Firm Dynamics and Pricing under Customer Capital Accumulation*

Sonia Gilbukh[†]
NYU Stern

Pau Roldan[‡]
New York University

June 8, 2017

PRELIMINARY DRAFT

[\[Click here for the latest version\]](#)

Abstract

What is the relationship between the rate at which firms accumulate their stock of demand over time and the prices that they set for their products? This paper analyzes the implications of customer capital accumulation for firms' pricing behavior and firm dynamics over the business cycle. We build an analytically tractable directed search model of the product market in which firms are ex-post heterogeneous in their customer base and commit to the prices they post. The model features dynamic long-term contracts with aggregate and idiosyncratic productivity shocks, endogenous entry of firms, and allows for an exact characterization of the firm distribution in and out of steady state. Price rigidity at the firm level emerges as an equilibrium outcome, and there is price dispersion in the cross-section because firms of different sizes use different pricing strategies to strike a balance between attracting new customers and exploiting incumbent ones. We show that our mechanism can generate realistic firm dynamics, a right-skewed firm size distribution, and size- and age-dependent markups which are in line with the data. Moreover, we analyze the response of the economy to aggregate supply and demand shocks, and find that our mechanism can give rise sluggish aggregate dynamics, countercyclical markups and substantial amplification.

JEL codes: D21, D83; E2; L11

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

*We are deeply indebted to Ricardo Lagos for his invaluable guidance and support throughout this project, and to Jess Benhabib, Boyan Jovanovic, and Edouard Schaal for continued discussions and suggestions. We have also benefited from insightful comments by Isaac Baley, Jaroslav Borovička, Luís Cabral, Laurent Cavenaile, Gian Luca Clementi, Diego Daruich, Jordi Galí, Mark Gertler, Simon Gilchrist, Nicholas Kozeniauskas, Julian Kozlowski, Joseba Martinez, Andrew McClellan, Virgiliu Midrigan, Simon Mongey, Jesse Perla, Leena Rudanko, Ennio Stacchetti, Venky Venkateswaran, and Basil Williams, as well as seminar audiences at NYU, and conference participants at SaM 2017 (Barcelona GSE Summer Forum), and CESC 2017 (Barcelona). All errors are our own responsibility. Comments and suggestions are welcome.

[†]Address: Department of Economics, New York University Stern School of Business, 44 W 4th street, 7th floor, New York, NY 10012. Email: sg2718@nyu.edu. Website: <https://sites.google.com/site/gilbukh>

[‡]*Corresponding author.* Address: Department of Economics, New York University, 19 W 4th street, 6th floor, New York, NY 10012. Email: pau.roldan@nyu.edu. Website: <https://wp.nyu.edu/pauroldan>

1 Introduction

It is a well-documented fact that firm heterogeneity is a relevant dimension for understanding firm behavior and aggregate flows in a variety of markets. Firms of different sizes, ages, and productivities experience persistently different growth paths along their life cycle. In particular, 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 \(2003\)](#)). Figure 1 illustrates this point for the United States at the establishment level by plotting the distribution of the employment share (left panel), establishment share (central panel), and the share of employees per establishment (right panel), across different age cohorts, averaged over the period 2003-2014. We see that, even though some large firms are young (i.e. are at most 5 years old), younger firms tend to be smaller than their more mature counterparts, and the firm size and age distributions are considerably right-skewed.

Figure 1: Establishment size differences across age cohorts (in years) in the U.S. private business sector, annual average rates (2003-2014).

Notes: This figure plots time-series averages (solid lines) and the min-max range (dashed lines) for $E_{it}/(\sum_j E_{jt})$ (left-most panel), $N_{it}/(\sum_j N_{jt})$ (central panel), and $(E_{it}/N_{it})/\sum_j (E_{it}/N_{it})$ (right-most panel), where E_{it} (respectively, N_{it}) denotes the number of employees (respectively, number of establishments) in cohort i in year t . An establishment is defined as the physical location where a firm's business is conducted. *Source:* Business Dynamics and Statistics of the U.S. Census Bureau (2003-2014), available at: <http://www.census.gov/ces/dataproducts/bds/data.html>.

A large theoretical literature, inspired by the seminal work of [Jovanovic \(1982\)](#) and [Hopenhayn \(1992\)](#), has traditionally attributed this evidence to a process of selection on the basis of productivity differences among firms, and has analyzed how these may in turn shape firm and industry dynamics in various meaningful ways. However, this interpretation has been recently challenged by [Foster *et al.* \(2008\)](#), who argue that, because empirical patterns of firm growth are usually based on revenue data (which cannot disentangle output prices from quantities), the productivity-based view on firm heterogeneity confounds selection on technological productivity with selection on profitability. As more disaggregated data have become available over subsequent years, the new empirical evidence has indicated that differences in firm performance are stemming, to a large extent, from differences in demand, rather than supply, fundamentals: as firms age, their revenue increases while TFP levels grow only slowly. Particularly, [Foster *et al.* \(2016\)](#) find, on the one hand, that within narrowly defined industries, new and young plants are typically

smaller than their more established competitors, and that they take time to catch up. On the other hand, while entrants have a slight productive advantage over incumbents, this difference dissipates within the first few years of operation. Finally, young firms set lower prices on average relative to their more mature competitors.

These observations suggest an important demand-sided channel of variation: frictions in demand accumulation could account for the behavior of small and young entrants. Using scanner data from the U.S. retail sector, [Hottman *et al.* \(2016\)](#) similarly show that most variation in firm sales and productivity are attributable to variation in demand components (e.g. firms’ “appeal” such as quality and taste, and product scope).¹ These empirical studies hence cast some doubt on the established notion that selection on productivity is indeed a prominent driver of firm and industry dynamics.

In this paper, we develop a firm dynamics model with aggregate and idiosyncratic shocks in which there is a meaningful role for a demand accumulation process. We interpret this process as the build-up of a customer base. In [Section 3](#), we present a search-and-matching model of a frictional product market in which buyers must search for sellers, and where the latter post price sequences which affect the rate at which demand accumulates. Sellers are heterogeneous in the number of consumers that they sell the product to, and they cannot price-discriminate between new and continuing customers.² Even though the model’s dynamics are rich, the environment admits a representation based on recursive dynamic contracts, in which sellers post complete, long-term contracts specifying the price and the continuation utilities for the buyers. Because sellers commit to the posted contract, valuable customer relationships emerge endogenously, as forward-looking buyers must internalize the future path of prices and the future size evolution that is implied by those pricing decisions. In addition, buyers can choose to endogenously separate from their supplier if they so desire. In equilibrium, when sellers first enter the product market, they have few customers and their demand is low, so they optimally decide to post low prices and extract a small share of the surplus in order to generate a higher rate of attracting new customers. Because of search frictions, the customer base of the firm is a slow-moving state variable, and the customer capital accumulation process takes time. As firms mature, they lower their future promises as they increasingly prefer to extract surplus from their current base. As a result, their markups tend to increase and the firm’s rate of growth slows down. On the other side of the market, customers trade off the current price they are being charged to the accessibility of the good, so they may not optimally choose to visit the lowest-price supplier because lower-price firms are also the ones with the longest queues. In equilibrium, the model admits a stationary distribution of buyers across firms’ idiosyncratic states (i.e. size and idiosyncratic productivity levels).

The firms’ pricing behavior thus breaks the [Diamond \(1971\)](#) paradox, as sellers do not typically extract all the surplus from buyers because customer markets are forward-looking. Instead, sellers prefer to keep prices strictly below the buyer’s valuation. Because sellers commit to the promised utility of the contracts

¹ Similar observations have been made for other countries, e.g. [Kugler and Verhoogen \(2012\)](#) (Colombia). [Cavenaile and Roldan \(2016\)](#) show that small firms have higher advertising-to-sales ratios, while [Eaton *et al.* \(2014\)](#) show that similar demand considerations are also prevalent in the pricing decisions of exporting firms.

² We believe this type of anonymity assumption to be present, for example, in the retail and wholesale sectors.

signed with their outstanding buyers, prices are indexed to the size of the customer base in a way so that an optimal balance is struck between high revenues (via contemporaneous prices) today and high market shares (via continuation values) in the future. In this sense, the model offers a micro-foundation for two prevailing features of the data: price stickiness at the firm level (Nakamura and Steinsson (2013)), and price dispersion in the cross section for narrowly defined product markets (Kaplan and Menzio (2015)). Moreover, because young and small firms set low prices upon entry, which in turn generates slow-moving dynamics, the model qualitatively replicates the aforementioned observations of plant-level dynamics by Foster *et al.* (2016).

In Section 3.3, we show that the contract that maximizes the seller’s expected value can be found as the contract that solves the joint surplus maximization problem. In the latter formulation, the price level can then be thought of as establishing a surplus-splitting rule between the seller and all its captive buyers. This allows for a decomposition of the optimal price into various cyclical components that reflect the underlying forces behind the firm’s customer attraction incentives. Using this decomposition, we find that interesting size non-monotonicities may arise in equilibrium. The equivalence between the seller’s and the joint surplus problems is also important because it reduces the dimensionality of the state space considerably, and renders an analytical characterization of the equilibrium dynamics. The result arises even though the model features rich dynamics, aggregate and idiosyncratic shocks, and two-sided heterogeneity (on sellers and buyers). Furthermore, we show that the allocation is constrained-efficient (that is, a planner that is subject to the search frictions would allocate buyers and sellers across markets in the same way as firms do through price contracts in the decentralized economy). This allows us to interpret the model as a *theory of efficient price markups*, in which sellers’ use of prices leads to a socially optimal allocation of customers across different product markets.

Another attractive feature of our setting is that we can fully describe the equilibrium allocation independently of the distribution of agents across states and time. We accomplish this by exploiting the property known as *block recursivity*, common in directed search models (e.g. Menzio and Shi (2010, 2011)), which implies that agents need not keep track of the distribution of firms. Since the firm size distribution can be derived independently of the optimal contract problem, the dynamics of firm size and pricing along both along the stationary solution and along transition paths are exactly characterized. We show how this construction works in Section 3.4.

