + \eta\big(\theta_{n+1,t}(z,\varphi)\big) + \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)\Big) g_{n,t}(z) \end{aligned}$$

Similarly, when $n = 1$ we have:

$$\begin{aligned} g_{1,t+\Delta}(z) =& \Big[\pi_z(z)\eta\big(\theta_{1,t+\Delta}(z,\varphi)\big)\Delta + o(\Delta)\Big]\frac{S_{0,t}(\varphi)}{\mathcal{S}_t} + 2\Big[\delta_c\Delta + o(\Delta)\Big] g_{2,t}(z) \\ &+ \sum_{\tilde{z}\neq z}\Big[\lambda_z(z|\tilde{z})\Delta + o(\Delta)\Big] g_{1,t}(\tilde{z}) + \Big[1 - \delta_f\Delta - \delta_c\Delta - \eta\big(\theta_{2,t+\Delta}(z,\varphi)\big)\Delta - \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)\Delta + o(\Delta)\Big] g_{1,t}(z) \end{aligned} \tag{G.1.2}$$

A similar derivation on (G.1.2) shows that, for $n = 1$,

$$\begin{aligned} \boldsymbol{\partial}_t g_{1,t}(z) =& \pi_z(z)\eta\big(\theta_{1,t}(z,\varphi)\big)\frac{S_{0,t}(\varphi)}{\mathcal{S}_t} + 2\delta_c g_{2,t}(z) \\ &+ \sum_{\tilde{z}\neq z}\lambda_z(z|\tilde{z}) g_{1,t}(\tilde{z}) - \Big(\delta_f + \delta_c + \eta\big(\theta_{2,t}(z,\varphi)\big) + \sum_{\tilde{z}\neq z}\lambda_z(\tilde{z}|z)\Big) g_{1,t}(z) \end{aligned}$$

It remains to show the law of motion for the measure of potential entrants, $S_{0,t}(\varphi)$. In this case, for given φ, we have:

$$S_{0,t+\Delta}(\varphi) = \Big[\delta_f\Delta + o(\Delta)\Big]\mathcal{S}_t + \Big[\delta_c\Delta + o(\Delta)\Big]\sum_z S_{1,t}(z) + \Big[1 - \sum_{z_0}\pi_z(z_0)\eta\big(\theta_{1,t+\Delta}(z_0,\varphi)\big)\Delta + o(\Delta)\Big] S_{0,t}(\varphi)$$

Taking the continuous-time limit in the usual way, we arrive at:

$$\boldsymbol{\partial}_t S_{0,t}(\varphi) = \Big(\delta_f + \delta_c \sum_z g_{1,t}(z)\Big)\mathcal{S}_t - \sum_{z_0} \pi_z(z_0)\eta\big(\theta_{1,t}(z_0,\varphi)\big)S_{0,t}(\varphi)$$

In the stationary solution, $\boldsymbol{\partial}_t g_{n,t}(z) = 0$ and $\boldsymbol{\partial}_t S_{0,t}(\varphi) = 0$. Then, we obtain a system of second-order equations which can be solved numerically on the state-space grid, $\mathcal{N} \times \mathcal{Z} \times \Phi$. This will yield a solution for the matrix $\{g_n(z)\}_{n,z}$, and the share of potential entrants per incumbent firm, $h_0(\varphi) \equiv S_0(\varphi)/\mathcal{S}$.

To compute aggregates, use (17) to obtain $b_n^A(z) \equiv \frac{B_n^A(z)}{\mathcal{S}}$ by:

$$b_n^A(z) = n g_n(z)$$

Then, $b^A \equiv B^A/\mathcal{S} = \sum_{n=1}^{+\infty}\sum_z n g_n(z)$. On the other hand, from equation (16) we know that $B_n^I(z,\varphi) = \mathcal{S}\theta_n(z,\varphi)g_{n-1}(z)$. Therefore, adding across $n \geq 2$ yields:

$$\mathcal{S}\sum_{n=2}^{+\infty}\sum_z \theta_n(z,\varphi)g_{n-1}(z) = \sum_{n=2}^{+\infty}\sum_z B_n^I(z,\varphi) = B^I - \sum_z B_1^I(z,\varphi) = 1 - B^A - \sum_z \theta_1(z,\varphi)S_0(\varphi)$$

Using the definitions above, we can then write:

$$\mathcal{S} = \frac{1 - \Big(b^A + h_0(\varphi)\sum_z \theta_1(z,\varphi)\Big)\mathcal{S}}{\sum_{n\geq 2}\sum_z \theta_n(z,\varphi)g_{n-1}(z)}$$

Solving for $\mathcal{S}$, we obtain the stationary measure of active sellers:

$$\mathcal{S} = \left(b^A + h_0(\varphi)\sum_z \theta_1(z,\varphi) + \sum_{n=1}^{+\infty}\sum_z \theta_{n+1}(z,\varphi)g_n(z)\right)^{-1}$$

Finally, the mass of potential entrants is $S_0(\varphi) = \mathcal{S}h_0(\varphi)$, the measure of incumbent sellers is $S_n = \mathcal{S}g_n$, the measure of active buyers is $B^A = \mathcal{S}b^A$, and that of inactive buyers is $B^I = 1 - B^A$.

G.2 Invariant Distribution (Special Case)

Assume an environment without exogenous (z,φ) shocks, and let $\sigma_n = S_n/(S_0 + \mathcal{S})$, for $n = 0, 1, 2, \ldots$ Then, when $\delta_f = 0$, we can re-write the flow equations in steady state as:

$$\eta(\theta_n)\sigma_{n-1} + (n+1)\delta_c\sigma_{n+1} - \big(\eta(\theta_{n+1}) + n\delta_c\big)\sigma_n = 0$$

for any $n \geq 1$, and $\delta_c\sigma_1 - \eta(\theta_1)\sigma_0 = 0$. Since $\sum_{n=0}^{+\infty}\sigma_n = 1$ by construction, $\{\sigma_n\}$ follows a stationary birth-death process, with Markov transition rates $\eta(\theta_{n+1})$ and $n\delta_c$ for transitions $n \to (n{+}1)$ and $n \to (n{-}1)$, respectively. Solving for $n \geq 1$ recursively, we find:

$$\sigma_n = \frac{1}{n!}\frac{\prod_{i=0}^{n-1}\eta(\theta_{i+1})}{(\delta_c)^n}\sigma_0 \tag{G.2.1}$$

Imposing that $\sum_{n=0}^{+\infty}\sigma_n = 1$ in equation (G.2.1) yields:

$$\sigma_0 = \left(1 + \sum_{n=1}^{+\infty}\frac{1}{n!}\frac{\prod_{i=0}^{n-1}\eta(\theta_{i+1})}{(\delta_c)^n}\right)^{-1} \tag{G.2.2}$$

From the last expression, it is clear that $\{\sigma_n\}$ admits an ergodic representation if, and only if:

$$\sum_{n=1}^{+\infty} \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{(\delta_c)^n} < +\infty \tag{G.2.3}$$

Under necessary condition (G.2.3), the stationary solution of the birth-death process $\{\sigma_n\}$ is given by (G.2.1)-(G.2.2). Using that $g_n = \sigma_n(1 + S_0/\mathcal{S})$ for $n \geq 1$, and $S_0/\mathcal{S} = \sigma_0/(1-\sigma_0)$, we then have:

$$g_n = \frac{S_0}{\mathcal{S}} \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{(\delta_c)^n}, \qquad \text{with } \frac{S_0}{\mathcal{S}} = \left[\sum_{n=1}^{+\infty} \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{(\delta_c)^n}\right]^{-1}$$

G.3 Existence of the Joint Surplus Function, given U^B

This section shows that there exists a unique joint surplus W, for each given U^B. First, we define the relevant functional space. Recall that the state space is $\mathcal{N} \times \mathcal{Z} \times \Phi = \{1, \ldots, \bar{n}\} \times \{z_i\}_{i=1}^{k_z} \times \{\varphi_j\}_{j=1}^{k_\varphi}$.