In Section 4 we present some of these predictions at the stationary solution and from a simulation exercise, for a numerical example. We show that our mechanism can generate a right-skewed firm size distribution (as in the data), as well as realistic firm dynamics of entry and exit. We also present a number of predictions regarding cross-sectional price dispersion, markups, and the distribution of price changes. In particular, we show that there exist asymmetries in the rate and frequency of price adjustments across firms, a feature of the micro-data (e.g. Nakamura and Steinsson (2010)). An important result of the model is that markups vary with firm age and size. Assuming marginal costs are fairly constant across firms (or that they scale only weakly with firm size), markups are increasing with the size of firms. Since in the model (as in the data) small firms are more likely to be young, markups are on average increasing in firm age, as well. Not only has this feature been documented using micro-data (e.g. Foster *et al.*

(2008), Peters (2016)), but it is also likely to have implications at the aggregate level, especially in terms of responses to business cycle dynamics and monetary non-neutrality.

To tackle this question, we also analyze the response of the economy to both aggregate demand and aggregate supply shocks for a non-calibrated economy. In this exercise, we find both level and distributional effects. Shocks that raise the marginal propensity to consume by buyers generate a boom in demand and additional entry of firms. Since new firms enter small and charge relatively low prices, increased competition lowers the average price level of the economy, and as a result the average size and growth rate of firms increases. Moreover, since firms size dynamics are slow-moving, our model with customer markets can offer a rationale for the sluggish response of macroeconomic variables to aggregate shocks, and a channel of amplification that could potentially be important as a mechanism of transmission of nominal shocks. Indeed, while countercyclical markups are a notoriously important channel of the macroeconomic transmission of demand shocks in models of sticky prices (Galí *et al.* (2007), Justiniano *et al.* (2010)), and may account for a substantial part of the cyclical variation in output (Rotemberg and Woodford (1999)), in our case they emerge endogenously in the absence of explicit menu costs or exogenously-given infrequent price-setting. Thus, the model may provide a rationale for the co-movement, the asymmetric response, and the incomplete pass-through to shocks that we see in the data.

Outline The remainder of this paper proceeds as follows. Section 2 summarizes related literature, and specifies our main contributions. In Section 3 we present our model of customer acquisition, pricing, and firm dynamics with aggregate and idiosyncratic productivity shocks, including the analytical derivation of the firm size distribution in Section 3.4. Section 4 then presents the computational implementation, a numerical example, and simulation results. Section 5 concludes and outlines the next steps of this project. An Appendix at the end presents derivations and additional material.

2 Related Literature

There is a large amount of survey evidence that suggests that the customer base of a firm and its pricing decisions are tightly linked. Blinder *et al.* (1998) show that the vast majority of firms report having implicit contracts with their customers, and that these contracts are a major source of price stickiness. For a variety of different countries, other studies such as Hall *et al.* (1997), Cason and Friedman (2003), Renner and Tyran (2004), and Apel *et al.* (2005), present similar survey evidence showing that customer loyalty is a sensitive concern for price-setting firms.³ More recently, using scanner data from a number of large retailers in the U.S., Paciello *et al.* (2016) are able to identify customer-retailer transactions and demonstrate that customer attrition rates are on average low over long spells (i.e. a retailer’s customer base is typically sticky). Moreover, they show that not only the probability that a customer leaves a retailer is significantly positively affected by changes in consumer prices, but also that longer relationships are associated with lower probabilities of the customer exit (i.e. the hazard rate is decreasing).

³ See also Fabiani *et al.* (2004) for exhaustive cross-country evidence in the Euro Area.

Our study is primarily related to a long tradition of building customer capital into macroeconomic models of firm pricing. Early attempts by Phelps and Winter (1970), Bils (1989), and Rotemberg and Woodford (1991, 1999), analyzed pricing behavior under customer retention concerns. In these papers, firms face an exogenously given law of motion for the customer base. A number of papers have further developed a variety of reasons why customers may be locked into a repeated-purchase relationship in the first place. For instance, Klemperer (1987, 1995) and Kleshchelski and Vincent (2009) propose that customers face switching costs, which can be broadly understood as the transaction costs associated to switching to a competitor, or the costs in terms of utility when the consumer has developed a loyalty toward a certain brand. In a similar vein, Ravn *et al.* (2006) and Nakamura and Steinsson (2011) consider that customers form good-specific habits for consumption, and for this reason have a preference for repeating purchases with the same sellers.⁴ While the literature has traditionally resorted to reduced-form formulations for customer capital formation, we contribute by offering an endogenous mechanism whereby customers become captive.

Whatever the reason, the common insight in the literature is that, when customers are locked into their suppliers, demand becomes forward-looking. In this situation, prices not only fulfill the usual distributive role of splitting gains from trade between buyers and sellers, as in a standard Walrasian economy, but may also play an allocative role and determine the duration of customer-seller relationships or the likelihood that new ones form. Consequently, the optimal price of the static profit maximization problem may differ from the dynamic one because firms must strike a dynamic trade-off between exploiting their current customers (by setting high prices) and attracting new customers in the future (by setting low prices). In short, low prices today serve as a tool to guarantee larger market shares in the future.⁵ Our model shares these features with the literature, and may explain for example the disparity between the short- and long-run demand elasticities that we see in the data.

An important focus of our work are the implications that customer-seller relationships have on firm dynamics, including firm growth, entry, and exit. In this dimension, our paper is also related to the literature that has introduced a role for various types of firm intangibles into models of firm and industry dynamics.⁶ The effects of intangibles on different aspects of the aggregate economy are well-understood, including labor wedges (Gourio and Rudanko (2014a)), aggregate productivity (McGrattan and Prescott (2014) and McGrattan (2015)), and household behavior (Hall (2008)). A number of papers have further analyzed how expenditures on intangibles can shape the evolution of firms and industries. Atkeson

⁴ As an application of this approach, Gilchrist *et al.* (2016) show that the inter-temporal pricing behavior of firms in customer markets interacts with their degree of financial constraints, and can rationalize the mild disinflation episode experienced in the United States during the Great Recession.

⁵ If firms are not committed to the price path, however, a well-known time-inconsistency problem arises: firms promise low prices to attract customers and, once these customers become captive, sellers renege on their earlier promises and charge a higher price. For instance, Nakamura and Steinsson (2011) show that, in this case, repeated interaction can lead to the development of implicit contracts which, through a set of properly defined trigger strategies, can prevent prices from increasing beyond a certain upper bound. Rudanko (2017) analyzes possible inefficiencies emerging from time-consistency issues. We show that, when there is commitment on the seller's side (e.g. the firm faces reputational concerns), a similar type of contractual environment can emerge.

⁶ Firm intangibles are a substantial share of firms' expenditure, and in the U.S. as much as 7.7% of GDP is devoted to marketing, with advertising expenditures alone averaging about 2.2% since the early 1980s (Arkolakis (2010)).

and Kehoe (2005) show that organizational capital (i.e. investment in new technologies, new markets, and new and higher-quality products) can drive the life-cycle of plants, and Hsieh and Klenow (2014) argue that these processes might account for differences in plant-specific total factor productivity (TFP) between different countries. Drozd and Nosal (2012), Eaton *et al.* (2014), and Arkolakis (2016) study how costs to market penetration can rationalize certain patterns of trade and firm growth among exporting firms. A related paper to ours is Gourio and Rudanko (2014b), who analyze the timing of firm responses to investment shocks by augmenting a neoclassical firm investment model with a search model of the product market in which firm use price discounts to attract new customers. Like them, we interpret customer acquisition as a search-and-matching process, but while in their paper firm expansion is limited by convex adjustment costs to customer accumulation, we achieve this endogenously via our dynamic contracts structure. Moreover, we emphasize how this shapes the firm’s life cycle, and analyze business-cycle dynamics. Luttmer (2006) and Fishman and Rob (2003) also study the implications of customer acquisition for the firm size distribution, but those papers do not allow for a meaningful role for prices. In contrast, Paciello *et al.* (2016) study the implications of customer markets for the cross-sectional price distribution and the pass-through of shocks, but place no emphasis on the implications for firm dynamics.⁷

We contribute to this literature by providing the missing link between market shares and firm dynamics in customer markets. In particular, a prevailing feature in the data is that the growth rate of firm size is size-, and age-, dependent (e.g. Sutton (1997), Caves (1998), and Rossi-Hansberg and Wright (2007)).⁸ Further, the size distribution is right-skewed in the data (e.g. Luttmer (2007)). Our model can qualitatively generate these patterns. Firstly, there exist differences in growth rates across sizes and ages: even though we assume a constant-returns-to-scale production technology, small firms promise higher continuation utilities, thereby attracting more customers and generating a higher likelihood of growing, while larger firms prefer exploiting their current customer base, typically by charging a higher price, thereby growing slower or even shrinking on average. Secondly, the distribution is right-skewed: since larger firms are visited less frequently because they set higher prices, the probability that a firm grows to be large is low, which generates the fat right-tail.

The link between firm dynamics and prices is also supported by a number of studies relating empirically demand-side fundamentals to the determination of prices at the firm level. Besides the aforementioned

⁷ To cite more examples in the literature of intangible and industry dynamics: Kaas and Kimasa (2016) embed the Gourio and Rudanko (2014b) framework into a frictional labor market to study the joint dynamics of prices and wages; Perla (2016) studies the implications of product sorting by uninformed consumers on the industry life cycle and the degree of market competition; Bai *et al.* (2012) incorporate a frictional goods market into a representative-agent neoclassical economy to study the role of demand shocks; and Petrosky-Nadeau and Wasmer (2015) combine the goods market friction with frictions in the credit market to analyze distortions in the labor market. Finally, Cavenaile and Roldan (2016) argue that interactions between different types of intangible expenditures can shape the firm size distribution as well as affect long-run economic growth.

⁸ The dependence of firm growth to firm characteristics has been an ongoing and is still very active debate, at least since Gibrat (1931) first argued that firm growth is independent of firm size. The empirical evidence is mixed, but it tends to favor negative deviations of Gibrat’s Law (e.g. Hall (1987), Geroski (1998), Audretsch *et al.* (2004)). Recently, Haltiwanger *et al.* (2013) have argued with U.S. Census data that, when measuring size by employment and growth by job creation, the size dependence becomes insignificant at the establishment level after controlling for age. In our model, firms start small and grow fast, so age and size are tightly linked.

work by Foster *et al.* (2008, 2016) and Peters (2016), Kugler and Verhoogen (2012) find a positive correlation between output prices and size at the plant level for Colombian firms, while Carlsson *et al.* (2014) find, using Swedish micro data, that a substantial component of output price variation remains unexplained after accounting for productivity differences. DeLoecker (2011) performs a similar analysis in the context of international trade. We rely on these observations to justify our demand-driven theory of firm dynamics. At the product level, Argente and Yeh (2016) have shown, using U.S. scanner data, that younger products tend to have lower prices and experience relatively more frequent and more drastic price changes. Since in our model each firm produces a distinct product, the theory can also rationalize these patterns.