Definition 2 *Let $\mathcal{W}$ be a closed and bounded subspace of vector-valued functions $\mathbf{W} : \mathcal{N} \times \mathcal{Z} \times \Phi \to \mathbb{R}$, with the following properties:*

1. *Increasing in n, i.e. $W_{n+1}(z, \varphi) > W_n(z, \varphi)$, $\forall n \in \mathcal{N}$.*
2. *Constant at the upper bound of $\mathcal{N}$, i.e. $W_{\overline{n}}(z, \varphi) = W_{\overline{n}+1}(z, \varphi)$.*

For a given U^B, the joint surplus can be written as follows:

$$\begin{aligned}
(r + \delta_f) W_n(z, \varphi) = \max_{\mathbf{x}'(n', z', \varphi')} \Bigg\{ & n\Big(v(\varphi) + (\delta_f + \delta_c) U^B(\varphi)\Big) - \Big(\mathcal{C}(n; z, \varphi) + \psi\Big(x'(n+1, z, \varphi); \varphi\Big) x'(n+1, z, \varphi)\Big) \\
& + n\delta_c \Big(W_{n-1}(z, \varphi) - W_n(z, \varphi)\Big) + \psi\Big(x'(n+1; z, \varphi); \varphi\Big)\Big(W_{n+1}(z, \varphi) - W_n(z, \varphi)\Big) \\
& + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z)\Big(W_n(z', \varphi) - W_n(z, \varphi)\Big) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)\Big(W_n(z, \varphi') - W_n(z, \varphi)\Big) \Bigg\}
\end{aligned} \tag{G.3.1}$$

where we have used the short-hand notation $\psi(x; \varphi) \equiv \eta \circ \mu^{-1}\left(\frac{\Gamma^B(\varphi)}{x - U^B(\varphi)}\right)$. Equation (G.3.1) is a continuous-time recursive problem. In order to use dynamic programming methods, we first transform it into a form that can be exploited in Blackwell's Theorem. We do this by a so-called *uniformization method*.[49] The objective is to construct a set of transition probabilities that mimic those of the continuous-time specification.

For a given vector of current states $\boldsymbol{\gamma} \equiv (n, z, \varphi)$, define $\Gamma' \equiv \{0, n-1, n+1\} \times \mathcal{Z} \times \Phi$ as the set of possible future states. Let $\boldsymbol{\zeta} \equiv \{\mathbf{x}'(n', z', \varphi') : (n', z', \varphi') \in \Gamma'\} \subseteq \mathcal{X}$ denote a set of policies. Let $\mathcal{P}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\boldsymbol{\zeta})$ denote the probability of a $\boldsymbol{\gamma}$-to-$\boldsymbol{\gamma}'$ transition under policy $\boldsymbol{\zeta}$. Finally, let $q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta})$ be the vector of Markov transition rates for a fixed $\boldsymbol{\gamma}$. Then, we have:

[49] See Ross (1996), Section 5.8. For an application in economics, see Acemoglu and Akcigit (2012).

$$\mathcal{P}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\boldsymbol{\zeta}) \equiv \frac{1}{q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta})} \cdot \begin{cases} \psi\Big(x'(n+1,z,\varphi);\varphi\Big) & \text{for } \boldsymbol{\gamma}' = (n+1,z,\varphi) \\ n\delta_c & \text{for } \boldsymbol{\gamma}' = (n-1,z,\varphi) \\ \delta_f & \text{for } \boldsymbol{\gamma}' = (0,z,\varphi) \\ \lambda_z(z'|z) & \text{for } \boldsymbol{\gamma}' = (n,z',\varphi),\ \text{any } z' \neq z \\ \lambda_\varphi(\varphi'|\varphi) & \text{for } \boldsymbol{\gamma}' = (n,z,\varphi'),\ \text{any } \varphi' \neq \varphi \end{cases}$$

and

$$q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta}) \equiv \psi\Big(x'(n+1,z,\varphi);\varphi\Big) + n\delta_c + \delta_f + \sum_{z'\neq z} \lambda_z(z'|z) + \sum_{\varphi'\neq\varphi} \lambda_\varphi(\varphi'|\varphi)$$

Since the state space is bounded, there exists a $\overline{q}^S < +\infty$ for which $q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta}) < \overline{q}^S$, for all states $\boldsymbol{\gamma}$, given $\boldsymbol{\zeta}$. Therefore, we can think of transitions actually occurring at rate $\overline{q}$, with a fraction $q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta})/\overline{q}$ of them being actual transitions out of state $\boldsymbol{\gamma}$, and the remainder being "fictitious" transitions. Thus, the Markov chain can be represented by the following transition probabilities, including transitions across different states as well as from each state into itself:

$$\widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\zeta) \equiv \begin{cases} \dfrac{q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta})}{\overline{q}} \mathcal{P}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\boldsymbol{\zeta}) & \text{for } \boldsymbol{\gamma}' \neq \boldsymbol{\gamma} \\ 1 - \dfrac{q_{\boldsymbol{\gamma}}(\boldsymbol{\zeta})}{\overline{q}} & \text{otherwise} \end{cases}$$

Finally, define the corresponding discount factor as $\beta \equiv \frac{\overline{q}}{r+\overline{q}}$, and the per-period payoff function in state (n,z,φ) as:

$$\widetilde{\Pi}_n(z,\varphi;\boldsymbol{\zeta}) \equiv \frac{1}{\overline{q}} \Bigg[n\Big(v(\varphi) + (\delta_f+\delta_c)U^B(\varphi)\Big) - \Big(\mathcal{C}(n;z,\varphi) + \psi\big(x'(n+1,z,\varphi);\varphi\big)x'(n+1,z,\varphi)\Big)\Bigg]$$

We can now state the dynamic optimization problem (G.3.1) in discretized form:

$$W_n(z,\varphi) = \max_{\boldsymbol{\zeta}\subseteq\mathcal{X}} \left\{ \widetilde{\Pi}_n(z,\varphi;\boldsymbol{\zeta}) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\boldsymbol{\zeta}) W_{n'}(z',\varphi') \right\} \tag{G.3.2}$$

We are now ready to prove the main result:

Lemma G.1 *For any $(n,z,\varphi) \in \mathcal{N}\times\mathcal{Z}\times\Phi$, the joint surplus problem (G.3.1) admits a unique solution. That is, the mapping $T : \mathcal{W} \to \mathcal{W}$ defined by:*

$$T.W_n(z,\varphi) = \max_{\boldsymbol{\zeta}\subseteq\mathcal{X}} \left\{ \widetilde{\Pi}_n(z,\varphi;\boldsymbol{\zeta}) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\boldsymbol{\zeta}) W_{n'}(z',\varphi') \right\}$$

has a fixed point $T.W_n(z,\varphi) = W_n(z,\varphi)$.

Proof. T is a well-defined mapping from $\mathcal{W}$ to $\mathcal{W}$. We want to show that it defines a contraction. Since $\mathcal{W}$ is closed and $\boldsymbol{\zeta}$ takes values in a compact set, the contraction property will be enough to invoke Banach's Fixed Point theorem. Hence, we check that T satisfies monotonicity and discounting.

- **Monotonicity:** Take $\mathbf{W}^a, \mathbf{W}^b \in \mathcal{W}$ such that $W^a_n(z,\varphi) \leq W^b_n(z,\varphi)$, $\forall (n,z,\varphi) \in \mathcal{N}\times\mathcal{Z}\times\Phi$. Denote the corresponding optimal policies by:

$$\hat{\boldsymbol{\zeta}}^i \equiv \arg\max_{\boldsymbol{\zeta}} \left\{ \widetilde{\Pi}_n(z,\varphi;\boldsymbol{\zeta}) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\boldsymbol{\zeta}) W^i_{n'}(z',\varphi') \right\}$$