Because we use search frictions to obtain a non-degenerate cross-sectional price distribution, this paper also contributes to the search literature on equilibrium price dispersion. Empirically, newly available micro-level evidence has shown that there is substantial price dispersion for identical goods (e.g. Kaplan and Menzio (2015)). Theoretically, search models of price dispersion have proliferated since the early work by Butters (1977), Varian (1980), and Burdett and Judd (1983). Recently, Menzio and Trachter (2015) and Kaplan *et al.* (2016) have shown that price dispersion can emerge from buyer heterogeneity in situations in which sellers can price-discriminate. In our model, in contrast, buyers are identical and there is no price-discrimination. Instead, it is differences among 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 for the evolution of the firm size distribution, nor do they analyze the implications of prices for firm dynamics over the business cycle. More recently, Rudanko (2017) has studied the role of both discriminatory and non-discriminatory pricing for firm growth, but while her paper focuses on the implications for efficiency of different commitment protocols in monopolist markets, we instead describe an efficient economy and analyze firm dynamics with an emphasis on the cross-sectional response to aggregate shocks.

Finally, our paper is related to search-and-matching models with large firms, where most advances have been made in the context of labor markets (e.g. Elsbj and Michaels (2013), Kaas and Kircher (2015), and Schaal (2017)). From a methodological point of view, we build on models of *directed* (or *competitive*) search (e.g. Moen (1997)). Particularly, we combine two technical insights from this literature, which turn out to be essential for analytically characterizing the equilibrium. First, we exploit the property of block recursivity (e.g. Menzio and Shi (2010, 2011)), which allows for a tractable characterization of the firm size distribution. This property implies that agents do not have to carry distributions as state variables in their optimization problems even though the model incorporates business cycles, thereby allowing us to study out-of-steady-state transitional dynamics in response to aggregate shocks. Secondly, we make use of recursive dynamic contracts (e.g. Moscarini and Postel-Vinay (2013), Schaal (2017)), which greatly reduces the dimensionality of the state space. In addition, we model firm size dynamics using a similar environment to that of Klette and Kortum (2004), in that the customer base of the firm evolves over a discrete grid in continuous time through endogenously-determined Poisson rates. We contribute to this strand of literature theoretically by showing that the combination of these ingredients leads to analytical solutions, even in the presence of aggregate shocks. These, in turn, allow us to solve the model exactly

without any need for numerical approximation methods.

3 Model

This section describes a directed search model of customer and firm dynamics with aggregate and idiosyncratic productivity shocks in which sellers must post pricing contracts in order to attract consumers. The key assumptions in the model are that firms offer long-term contracts to their customers (because of perfect commitment to the posted terms of trade), and that firms are unable to discriminate between incoming and incumbent customers (i.e. customers are *anonymous* from the point of view of suppliers). While plausible for markets with large numbers of potential buyers (for instance retail markets), these two assumptions are not innocuous: when firms commit to their prices, a dynamic trade-off emerges for the seller between the added value of new customers and the loss of profits on incumbent ones. In turn, demand becomes forward-looking.

3.1 Environment

Time is continuous and goes on forever, with a time instant indexed by $t \in \mathbb{R}_+$. The economy is populated by a mass-one continuum of risk-neutral, infinitely-lived, ex-ante identical *buyers*, and a continuum of risk-neutral *firms* (sometimes referred to as *sellers*). The total measure of firms is determined in equilibrium. Buyers and sellers both discount future payoffs with a common and exogenous rate, $r > 0$. All payoffs and payoff-relevant states are public information among all agents.⁹ The aggregate state of nature is denoted by the random variable φ , taking values in a discrete and finite grid $\Phi := \{\underline{\varphi} < \dots < \overline{\varphi}\}$, with cardinality $|\Phi| = k_\varphi \geq 2$. The aggregate state follows a continuous-time Markov chain with Poisson transition rates $\lambda_\varphi(\varphi'|\varphi) \geq 0$ for a φ -to- φ' transition. The transition rates satisfy the condition $0 < \sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) < +\infty$, $\forall \varphi \in \Phi$, establishing that no state φ is absorbing (first inequality), and that, when visited, the economy always spends a non-zero measure of time on any given state (second inequality).

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: congestion effects prevent all searchers from finding a match with certainty at any given instant.¹⁰

Buyers All buyers value the consumption of the good by the same fixed utility flow, $v > 0$. At any instant in time, a buyer is either *active*, if matched with a firm and consuming the good, or *inactive*, if

⁹ For a search model in which private information about buyers' payoffs can generate customer relationships, see Shi (2016). Menzio (2007) analyzes the dynamics of prices when the cost structure of firms is private information.

¹⁰ An interpretation of the product market frictions is that there exist informational asymmetries regarding product characteristics, as argued by Gourio and Rudanko (2014b). Perla (2016) provides a micro-foundation based on the evolution of buyers' consideration sets.

unmatched and searching for a seller at a cost.¹¹ While the total mass of buyers is exogenous and normalized to unity, the buyers' composition between active and inactive across aggregate states is determined in equilibrium.

If the buyer is matched with a firm who is charging price p_j at time $j \geq t$, then as of time $t \geq 0$ the buyer obtains the expected value:

$$V^B(\varphi_t) = \mathbb{E}_t \left\{ \int_t^{\tau_B} e^{-r(j-t)} (v - p_j) dj + e^{-r(\tau_B-t)} U^B(\varphi_{\tau_B}) \right\}$$

where $\mathbb{E}_t\{\cdot\}$ is an expectation operator, the random variable $\tau_B \geq t$ represents the time at which a buyer who is matched at t will cease to purchase from the firm (due to separation, which we describe below), and U^B denotes the expected continuation value of an inactive buyer. In words, an active buyer enjoys the flow surplus from consumption until separation occurs at some uncertain time $\tau_B \geq t$.

Symmetrically, if inactive at time t , a buyer obtains expected value:

$$U^B(\varphi_t) = \mathbb{E}_t \left\{ - \int_t^{\tilde{\tau}_B} e^{-r(j-t)} c(\varphi_j) dj + e^{-r(\tilde{\tau}_B-t)} V^B(\varphi_{\tilde{\tau}_B}) \right\}$$

where $\tilde{\tau}_B \geq t$ is the time at which a successful match is formed, and $c(\varphi) > 0$ is the flow cost of search for inactive buyers, which we interpret as a proxy for the disutility of shopping. Without loss of generality, the elements in Φ are ordered so that $c(\varphi_i) \leq c(\varphi_{i-1})$, for all $i = 2, \dots, k_\varphi$. Since the search cost relates directly to buyers' preferences, allowing for c to vary with the business cycle incorporates the possibility of aggregate demand shocks into the model.¹²

We assume the inactive buyer is always searching (and thus, paying c every instant until a match is formed), because the value of its outside option is no greater than U^B . When an inactive buyer finds a firm, the buyer becomes active, and she perceives value V^B . Note that in both of these values the buyer must forecast future market (and particularly seller-specific) conditions, as the times of separation (τ_B) and matching ($\tilde{\tau}_B$) are uncertain as of time t . Finally, we assume that no buyer is ever allowed to borrow against its future income.

Sellers There exist two types of sellers: incumbent (or *active*) sellers, and potential entrant (or *inactive*) sellers.

At any given time t , a typical incumbent seller has a customer base of $n_t \in \mathbb{N} := \{1, 2, 3, \dots\}$ customers, which we subsequently call the *size* of the firm. Each firm is also characterized by the realization of an idiosyncratic productivity level z , taking values on a discrete and finite grid $\mathcal{Z} := \{z, \dots, \bar{z}\}$ of cardinality $|\mathcal{Z}| = k_z \geq 2$. The idiosyncratic state is Markov with transition rates $\lambda_z(z'|z) \geq 0$ for a z -to- z' transition, where $0 < \sum_{z'} \lambda_z(z'|z) < +\infty$, $\forall z \in \mathcal{Z}$. The realization of the idiosyncratic state is observable and public information.

¹¹ Search costs are a stand-in for product switching costs, and they have been found to be relevant for certain retail markets (e.g. [Sorensen \(2000\)](#)).

¹² The cyclical variation in shopping disutility can be thought of as reflecting the cyclical nature of household shopping behavior, which has been documented by [Petrosky-Nadeau et al. \(2016\)](#) for the United States.

An incumbent firm'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 firm equals the number of customers in the firm, with each customer consuming one unit. The firm also faces operating variable flow costs of $\mathcal{C}(n; z, \varphi)$, which depend on the idiosyncratic state (n, z) , as well as possibly the aggregate state φ . Further, we make the following assumptions:

Assumption 1 (i) $\mathcal{C}$ is a continuous, increasing, and time-invariant function of n ;
(ii) For any $(z, \varphi) \in \mathcal{Z} \times \Phi$, $\mathcal{C}(n; z, \varphi) \geq 0$, $\forall n \in \mathbb{N}$, and $\mathcal{C}(0; z, \varphi) = 0$.

Assumption 1 imposes mild regularity condition on the firms' technology. In particular, 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 in the firm's productive technology. For now, we need not make an explicit assumption in this respect. As we shall see, in equilibrium smaller firms experience faster growth even under constant returns in technology.

Besides producing, incumbent sellers post prices in the product market. Posting a price bears no explicit cost for an incumbent. Moreover, incumbent sellers exit the market (and thereby enter the pool of potential entrants) in either one of two ways: because they go bankrupt, at a constant Poisson rate $\delta_f > 0$, or if they separate from their last remaining customer (because the buyer abandons the firm), at a rate $\delta_c > 0$. These events are assumed to be mutually independent, and orthogonal to the idiosyncratic and aggregate shocks.

Since the firm is not allowed to price-discriminate, its total revenues are pn . The sequence problem of an incumbent firm with n_t customers at time t is, therefore:

$$V^S(n_t; z_t, \varphi_t) = \max_{(p_j; j \geq t)} \mathbb{E}_t \left\{ \int_t^{\tau_S} e^{-r(j-t)} (p_j n_j - \mathcal{C}(n_j; z_j, \varphi_j)) dj + e^{-r(\tau_S-t)} V_0^S(\varphi_{\tau_S}) \right\}$$

subject to the law of motion for n_t (which we specify later on), where $\tau_S \geq t$ is the random time at which the firm exits the market, and V_0^S denotes the value of having no customers. In words, the incumbent seller enjoys flow profits until being separated from all of its customers, at which point the firm exits and re-enters the pool of inactive firms.

Like incumbent firms, inactive firms are posting prices in order to attract customers and start operating in the product market. Unlike them, however, they must incur a flow cost $\kappa > 0$ for doing so, which can be thought of a set-up cost (e.g. the fixed costs of advertising a new business idea, which may or may not result in a new successful entrant). Therefore, the value of inactive firms is:

$$V_0^S(\varphi_t) = \mathbb{E}_t \left\{ - \left(1 - e^{-r(\tilde{\tau}_S-t)} \right) \frac{\kappa}{r} + e^{-r(\tilde{\tau}_S-t)} \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) V^S(1; z_0, \varphi_{\tilde{\tau}_S}) \right\}$$

where $\tilde{\tau}_S \geq t$ is the time at which the firm successfully attracts its first customer, and $z_0 \sim \pi_z$ is the initial productivity draw for an entrant, where $\pi_z(z) \geq 0$, and $\sum_{z \in \mathcal{Z}} \pi_z = 1$. We assume there is free entry of firms into the product market.

Contracts 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 length $j \geq 0$ of the match, conditional on no separation. Contracts are complete and fully 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 $p_{t+j} = p(n^{t+j}; z^{t+j}, \varphi^{t+j})$, $\forall j, t$.¹³ We assume all prices p_{t+j} must be non-negative, and are bounded above by v .

All agents are able to direct their search. Sellers announce price contracts in order to attract buyers. Buyers, on the other hand, can perfectly observe the posted contract and are able to discern the identity (i.e. the size n) of the firm who is posting it. There are search frictions in the product market because buyers are unable to coordinate their actions, or alternatively because some aspects of the supply side of the economy are unknown to the potential customer (e.g. the exact location of stores). For this reason, inactive buyers must pay a search cost c in order to sample contracts.

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).¹⁴ On the sellers' side, we make two main assumptions. First, unlike the buyer, 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 be made incentive-compatible.¹⁵ Second, we assume anonymity among buyers, in that the firm is unable to index the contract to the identity of each buyer. This implies that, when setting a price path optimally, the firm must internalize that any additional revenue from expanding the number of customers comes at the expense of potentially lowering the average revenue for the incumbent base. This component of the firm's decision renders the customer base a valuable form of capital stock for the seller.

Product Markets As it is customary in the directed search literature, a sufficient statistic for each long-term pricing contract is the promised life-time value 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 denote this value by x , and assume that all sellers advertising the same value x compete in all such contracts. Moreover, buyers cannot coordinate their decisions among themselves. Up to the observable idiosyncratic state (n, z) , sellers offering the same value x are virtually indistinguishable to the buyer. Thus, x effectively indexes a *market*. Consequently, the economy can be thought of being composed of a nonempty set $\mathcal{X} = [\underline{x}, \bar{x}] \subseteq \mathbb{R}_+$ of markets.

Each seller can simultaneously post, and each buyer can simultaneously search, in at most one submarket. Within each submarket $x \in \mathcal{X}$, a certain mass $B(x; \varphi) \in [0, 1]$ of buyers apply to the contract,

¹³ The distribution of agents over states is itself an aggregate state of the economy, and thus should be incorporated as a relevant dimension in the price function. However, as we discuss later, in our framework the only *payoff-relevant* aggregate state is φ .

¹⁴ The reason is that there are endogenous switching costs for buyers. In particular, customer loyalty emerges endogenously because of the opportunity cost (i.e. forgone expected value) of leaving the seller of size n .

¹⁵ A possible interpretation of this assumption is that firms have a reputational concern, so that reneging on previous promises entails unaffordable costs. Nakamura and Steinsson (2011) show that, in the absence of the assumption commitment, implicit contracts with sticky prices may emerge in equilibrium in a large variety of ways.

which is posted by a mass $S(x; \varphi) \geq 0$ of sellers. Because buyers cannot screen sellers within a market x , their *within-market* search is random. A market is said to be *active* (or *open*) if:

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

where $\theta(x; \varphi)$ is the buyer-to-seller ratio in market segment x , also referred to as the *market's tightness*. Importantly, agents take the mapping $\theta : \mathcal{X} \times \Phi \rightarrow \mathbb{R}_+$ as given when directing their search toward specific offers. Meeting rates within a market can be fully summarized by its tightness: in a typical $x \in \mathcal{X}$, a single applicant obtains offer x at the Poisson arrival rate $\mu(\theta(x; \varphi)) \geq 0$, while a seller successfully finds an applicant for offer x at the Poisson arrival rate $\eta(\theta(x; \varphi)) \geq 0$, where $\eta(\theta) = \theta\mu(\theta)$, $\forall \theta \geq 0$.¹⁶ We impose the following additional regularity conditions:

Assumption 2 *The meeting rates satisfy:*

- (i) $\eta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and $\mu : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ are twice continuously differentiable and time-invariant functions;
- (ii) η is increasing and concave; μ is decreasing and convex;
- (iii) $\eta \circ \mu^{-1}$ is strictly concave;
- (iv) $\eta(0) = \lim_{\theta \nearrow +\infty} \mu(\theta) = 0$, and $\lim_{\theta \nearrow +\infty} \eta(\theta) = \lim_{\theta \searrow 0} \mu(\theta) = +\infty$.

As it will become clear shortly, assumptions 2(i)-(ii) guarantee that the problems of the buyer and the seller are well-defined; the assumption on the composition $\eta \circ \mu^{-1}$ guarantees that the price-posting problem of the seller has a unique interior solution; and assumption 2(iv) imposes a transversality condition.

A common micro-foundation of these assumptions is to suppose that each market $x \in \mathcal{X}$ is endowed with a constant-returns-to-scale matching function $\mathcal{M}(B, S)$ that is equipped with the appropriate Inada conditions. Pairwise matching then requires that $\mu(\theta) = \mathcal{M}(1, \theta^{-1})$ and $\eta(\theta) = \theta\mu(\theta)$. Intuitively, the seller's meeting rate is found as the measure of meetings per seller, and because of congestion effects in the product market, longer queues of applicants for a contract yield lower (respectively, higher) rates of matching for the buyer (respectively, the seller).

3.2 Block-Recursive Formulation

We seek to solve for the *symmetric Markov perfect equilibrium* of this economy. We narrow attention to this class of equilibrium in the following sense. Markov-perfection implies that the equilibrium policies depend solely on the firm's vector of payoff-relevant states. For a firm with $n \in \mathbb{N}$ customers promising a life-time value of $x \in \mathcal{X}$ to each of them, this vector is $(n, x; \mathbf{s})$, where henceforth we use $\mathbf{s}$ to denote the vector of payoff-relevant states that are exogenous from the point of view of an individual agent (and

¹⁶ The assumption that meeting rates are Poisson is made for simplicity, as it guarantees that the event that two or more meetings occur within an infinitesimal instant is, while possible, measure-zero. Ultimately, this implies that sellers accumulate customers over time on a one-by-one basis, which simplifies the analysis of state dynamics in the model (see Section 3.4).

which we specify below). 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 submarket, and the fact that the firm's states are fully observable. 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.¹⁷

Because a dynamic pricing contract is a time path 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 the object:

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

The elements of a recursive contract ω are the following. Firstly, the contract specifies the price p that is to be charged to each one of the n incumbent customers of the firm. Secondly, the contract specifies the vector $\mathbf{x}'(n'; \mathbf{s}')$ of continuation payoffs that are promised by the firm to each buyer under every new size n' and/or a new exogenous state $\mathbf{s}'$.¹⁸ Hence, by stationarity, conditional on a fixed exogenous state $\mathbf{s}$, contracts are rewritten every time the seller changes sizes, and they remain in place for as long as the firm remains with the same size. Similarly, conditional on a fixed size n , a new contract is implemented when the exogenous state vector $\mathbf{s}$ first changes, but otherwise the contract remains in place. For as long as neither state changes, $x'(n; \mathbf{s}) = x$.¹⁹

Importantly, since meeting rates in a market x are sufficiently summarized by the market's tightness $\theta(x; \varphi)$, which is taken parametrically by both sellers and buyers, and because in equilibrium market tightness will adjust via a revenue-equalization condition across markets, individuals in this economy need not keep track of the evolution of the cross-sectional distribution of agents over time. This means that the distribution of agents over aggregate states is not a payoff-relevant state from the point of view of the individual agent. As we shall see in more detail in Section 3.4, the distribution can be characterized precisely and in isolation once the set of equilibrium markets and their corresponding tightnesses have been determined. In this sense, we say that our recursive Markov perfect equilibrium is, in fact, *block-recursive*: the equilibrium policy functions will depend on the aggregate state of the economy only through the exogenous Markov state $\varphi \in \Phi$, but not through the distribution of agents across individual states (n, z) .

In anticipation of this result, we can now establish that the vector of *payoff-relevant* states that are exogenous to the agent reduces to the two exogenous Markov states:

$$\mathbf{s} = (z, \varphi) \in \mathcal{Z} \times \Phi$$

¹⁷ Though we do not rule out the possibility of non-stationary equilibria, we focus our analysis on the stationary solution for simplicity. If agents were risk-averse, for instance, non-stationary solutions would emerge naturally. One such case is the model of the labor market with job-to-job transitions presented in Shi (2009), where workers' utilities have some curvature and firms optimally choose a front-loaded wage profile in time to entice workers to remain matched to them.

¹⁸ Conventionally, we subsequently use bold characters to denote vectors. Scalars are written in light characters.

¹⁹ As we shall see, there is a one-to-one mapping between continuation utilities and price levels, so that p also remains fixed for as long as the firm does not experience a change in its state vector.

The dimensionality of the problem has indeed been greatly reduced. Since distributions are not part of the agent's state vector, the state space has finite dimension. Moreover, because size is a discrete random variable and the firm meets or loses at most one customer every instant, we have $n' \in \{n-1, n, n+1\}$. Therefore, for any given $(n, \mathbf{s})$, the object $\mathbf{x}'(n'; \mathbf{s}') \in \omega$ has dimension $3 \times k_z \times k_\varphi$, and the firm must choose $\omega \in \Omega$, where $\Omega := [0, v] \times \mathcal{X}^{\bar{k}}$ denotes the set of admissible contracts, and $\bar{k} := (3k_z k_\varphi - 1) \in \mathbb{N}$.