for each $i = a, b$. Then:

$$\begin{aligned} T.W^b_n(z,\varphi) &\geq \widetilde{\Pi}_n(z,\varphi;\hat{\boldsymbol{\zeta}}^a) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\hat{\boldsymbol{\zeta}}^a) W^b_{n'}(z',\varphi') \\ &\geq \widetilde{\Pi}_n(z,\varphi;\hat{\boldsymbol{\zeta}}^a) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\hat{\boldsymbol{\zeta}}^a) W^a_{n'}(z',\varphi') \\ &= T.W^a_n(z,\varphi) \end{aligned}$$

for any $(n,z,\varphi) \in \mathcal{N} \times \mathcal{Z} \times \Phi$, where the first inequality follows by optimality, and the second one follows from $\mathbf{W}^a \leq \mathbf{W}^b$.

- **Discounting:** Let $a \geq 0$ and $\mathbf{W} \in \mathcal{W}$, and denote the optimal policy by $\hat{\boldsymbol{\zeta}}$. Since a is a constant, we have that:

$$\begin{aligned} T.\big[W + a\big]_n(z,\varphi) &= \widetilde{\Pi}_n(z,\varphi;\hat{\boldsymbol{\zeta}}) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\hat{\boldsymbol{\zeta}}) \Big(W_{n'}(z',\varphi') + a \Big) \\ &= \widetilde{\Pi}_n(z,\varphi;\hat{\boldsymbol{\zeta}}) + \beta \sum_{\boldsymbol{\gamma}'\in\Gamma'} \widetilde{\mathcal{P}}_{\boldsymbol{\gamma},\boldsymbol{\gamma}'}(\hat{\boldsymbol{\zeta}}) W_{n'}(z',\varphi') + a\beta \\ &= T.W_n(z,\varphi) + a\beta \end{aligned}$$

  for any $(n,z,\varphi) \in \mathcal{N} \times \mathcal{Z} \times \Phi$. Since $\beta < 1$, discounting obtains.

Therefore, for a given U^B, T defines a contraction in $\mathcal{W}$ with modulus β, and by Banach's fixed-point theorem there exists a unique value function $W_n(z,\varphi)$ such that $T.W_n(z,\varphi) = W_n(z,\varphi)$. $\square$

G.4 Conditional and Aggregate Price Statistics

This section shows how to calculate, using the model's stationary solution, the conditional and aggregate price statistics that we use in the validation exercise.

Consider a price spell whose starting date is normalized to $\underline{t} = 0$ and which lasts until some unknown time $\bar{t} \geq 0$. Let T denote the total duration of the price spell, and let $\mathcal{F} : \mathbb{R}_+ \to [0,1]$ be the c.d.f. of T. We define the *survival function* associated to duration T as $\mathcal{S}^T_t \equiv \mathbf{Pr}[T \geq t] = 1 - \mathcal{F}_t$. The probability that the price spell will end in the $[t, t+\Delta]$ interval is:

$$\mathbf{Pr}[t < T \leq t + \Delta] = \mathcal{S}^T_t - \mathcal{S}^T_{t+\Delta}$$

The *hazard function* is defined as $\mathbf{Pr}[t < T \leq t + \Delta | T > t]$. Using Bayes' rule, we can write the hazard function in terms of the survival function as follows: $\mathbf{Pr}[t < T \leq t + \Delta | T > t] = 1 - \mathcal{S}^T_{t+\Delta} / \mathcal{S}^T_t$. The *instantaneous hazard rate* is then defined by the continuous-time limit, $h_t \equiv \lim_{\Delta \searrow 0} \frac{1}{\Delta} \left(1 - \mathcal{S}^T_{t+\Delta} / \mathcal{S}^T_t\right)$. Using L'Hôpital's rule:

$$h_t = -\boldsymbol{\partial}_t \log \mathcal{S}^T_t \tag{G.4.1}$$

Hence, defining the *cumulative hazard* as $\mathcal{H}_t \equiv \int_0^t h_s \mathrm{d}s$, the cumulative hazard and the survival functions are related by $\mathcal{S}_t^T = \exp\{-\mathcal{H}_t\}$ (as $\mathcal{S}_0^T = 1 - \mathcal{H}_0 = 1$). Using this result, we can write:

$$\mathbf{Pr}[t < T \leq t + \Delta | T > t] = 1 - \exp\left\{ - \int_t^{t+\Delta} h_s \mathrm{d}s \right\} \tag{G.4.2}$$

Finally, the expected duration of price spells is given by $\mathbb{E}\{T\} = \int_0^{+\infty} t \mathrm{d}\mathcal{F}_t$. Integrating by parts and using that $\mathcal{S}_t^T = 1 - \mathcal{F}_t$, we obtain:

$$\mathbb{E}\{T\} = \int_0^{+\infty} \mathcal{S}_t^T \mathrm{d}t \tag{G.4.3}$$

Instantaneous Hazard Rate Let $T_n(z,\varphi)$ denote the duration of price spells of firm $(n_t, z_t) = (n, z)$ in aggregate state $\varphi \in \Phi$. Conditional on survival, the probability of a price change during the interval $[t, t+\Delta]$, given that the price spell was still ongoing at date t, is:

$$\begin{aligned}\mathbf{Pr}\Big[t < T_n(z,\varphi) \leq t + \Delta \Big| T_n(z,\varphi) > t\Big] = &\Big[\eta\big(\theta_{n+1,t+\Delta}(z,\varphi)\big)\Delta + o(\Delta)\Big] + n\Big[\delta_c \Delta + o(\Delta)\Big] \\ &+ \sum_{\tilde{z} \neq z} \Big[\lambda_z(\tilde{z}|z)\Delta + o(\Delta)\Big] + \sum_{\tilde{\varphi} \neq \varphi} \Big[\lambda_\varphi(\tilde{\varphi}|\varphi)\Delta + o(\Delta)\Big]\end{aligned}$$

where $o(\Delta)$ collects higher-order terms. The *instantaneous hazard rate* (as defined in (G.4.1)) is:

$$h_n(z,\varphi) = \eta\big(\theta_{n+1}(z,\varphi)\big) + n\delta_c + \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) + \sum_{\tilde{\varphi} \neq \varphi} \lambda_\varphi(\tilde{\varphi}|\varphi)$$

Note the absence of time subscripts in the above expression. This is a convenient implication of our block-recursive structure. Two relevant implications of this result follow:

- The firm-level cumulative hazard is linear in time (though non-linear in the aggregate state):

$$\mathcal{H}_{n,t}(z,\varphi) = h_n(z,\varphi)t \tag{G.4.4}$$

  The survival function, in turn, takes the simple form $\mathcal{S}_{n,t}^T(z,\varphi) = \exp\big\{ - h_n(z,\varphi)t \big\}$.

- The cross-sectional average hazard of price changes is:

$$H_t(\varphi) \equiv \sum_{n \in \mathbb{N}} \sum_{z \in \mathcal{Z}} g_{n,t}(z) h_n(z,\varphi)$$

  where $g_n(z) = S_{n,t}(z) \Big/ \sum_{n,z} S_{n,t}(z)$ is the firm-size probability mass function (p.m.f.).

Frequency of Price Changes We define the frequency of price changes over a time window of length one (i.e. $1/\Delta$ sub-periods) as the cumulative probability of a price change after a spell of such length. Using equations (G.4.2) and (G.4.4)), this probability is:

$$f_n(z,\varphi) = 1 - \exp\Big\{ - h_n(z,\varphi) \Big\} \tag{G.4.5}$$

Therefore, the frequency of price changes at the firm-level is a jump variable. The average frequency of price adjustment in the cross-section of firms is:

$$F_t(\varphi) \equiv \sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} g_{n,t}(z) f_n(z,\varphi)$$

Hence, the aggregate frequency of price changes evolves over time according to the underlying distribution dynamics.