Let us now characterize the value functions of each type of agent in the economy. These functions must internalize the firm's size evolution, as sellers of different sizes might optimally choose to post new contracts and, by the no-price-discrimination assumption, these new contracts will in turn affect captive customers' future life-time utilities. Likewise, the value function of buyers must be internalized by sellers, as the latter must understand how buyers sort across firms in order to design contracts optimally.

Inactive Buyers If a buyer is presently inactive, its expected value is equal to $U^B(\varphi)$ in state $\varphi \in \Phi$. The buyer enters the submarket that offers the highest valuation, and therefore:

$$U^B(\varphi) = \max_{\hat{x}(\varphi) \in \mathcal{X}} u^B(\hat{x}(\varphi); \varphi) \quad (1a)$$

where $u^B(\cdot)$ satisfies the Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} ru^B(x; \varphi) = & -c(\varphi) + \mu(\theta(x; \varphi)) \left(x - u^B(x; \varphi) \right)^+ \\ & + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi' | \varphi) \left(u^B(x; \varphi') - u^B(x; \varphi) \right) \end{aligned} \quad (1b)$$

for any $x \in \mathcal{X}$, and the notation $(\cdot)^+$ stands for $\max\{\cdot, 0\}$.²⁰ Equation (1a) states that the inactive buyer searches in the product market that promises the highest expected value, $\hat{x}(\varphi)$. This value is given by equation (1b) evaluated at $x = \hat{x}(\varphi)$. Generally, the value of entering market x incorporates the search cost $c(\varphi) > 0$, and the option value of matching with any one firm within said market. The meeting rate depends on how “crowded” the marketplace is, as measured by the prevailing tightness schedule $\theta(x; \varphi)$, which is taken as given by agents. In case of a successful match, and because sellers can only meet at most one customer every instant, the buyer will instantly join the seller's customer base as long as she finds it profitable to do so, i.e. if $x > u^B(x; \varphi)$. In this event, the size of the seller's customer base will change, and the previously unmatched buyer will obtain the promised value x .²¹ The second line in equation (1b) incorporates the expected change in value due to a change in the aggregate state, from φ to some φ' , occurring at rate $\lambda_\varphi(\varphi' | \varphi)$.²²

²⁰ For a derivation of all the HJB equations in the main text, see Appendix A.1.

²¹ Note that we assume that the inactive buyer returns to market x if unsuccessful in his search. As we argue shortly, this entails no loss in generality.

²² We should point out that notation has been economized in two ways here. First, since the value of inactivity is itself

Since inactive buyers choose to apply to the highest-valuation offer, they must be indifferent across all active market segments. Thus:

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

In equilibrium, a non-zero measure of markets is open, and we let $\mathcal{X}^*(\varphi) := \{x \in \mathcal{X} : \theta(x; \varphi) > 0\} \subseteq \mathcal{X}$ be the *equilibrium* set of markets in state $\varphi \in \Phi$. Note, in particular, that if a market segment is active, then it must be visited by a non-zero mass of inactive buyers. Hence, for any given aggregate state $\varphi \in \Phi$, we have:

$$\mu(\theta(x; \varphi)) \left(x - u^B(x; \varphi) \right)^+ = \gamma^B(x; \varphi) \quad (2)$$

for all $x \in \mathcal{X}$, where:

$$\gamma^B(x; \varphi) := c(\varphi) + ru^B(x; \varphi) - \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi' | \varphi) \left(u^B(x; \varphi') - u^B(x; \varphi) \right) \quad (3)$$

is the expected option value of matching for the buyer in market x . Equation (2) describes how inactive buyers sort in equilibrium. It states that all active market segments must make inactive buyers ex-ante indifferent between any value that is offered in equilibrium. In particular, across active market segments x , we have $u^B(x; \varphi) = U^B(\varphi)$, and thus:

$$\Gamma^B(\varphi) := \max_{x \in \mathcal{X}} \gamma^B(x; \varphi)$$

In words, the option value of matching is independent across all active markets. This indifference then uniquely pins down the market tightness of any one market in $\mathcal{X}^*$.

A few remarks are in order. First, if x belongs to the equilibrium set of markets for a given aggregate state φ (so that $\theta(x; \varphi) > 0$), then it must be that $(x - U^B(\varphi))^+ > 0$, or else no buyer would visit the market. Thus, any matched customer is ex-post better off than any inactive buyer. Inactive buyers are always willing to enter into the product market, and therefore the seller's optimal design of the contract ω need not incorporate a participation constraint. Second, equation (2) defines the $\theta(\cdot; \varphi)$ schedule over the entire support $\mathcal{X}$, and thus it is used by agents to form beliefs about market tightness on *both* equilibrium and off-equilibrium markets. This restriction, which is implicit in the bulk of the competitive search literature going back to [Moen \(1997\)](#), imposes a form of trembling-hand stability in beliefs, and ensures the existence of a stable rational-expectation equilibrium.²³ In particular, no firm (or coalition of firms) can possibly make a profitable off-equilibrium deviation, for in this case beliefs dictate that buyers would remain indifferent and thus the equilibrium allocation would be unaffected. Although we

an equilibrium object, we write $\theta(x; \varphi)$ to mean $\theta(x; \varphi, U^B(\varphi))$. In Section 3.5 we shall explore how to pin down $U^B(\varphi)$ via a fixed-point argument. Second, since θ is parametric for the agent, $u^B(x; \varphi)$ is actually short-hand notation for $u^B(x; \varphi, \theta(x; \varphi))$. Similar concise notation will be used on the remaining value functions.

²³ For an in-depth discussion on the game-theoretical foundations of this assumption, see [Galenianos and Kircher \(2009, 2012\)](#).

recognize the possibility that other type of equilibria may exist under alternative characterizations of agents' beliefs, in what follows we will focus on this type of rational-expectations equilibrium for the sake of tractability.

With these remarks in place, we note that an implication of equation (2) is that, for each $\varphi \in \Phi$, $\theta(x; \varphi)$ is an increasing function of $x \in \mathcal{X}$. This result is intuitive: more profitable offers attract a larger mass of buyers per seller, while sellers offering less favorable contracts to the buyer can expect to find a match sooner. In equilibrium, firms design contracts for which a low buyer meeting rate μ can be compensated with a high enough promised expected continuation value x . The buyer-to-seller ratio is also increasing in $U^B(\varphi)$: 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 customer who is currently consuming the homogeneous good from a firm of size $n \in \mathbb{N}$ and idiosyncratic productivity $z \in \mathcal{Z}$, under contract $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$. The contract delivers the promised value x to the customer, and it specifies the current price $p \in [0, v]$ and the new continuation promises $\mathbf{x}'(n'; \mathbf{s}')$, to be delivered by the seller after some n -to- n' or $\mathbf{s}$ -to- $\mathbf{s}'$ transition.

The value for the buyer is given by the following HJB equation:

$$\begin{aligned} rV^B(n, \omega; \mathbf{s}) = & v - p + (\delta_f + \delta_c) \left(U^B(\varphi) - V^B(n, \omega; \mathbf{s}) \right) + (n-1)\delta_c \left(x'(n-1; \mathbf{s}) - V^B(n, \omega; \mathbf{s}) \right) \\ & + \eta \left(\theta(x'(n+1; \mathbf{s}); \varphi) \right) \left(x'(n+1; \mathbf{s}) - V^B(n, \omega; \mathbf{s}) \right) \\ & + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left(x'(n; z', \varphi) - V^B(n, \omega; \mathbf{s}) \right) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left(x'(n; z, \varphi') - V^B(n, \omega; \mathbf{s}) \right) \end{aligned} \quad (4)$$

where $U^B(\varphi)$ is defined in equation (1a). The right side of equation (4) has different additive terms. In the first line, the first term $(v - p)$ shows flow surplus for the agreed-upon price $p \in \omega$; the second term states the possibility of separation, due to either the destruction of the firm or the destruction of the match, in which case the customer ceases to consume and becomes inactive; the third term includes the event in which any customer of the firm (except for the buyer in question) separates, in which case the firm becomes size $n - 1$ and changes the promised value to $x'(n - 1; \mathbf{s}) \in \omega$ for all those customers that remain captive. The second line is the expected change in value due to the firm successfully attracting a customer with its currently posted offer, in which event the seller becomes size $n + 1$ and implements value $x'(n + 1; \mathbf{s}) \in \omega$. Because the seller cannot price-discriminate between the n incumbent customers and the newcomer, this event affects the match value for all captive buyers in the same way. The likelihood of the event depends upon how tight market $x'(n + 1; \mathbf{s})$ is. Finally, the last line of equation (4) includes the change in value due to an exogenous shock, whether of idiosyncratic (first term) or aggregate (second term) nature.

Importantly, equation (4) shows the sense in which the customer must anticipate the future path of prices. 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 thus arises endogenously because the seller is unable to discriminate across its customers. Let us now describe how the seller optimally chooses to do so.

Incumbent Sellers Consider a seller with idiosyncratic productivity $z \in \mathcal{Z}$ who is endowed with $n \in \mathbb{N}$ captive customers. This seller currently follows the price strategy set up by its past contracts, under which its customers agreed to trade in exchange for a promised value of x . The problem of such a seller, whose expected value is denoted by $V^S(n, x; \mathbf{s})$, is to select a new contract $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$ for all of its n customers so as to maximize its life-time value. The seller's HJB equation is:

$$rV^S(n, x; \mathbf{s}) = \max_{\omega \in \Omega} \left\{ pn - \mathcal{C}(n; \mathbf{s}) + \delta_f \left(V_0^S(\varphi) - V^S(n, x; \mathbf{s}) \right) \right. \\ + n\delta_c \left(V^S(n-1, x'(n-1; \mathbf{s}); \mathbf{s}) - V^S(n, x; \mathbf{s}) \right) \\ + \eta \left(\theta(x'(n+1; \mathbf{s}); \varphi) \right) \left(V^S(n+1, x'(n+1; \mathbf{s}); \mathbf{s}) - V^S(n, x; \mathbf{s}) \right) \\ + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left(V^S(n, x'(n; z', \varphi); z', \varphi) - V^S(n, x; \mathbf{s}) \right) \\ \left. + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left(V^S(n, x'(n; z, \varphi'); z, \varphi') - V^S(n, x; \mathbf{s}) \right) \right\} \quad (5)$$

where $V_0^S(\varphi)$ denotes the value of having no customers (which we derive below).²⁴ The right side of equation (5) has the following components. The term $[pn - \mathcal{C}(n; \mathbf{s})]$ is the seller's flow profits, composed of revenue from selling n units, 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 she instantly loses all customers and enters the pool of potential entrants. The second line states that the seller faces the probability that any one of its n customers separates from the match, in which case the seller shrinks down to size $(n-1)$ and delivers the promised value $x'(n-1; \mathbf{s}) \in \omega$. The third line shows that, by posting a new offer $x'(n+1; \mathbf{s}) \in \omega$, the seller attracts a certain mass of buyers and faces a probability of increasing its size to $n+1$, which the seller takes as given. Namely, the seller understands the sorting behavior of buyers across states for different promised values. In the event of a successful match, the seller would implement the new continuation value, and its state vector would transition from $(n, x; \mathbf{s})$ to $(n+1, x'(n+1; \mathbf{s}); \mathbf{s})$. Finally, the value of the firm could change exogenously because of a state transition from $\mathbf{s} = (z, \varphi)$ to either (z', φ) or (z, φ') , as captured by the last two lines in equation (5).

When choosing a contract ω , a seller in state $(n, x; \mathbf{s})$ is constrained by the following condition:

$$V^B(n, \omega; \mathbf{s}) \geq x \quad (6)$$

Equation (6) is a *promise-keeping (PK) constraint*: the value that each buyer of the firm obtains under the contract must be weakly greater than the value x that was promised to her. In other words, in

²⁴ For $n = 1$, we note that $x'(n-1; \mathbf{s}) = \emptyset$, $\forall \mathbf{s} \in \mathcal{Z} \times \Phi$, and denote $V^S(n-1, x'(n-1; \mathbf{s}); \mathbf{s})$ by $V_0^S(\varphi)$.

its choice of the contract, the seller must honor the promises that were made in the past. This constraint is in place due to our commitment assumption on the seller's side.

Potential Entrants Finally, let us describe the problem of an outside firm. These 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 an initial contract. Prior to start selling the good, they must also draw an initial productivity level z_0 from the π_z distribution. Given the realization of z_0 , the contract is the singleton $\{x'(1; z_0, \varphi)\}$, specifying the utility promised to the first customer of the firm under state (z_0, φ) . Thus, the potential entrant chooses amongst a menu of contracts, $\omega_0(\varphi) := \{x'(1; z_0, \varphi)\}_{z_0 \in \mathcal{Z}}$, contingent on each realization of productivity at entry. Note, in particular, that the potential entrant's contract does not specify a price level for the first customer, for this choice is made ex-post, i.e. once the customer has been acquired (and the seller faces problem (5) for $n = 1$).

The ex-ante value of the potential entrant in aggregate state φ is, therefore:

$$rV_0^S(\varphi) = -\kappa + \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) v_0^S(z_0, \varphi) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi' | \varphi) (V_0^S(\varphi') - V_0^S(\varphi)) \quad (7)$$

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

$$v_0^S(z_0, \varphi) := \max_{x'(1; z_0, \varphi)} \eta \left(\theta(x'(1; z_0, \varphi); \varphi) \right) \left(V^S(1, x'(1; z_0, \varphi); z_0, \varphi) - V_0^S(\varphi) \right)$$

Note that, because this firm does not yet have any customers at the time of choosing the contract, the potential entrant is not subject to a PK constraint. Once again, the firm understands how inactive buyers sort across markets, as the $\theta(\cdot; \varphi)$ schedule is taken as given.

We assume free entry into the product market for the first customer. Since the total mass of sellers adjusts freely, this assumption implies that, in equilibrium, more firms will enter the economy as long as the expected value of posting a contract exceeds the set-up cost κ . As more potential entrants flood into the market, this expected value is pushed down to $\kappa > 0$. Therefore, in an equilibrium with positive entry in all aggregate states, it must be that:

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

This condition is informative about the average tightness in the entry market. Since, by construction, firms enter with one customer, the free-entry condition pins down the average market tightness among firms of size one in the cross-section of initial productivity levels. In particular, from equation (7), for all states $z_0 \in \mathcal{Z}$ such that $\theta(x'(1; z_0, \varphi); \varphi) > 0$, it must be the case that:

$$\kappa = \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) v_0^S(z_0, \varphi) \quad (8)$$

for any given $\varphi \in \Phi$.

3.3 Optimal 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 profit-maximizing contract leads to an allocation of utilities in which the *joint surplus* (i.e. the sum of the expected values of the seller and all of its customers) is maximized. Moreover, for any allocation that maximizes the joint surplus, there always exists a price that maximizes the seller's profit subject to the PK constraint. Thus, the seller's and the joint surplus problems are equivalent. As we shall see, this observation allows us to simplify the state space greatly, and it renders an analytical equilibrium characterization.

To start, consider a typical firm whose state vector is $(n, x; \mathbf{s})$, where recall that $\mathbf{s} := (z, \varphi)$ collects the exogenous states. As seen in the last section, the optimal contract $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$ can be obtained as the solution to the problem of the seller, described in (5). 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 (6) will hold with equality. Thus, to economize on notation, for the remainder of the paper we write x in place of $V^B(n, \omega; \mathbf{s})$ with no loss of generality.

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. Because the contract cannot discriminate across different customers, the definition of the joint surplus takes the following simple form:

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

In Appendix A.2, we show that the joint surplus can be written in the following HJB representation:

$$\begin{aligned} rW(n, x; \mathbf{s}) = \max_{\mathbf{x}'(n'; \mathbf{s}')} & \left\{ n \left(v + (\delta_f + \delta_c) U^B(\varphi) \right) - \left(\mathcal{C}(n; \mathbf{s}) + \eta \left(\theta(x'(n+1; \mathbf{s}); \varphi) \right) x'(n+1; \mathbf{s}) \right) \right. \\ & + n\delta_c \left(W(n-1, x'(n-1; \mathbf{s}); \mathbf{s}) - W(n, x; \mathbf{s}) \right) - \delta_f W(n, x; \mathbf{s}) \\ & + \eta \left(\theta(x'(n+1; \mathbf{s}); \varphi) \right) \left(W(n+1, x'(n+1; \mathbf{s}); \mathbf{s}) - W(n, x; \mathbf{s}) \right) \\ & + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left(W(n, x(z', \varphi); z', \varphi) - W(n, x; \mathbf{s}) \right) \\ & \left. + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left(W(n, x(z, \varphi'); z, \varphi') - W(n, x; \mathbf{s}) \right) \right\} \quad (9) \end{aligned}$$

Intuitively, equation (9) 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 $n(v + (\delta_f + \delta_c) U^B(\varphi))$ represents the aggregate flow surplus for all the n customers of the firm, which is composed of the sum of per-customer

consumption gains, v , and the expected per-customer gains from separation, $(\delta_f + \delta_c)U^B(\varphi)$. The second component in parentheses are the total costs of the match for the seller, which include total operating costs, $\mathcal{C}(n; \mathbf{s})$, and the expected costs of offering the life-time a continuation value of $x'(n+1)$ to the new incoming customer, adjusted by the endogenous rate at which a new customer joins the match. The second and third lines include the change in the expected joint surplus when the match shrinks (because any one of the n customers leaves, or the firm is destroyed), or grows (because a new customer joins). Finally, the last two terms incorporate expected changes in the joint surplus that are due to exogenous shocks.

We can now state the main result:

Proposition 1 (Joint Surplus Problem) *The following properties hold:*

1. *The firm's and the joint surplus problems are equivalent. In particular:*
 - (a) *Given a contract $\omega^* = \{p, \mathbf{x}'(n', \mathbf{s}')\}$ that maximizes problem (5), $\mathbf{x}'(n', \mathbf{s}')$ is a solution to problem (9).*
 - (b) *For every vector $\mathbf{x}'(n', \mathbf{s}')$ that solves (9), there exists a price level p for which $\{p, \mathbf{x}'(n', \mathbf{s}')\}$ is a solution to problem (5).*
2. *For any state $(n, x; \mathbf{s})$, the joint surplus is invariant to x , namely we can write $W(n, x; \mathbf{s}) = W_n(\mathbf{s})$.*

Proof. See Appendix A.2.

The first part of Proposition 1 establishes that the profit-maximizing contract can be found by solving an alternative and equivalent problem, given by (9). In this problem, the contract is designed so as to maximize the profits of all the parties involved in a utilitarian manner, 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 configuration of future states), and both agents have linear preferences, for any configuration of states there always exists a menu composed of a price and promised utility pair that redistributes rents among the seller and its customers for such an allocation to be implemented. Moreover, because the seller commits to the terms of the contract, the allocation is unique in that no deviation from joint surplus maximization can deliver higher profits for the seller.

The second part of the proposition thus follows immediately from the first one, and explains why problem (9) is much simpler to solve than the firm's problem, (5). It states that the maximized surplus is invariant to the rent-sharing components of the contract. Conveniently, this means that the problem can be split in two stages. In the first stage, 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 the second stage, the price level is set so as to implement such an allocation, thereby splitting and distributing rents among the $(n+1)$ agents involved in an incentive-compatible way (that is, ensuring that PK binds). Further, the surplus is also constant in the firm's previous promise, since x is a predetermined state that was chosen optimally in the prior stage of the firm. By Markov perfection and completeness, the size n

and current exogenous state $\mathbf{s}$ serve as sufficient statistics to determine the current surplus-maximizing policies.

While the equivalence between the joint-surplus problem and the decentralized problem is a familiar result in the literature on complete contracts with commitment and transferrable utility, here we show that it can also result from, and provide great analytical tractability to, a dynamic model with two-sided heterogeneity and meaningful firm dynamics.²⁵ In particular, by the invariance of W to (p, x) , the joint surplus can be expressed as:

$$W_n(\mathbf{s}) = W(n, x; \mathbf{s}), \quad \forall n \in \mathbb{N}$$

for some sequence $\{W_n(\mathbf{s})\}_{n=1}^{+\infty}$ of positive real numbers, given $\mathbf{s}$. As a result, the policy that solves problem (9) is not a function of x , and neither is the optimal price level.