Expected Duration of Price Spells From equation (G.4.4), the price duration $T_n(z,\varphi)$ follows an exponential distribution with parameter $h_n(z,\varphi)$. The average duration (equation (G.4.3)) is then simply the reciprocal of the instantaneous hazard. Expressed in terms of frequency, this means:

$$\mathbb{E}\big\{T_n(z,\varphi)\big\} = -\frac{1}{\log\big(1 - f_n(z,\varphi)\big)} \tag{G.4.6}$$

Then:

$$D_t(\varphi) \equiv \sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} \frac{g_{n,t}(z)}{h_n(z,\varphi)}$$

is the average expected duration of prices at time t.

Moments of the Distribution of Price Changes Finally, we report moments of the distribution of (non-zero) price log-changes.

- The *expected absolute price change* in market (n,z) is defined as the average log change in prices. Denoting $\widehat{p} \equiv \log p$, and using $\big|.\big|$ for the absolute value, we have:

$$\begin{aligned}\mu_n^{\Delta}(z,\varphi) = {} & \eta\big(\theta_{n+1}(z,\varphi)\big)\Big|\widehat{p}_{n+1}(z,\varphi) - \widehat{p}_n(z,\varphi)\Big| + n\delta_c\Big|\widehat{p}_n(z,\varphi) - \widehat{p}_{n-1}(z,\varphi)\Big| \\ & + \sum_{\tilde{z}\neq z} \lambda_z(\tilde{z}|z)\Big|\widehat{p}_n(\tilde{z},\varphi) - \widehat{p}_n(z,\varphi)\Big| + \sum_{\tilde{\varphi}\neq\varphi} \lambda_\varphi(\tilde{\varphi}|\varphi)\Big|\widehat{p}_n(z,\tilde{\varphi}) - \widehat{p}_n(z,\varphi)\Big|\end{aligned}$$

- The *variance* of the distribution of price changes is given by:

$$\begin{aligned}\sigma_n^{\Delta}(z,\varphi) = {} & \eta\big(\theta_{n+1}(z,\varphi)\big)\Big(\Big|\widehat{p}_{n+1}(z,\varphi) - \widehat{p}_n(z,\varphi)\Big| - \mu_n^{\Delta}(z,\varphi)\Big)^2 + n\delta_c\Big(\Big|\widehat{p}_n(z,\varphi) - \widehat{p}_{n-1}(z,\varphi)\Big| - \mu_n^{\Delta}(z,\varphi)\Big)^2 \\ & + \sum_{\tilde{z}\neq z} \lambda_z(\tilde{z}|z)\Big(\Big|\widehat{p}_n(\tilde{z},\varphi) - \widehat{p}_n(z,\varphi)\Big| - \mu_n^{\Delta}(z,\varphi)\Big)^2 + \sum_{\tilde{\varphi}\neq\varphi} \lambda_\varphi(\tilde{\varphi}|\varphi)\Big(\Big|\widehat{p}_n(z,\tilde{\varphi}) - \widehat{p}_n(z,\varphi)\Big| - \mu_n^{\Delta}(z,\varphi)\Big)^2\end{aligned}$$

At the population level, these moments cannot be aggregated using g (the unconditional firm distribution), for not all firms change prices every period. Instead, we use the so-called *renewal distribution* of firms, that is, the distribution of firms conditional on a price adjustment:

$$r_{n,t}(z,\varphi) \equiv \frac{g_{n,t}(z) f_n(z,\varphi)}{\sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} g_{n,t}(z) f_n(z,\varphi)}$$

Then, the average expected size and the average standard deviation of price changes are, respectively:

$$\mathcal{M}_t^{\Delta}(\varphi) \equiv \sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} r_{n,t}(z,\varphi)\mu_n^{\Delta}(z,\varphi) \qquad \text{and} \qquad \Sigma_t^{\Delta}(\varphi) \equiv \sum_{n\in\mathbb{N}} \sum_{z\in\mathcal{Z}} r_{n,t}(z,\varphi)\sqrt{\sigma_n^{\Delta}(z,\varphi)}$$