Characterization Let us characterize the equilibrium policies that result from problem (9). Firstly, by the indifference condition among inactive buyers (equation (2)), we know that:

$$\eta(\theta(x(\mathbf{s}); \varphi)) = \eta \circ \mu^{-1} \left(\frac{\Gamma^B(\varphi)}{x(\mathbf{s}) - U^B(\varphi)} \right) \quad (10)$$

for all $\mathbf{s} \in \mathcal{Z} \times \Phi$. Recall that the option value of matching for the buyer is constant across markets by buyers' indifference, i.e. for each φ we have that $\Gamma^B(\varphi) = \gamma^B(x(\mathbf{s}); \varphi)$, $\forall x(\mathbf{s})$, where γ^B was defined in equation (3). By Assumption 2(i) and continuity of $\theta(x; \varphi)$ on $x \in \mathcal{X}$, equation (A.8a) describes the maximization of a continuous function over a compact support $\mathcal{X}^{\bar{k}}$, so there exist $\mathbf{x}_n(\mathbf{s}')$, $x_n^+(\mathbf{s})$, and $x_n^-(\mathbf{s})$ which are solutions to the joint surplus problem. Once again, note that we index these policies by n , but not x .

Henceforth, denote a generic policy by $\{p_n(\mathbf{s}); x_n^+(\mathbf{s}), x_n^-(\mathbf{s}), \mathbf{x}_n(\mathbf{s}')\}$, where the “+” (respectively, “−”) superscript denotes the upsize (respectively, downsize) decision. Let us begin with the choice of x^+ . Firstly, using equation (10) and concavity of $\eta \circ \mu^{-1}$, then $\eta(\theta(x_n^+(\mathbf{s}); \varphi))$ is increasing and concave in $x_n^+(\mathbf{s})$, and there is a unique and interior solution for $x_n^+(\mathbf{s})$ in $\mathcal{X}^{k_z \times k_\varphi}$. Further, since η is differentiable, the following first-order condition is sufficient for optimality:

$$\frac{\partial \eta(\theta(x_n^+(\mathbf{s}); \varphi))}{\partial x_n^+(\mathbf{s})} [W_{n+1}(\mathbf{s}) - W_n(\mathbf{s})] = \frac{\partial \eta(\theta(x_n^+(\mathbf{s}); \varphi))}{\partial x_n^+(\mathbf{s})} x_n^+(\mathbf{s}) + \eta(\theta(x_n^+(\mathbf{s}); \varphi)) \quad (11)$$

Intuitively, the optimal continuation value $x_n^+(\mathbf{s})$ 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 one hand, an increase by one dollar in the promised value x^+ increases the joint surplus by the amount $[W_{n+1} - W_n] > 0$ in case the seller makes a size transition.²⁶ These gains

²⁵ For an application of this principle to a search model of firm dynamics in the labor market, see [Schaal \(2017\)](#).

²⁶ By equation (10), a necessary condition for market segment x_n^+ to be active (i.e. $\theta(x_n^+) > 0$) is $x_n^+ > U^B \geq 0$. Therefore, under certain parametric restrictions on Φ , the right-hand side of (11) is strictly positive, and thus it must be the case that $[W_{n+1} - W_n] > 0$ on the left-hand side. Namely, for a given $\mathbf{s}$, $\{W_n(\mathbf{s})\}$ is a strictly increasing sequence in n .

must then be weighted by the marginal effect of x_n^+ on the likelihood that the firm meets a new customer. On the other hand, for every dollar spent on the new continuation value x_n^+ , 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; and second, the decrease in the price level, by $\eta(\theta(x_n^+; \varphi))$ dollars, which is required by incentive compatibility (see equation (A.6) in Appendix A.2).

As for the choices of $\mathbf{x}(\mathbf{s}')$ and $x^-(\mathbf{s})$, note that these do not feature anywhere in equation (9) once we impose that the joint surplus is invariant to promised utilities. Therefore, $\{\mathbf{x}_n(\mathbf{s}'), x_n^-(\mathbf{s})\}$ cannot be determined by a surplus-maximizing condition similar to (11). Instead, these values are purely redistributive: the only dimension in which they matter is the price level, and thus they affect only the way in which the total surplus is split (i.e. the terms of trade). In particular, since beliefs are pinned down by equation (10), the firm's choice must be consistent with the sorting behavior of inactive buyers. By symmetry, the optimal downsizing choice for a size- n must be consistent with the optimal upsizing choice for a firm of size $(n - 2)$, or $x_n^-(\mathbf{s}) = x_{n-2}^+(\mathbf{s})$. Similarly, when transitioning to another state, equation (10) and symmetry require that $x_n(\mathbf{s}') = x_{n-1}^+(\mathbf{s}')$. Therefore, a firm of a certain size must promise the same utility regardless of its previous size, given the state. Similarly, a firm of a certain productivity type must promise the same utility regardless of previous productivities, given its size. In short, a firm's optimal continuation utility is history-independent.

Since, as we have just argued, there are unique solutions to the two stage problems, the optimal contract is unique for each $(n; \mathbf{s})$, and denoted by $\omega_n(\mathbf{s})$. By commitment, we know that firms must deliver what they promised in their prior contracts, so:

$$x_n(\mathbf{s}) = x_{n-1}^+(\mathbf{s}), \quad \forall (n; \mathbf{s}) \in \mathbb{N} \times \mathcal{Z} \times \Phi$$

with $x_{n-1}^+(\mathbf{s})$ given by equation (11). Thus, the set $\mathcal{X}^*(\varphi)$ of active market segments in equilibrium is comprised of a countable set of sequences, namely $\mathcal{X}^*(\varphi) = \{\mathcal{X}_{\underline{z}}(\varphi), \dots, \mathcal{X}_{\bar{z}}(\varphi)\}$, with a typical sequence given by:

$$\mathcal{X}_z(\varphi) := \{x_n(z, \varphi)\}_{n=1}^{+\infty}$$

where each element $x_n(z, \varphi)$ is given by the solution to (11). In equilibrium, each $\mathcal{X}_z(\varphi)$ is a *decreasing* sequence in n : though indifferent ex-ante, buyers ex-post prefer to be matched to smaller firms.²⁷ Figure 2 depicts the different markets in equilibrium, for a given state $\mathbf{s}$. All equilibrium markets are distributed across the θ schedule, in blue, and inactive buyers are ex-ante indifferent between all contracts.

Once the set of markets in equilibrium has been determined the remaining equilibrium objects follow directly. Defining for brevity $\theta_n(z, \varphi) := \theta(x_n(z, \varphi); \varphi)$, the equilibrium set of market tightnesses is given by $\Theta^*(\varphi) = \{\Theta_{\underline{z}}(\varphi), \dots, \Theta_{\bar{z}}(\varphi)\}$, where $\Theta_z(\varphi) := \{\theta_n(z, \varphi)\}_{n=1}^{+\infty}$, and each element $\theta_n(z, \varphi)$ is given by equation (10). Note that, because we have found that $\theta(x; \varphi)$ is an increasing and continuous function of x , each $\Theta_z(\varphi)$ sequence is decreasing in n . Intuitively, by offering higher ex-post values, smaller firms attract more buyers, so the buyer-to-seller ratio is higher in those markets. Since smaller firms attract a

²⁷ A proof of equilibrium existence, uniqueness, and monotonicity is currently work in progress.

Figure 2: Equilibrium tightness function $\theta(x) = \mu^{-1} \left(\frac{\Gamma^B}{x - U^B} \right)$, and equilibrium markets across different n 's, given $\mathbf{s}$.

higher number of customers within any time interval, these firms grow faster as a result. Indeed, after a small interval of size $\Delta \approx 0$, the change in the size of the firm is given by:

$$n_{t+\Delta} - n_t = \begin{cases} 1 & \text{w/prob. } \eta(\theta_{n_t+1}(z, \varphi))\Delta + o(\Delta) \\ -1 & \text{w/prob. } n_t\delta_c\Delta + o(\Delta) \\ -n_t & \text{w/prob. } \delta_f\Delta + o(\Delta) \\ 0 & \text{else} \end{cases} \quad (12)$$

where $o(\Delta)$ collects higher-order terms and satisfies $\lim_{\Delta \rightarrow 0} \frac{o(\Delta)}{\Delta} = 0$. For small firms, the probability of attracting a new customer is relatively higher. As the firm grows, the attrition probability increases proportionally to firm's size as the attraction probability decays due to a decrease in x_n , leading to slower firm growth. Eventually, conditional on survival, firm converge to a stationary size. In this way, the model rationalizes the observation that smaller and younger firms tend to grow more rapidly than larger and more mature ones (see footnote 8 for a brief review of this literature).

Using similar notation, the equilibrium set of joint surpluses is the set $\mathcal{W}^*(\varphi)$, whose typical element is a sequence $\mathcal{W}_z(\varphi) := \{W_n(z, \varphi)\}_{n=1}^{+\infty}$. In turn, using equation (9), each sequence $\mathcal{W}_z(\varphi)$ can be found as the solution to the following second-order difference equation in n :

$$\begin{aligned} (r + \delta_f)W_n(z, \varphi) = & n \left(v + (\delta_f + \delta_c)U^B(\varphi) - \delta_c(W_n(z, \varphi) - W_{n-1}(z, \varphi)) \right) - \mathcal{C}(n; z, \varphi) \\ & + \eta(\theta_{n+1}(z, \varphi)) \left(W_{n+1}(z, \varphi) - W_n(z, \varphi) - x_{n+1}(z, \varphi) \right) \\ & + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) \left(W_n(z', \varphi) - W_n(z, \varphi) \right) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) \left(W_n(z, \varphi') - W_n(z, \varphi) \right) \end{aligned} \quad (13)$$

Finally, the set of equilibrium price sequences is $\mathcal{P}^*(\varphi)$, with typical sequence $\mathcal{P}_z(\varphi) = \{p_n(z, \varphi)\}_{n=1}^{+\infty}$ for firms of type z , and each price given implicitly by the PK constraint. Using equation (4), this means:

$$\begin{aligned}
p_n(z, \varphi) = & \underbrace{v - rx_n(z, \varphi)}_{\substack{\in [0, v] \\ \text{Baseline price}}} + \underbrace{\delta_f(U^B(\varphi) - x_n(z, \varphi))}_{\substack{\leq 0 \\ \text{Exit discount}}} + \underbrace{\eta(\theta_{n+1}(z, \varphi))(x_{n+1}(z, \varphi) - x_n(z, \varphi))}_{\substack{\leq 0 \\ \text{Growth discount}}} \\
& + \underbrace{n\delta_c \left(\frac{U^B(\varphi) + (n-1)x_{n-1}(z, \varphi)}{n} - x_n(z, \varphi) \right)}_{\substack{\geq 0 \\ \text{Separation penalty}}} + \underbrace{\sum_{z' \in \mathcal{Z}} \lambda_z(z'|z)(x_n(z', \varphi) - x_n(z, \varphi))}_{\text{Idiosyncratic-shock adjustment}} \\
& + \underbrace{\sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)(x_n(z, \varphi') - x_n(z, \varphi))}_{\text{Aggregate-shock adjustments}}
\end{aligned} \tag{14}$$

The optimal price level for a firm of type (n, z) can be decomposed into the following parts. The first one, $v - rx_n(z, \varphi)$, 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 the firm did not change size going forward. We call this term the *baseline price level*. The remaining terms in (14) introduce the necessary adjustments for possible changes in firm states. These adjustments entice the customers to accept the terms of trade. First, the firm offers a price discount of $\delta_f(U^B(\varphi) - x_n) \leq 0$ to compensate the customer for the expected loss in value in the event of the firm exiting the market. We label this type of discount as the *exit discount*. The term $\eta(\theta_{n+1})(x_{n+1} - x_n)$ is a discount for the possibility that the firm grows, because in that case the total surplus is split among more buyers and each individual customer obtains a lower share (indeed, recall $x_{n+1} \leq x_n$). This discount is labeled as the *growth discount*. In contrast, the firm penalizes customers for separating from the match, because a separation hurts the seller but has a positive externality on all the customers that remain matched (as $x_{n-1} \geq x_n$). We call this the *separation penalty*. If a separation occurs, then the separating customer obtains U^B , and the remaining non-separating customers each obtain x_{n-1} . This amounts to an average value of $\frac{U^B + (n-1)x_{n-1}}{n}$ per customer, which determines the size of the penalty. The remaining two terms in equation (14) adjust the price level for expected changes in the exogenous states. These may adjust the price either upward or downward, depending on the nature of the shock.

Note that the $\{p_n(\mathbf{s})\}_{n=1}^{+\infty}$ sequence is non-monotonic in x and, therefore, there may exist size non-monotonicities in price levels in equilibrium. In particular, while the baseline and exit discount components are monotonically increasing in x , the growth discount and separation penalty components respond non-monotonically to changes in x . When x is a decreasing function of n (so that smaller firms offer more ex-post attractive contracts to buyers and grow faster as a result), prices are relatively low and increase slowly. However, depending on the curvature of x_n to n , larger firms may prefer to begin lowering their price, as further firm growth requires more generous discounts that, by commitment, are needed to entice customers to remain matched. Figure 3 shows one such example.

Taking stock This completes the characterization of the equilibrium. To recap, we have obtained price dispersion on the basis of firm heterogeneity in size and productivity. When firms are small, their

Figure 3: Example of a non-monotonic equilibrium price in firm size. Price components expressed in absolute values. In this example, small firms charge an increasing price schedule with size. When firms reach a certain size, however, growth discounts must become relatively more generous, and prices start declining with size.

current revenue is relatively small, so they prefer to promise a high continuation value to their buyers in order to generate a high meeting rate. Accordingly, market tightness is high because most buyers are attracted to such type of firms. As the firm matures, the price increases because the seller becomes increasingly concerned with extracting surplus from its base as the base grows larger. Consequently, the value that accrues to each captive buyer in the firm declines, and so do queue lengths and firm meeting rates. In other words, as in the original intuition behind [Rotemberg and Woodford \(1991\)](#), small and young firms use low prices as a way to invest into larger market shares in the future.

3.4 Block Recursivity and Distributions Dynamics

The equilibrium of the economy described in Section 3.2 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, the characterization of individuals' decisions need has been silent on the exact composition of buyers and sellers within each specific market segment, or the evolution of this distribution. This is because the equilibrium allocation is *block-recursive*.

In our model, block recursivity arises from two key ingredients. On the one hand, we assume that search is directed, and thus sellers' offers are not contingent on the identity of the applicant (in particular, they are not contingent on the applicant's outside option). 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 not have to forecast the evolution of aggregates over future states of the economy. On the other hand, the indifference condition across all markets among inactive buyers (equation (2)), and free entry of firms in the market for one customer (equation (8)), together imply that the equilibrium tightness on each market adjusts to be consistent with agents' beliefs. This has allowed us to inductively construct the entire sequence of buyer-to-seller ratios without ever having to specify the exact composition of agent types within each market segment.²⁸

²⁸ [Kaas and Kircher \(2015\)](#) exploit similar insights to obtain tractability. In our continuous-time environment, the indifference

In fact, market tightness is also a sufficient statistic for characterizing the composition of agents across markets. Further, since in equilibrium markets are uniquely indexed by the idiosyncratic state of the firms who inhabit them, the distribution of sellers' sorting across markets corresponds exactly to the firm distribution of the economy. Thus, the model allows for an exact characterization of the distribution of agents (on and off equilibrium), and does not require approximation techniques such as those of [Krusell and Smith \(1998\)](#). This makes our environment particularly apt to study aggregate product market dynamics.

Let us describe state dynamics precisely. For this section, we index all equilibrium measures explicitly by time t , because the flow equations describing their corresponding dynamics will hold both in and out of steady state. Let us also subsequently denote all *jump variables* (i.e. variables which change discontinuously immediately following an aggregate shock) by making their dependence on φ explicit. In contrast, *stock variables*, which evolve smoothly over the cycle, are not indexed by φ because, in this case, the state-dependence is fully encoded in the time subscript.

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 $t \geq 0$. Using the results from the previous section, recall that all such firms are seeking new customers in market $x_{n+1,t}(z, \varphi)$. Therefore, letting $B_{n+1,t}^I(z, \varphi)$ be the measure of (inactive and searching) buyers within market $x_{n+1,t}(z, \varphi)$, market tightness must adjust at every $t \geq 0$ to guarantee that:

$$B_{n+1,t}^I(z, \varphi) = \theta_{n+1,t}(z, \varphi) \cdot S_{n,t}(z) \quad (15)$$

for all $n \in \mathbb{N}$, where recall that $\theta_{n+1}(\cdot)$ is short-hand notation for the tightness in market $x_{n+1}(\cdot)$, i.e. $\theta_{n+1}(\cdot) \equiv \theta(x_{n+1}(\cdot); \varphi)$. Using that $\eta(\theta) = \theta\mu(\theta)$, equation (15) can be written as $\mu(\theta_n(z, \varphi))B_n^I(z, \varphi) = \eta(\theta_n(z, \varphi))S_{n-1,t}(z)$, 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 an additional customer.²⁹

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

$$B_{n,t}^A(z) = nS_{n,t}(z) \quad (16)$$

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

$$\forall \varphi \in \Phi, \forall t \geq 0 : \underbrace{\sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} B_{n,t}^A(z)}_{=B_t^A} + \underbrace{\sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} B_{n,t}^I(z, \varphi)}_{=B_t^I} = 1 \quad (17)$$

and free-entry structure are natural: since firms can only gain one customer at a time, we cannot impose free entry across firms of all sizes. In discrete time, an alternative approach would have been to dispense of the indifference condition among inactive buyers, and assume instead free entry of firms across *all* contracts (not just in the entrant's market). This is the approach followed in the on-the-job search models of [Menzio and Shi \(2010, 2011\)](#) and [Schaal \(2017\)](#).

²⁹ Adding up across idiosyncratic states (n, z) , we obtain the more general identity that the total measure of inactive buyers who become customers of a seller must be equal to the total measure of sellers who obtain a new customer.

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

Note that we have written $B_{n,t}^I(z, \varphi)$ as a jump variable (making it explicitly depend on φ), because the measure of customers looking to be matched with a specific type of firm responds instantaneously to the business cycle to guarantee that the indifference condition among unmatched buyers (equation (2)) holds at all dates. In contrast, $S_{n,t}(z)$, for $n \geq 1$, is a stock variable evolving smoothly and it does not respond immediately to changes in φ . Hence, through equation (16), $B_{n,t}^A(z)$ changes smoothly over time, while through equation (15), market tightness $\theta_{n,t}(z, \varphi)$ must jump instantaneously in response to aggregate shocks. Similarly, the mass of potential entrants, denoted $S_{0,t}(\varphi)$, jumps following a φ -shock. By equation (17), then, the distribution of inactive buyers searching on each market shifts around with the business cycle in such a way for the *aggregate* measure of customers to adjust sluggishly over time. Other economic aggregates, such as the average price level and aggregate output, will inherit this property as well.

In sum, the policy functions derived in Section 3.3 are jump variables, while the distributions of active agents (i.e. incumbent firms and active customers) respond slowly to aggregate shocks. The latter are described by a set of flow equations, which we now discuss. Because of this slow adjustment, and because firms take time to move around in the size distribution before they reach a stationary size, the model features sluggish aggregate dynamics.

Figure 4: Firm size transitions in the stationary equilibrium, for a typical firm (right-hand side block) and for entrants (left-hand side block), where z is fixed for expositional ease. Arrow labels indicate flow rates.

Figure 4 provides a graphical depiction of a typical (n, z) -firm's transitions, where we keep the exogenous state z fixed for visual clarity. Formally, the dynamics of the distribution of firms over idiosyncratic states (n, z) are given by a continuous-time Markov process that can be summarized by a set of Kolmogorov Forward (KF) equations taking values on a discrete support. The derivation of the KF equations to follow can be found in Appendix A.3.1.

First, for $n = 1$ we have:

$$\begin{aligned}
\partial_t S_{1,t}(z) &= \pi_z(z) \eta(\theta_{1,t}(z, \varphi)) S_{0,t}(\varphi) + 2\delta_c S_{2,t}(z) + \sum_{\tilde{z} \neq z} \lambda_z(z|\tilde{z}) S_{1,t}(\tilde{z}) \\
&\quad - \left(\delta_f + \delta_c + \eta(\theta_{2,t}(z, \varphi)) + \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) \right) S_{1,t}(z)
\end{aligned} \tag{18}$$

for any $z \in \mathcal{Z}$, where ∂_t denotes the partial derivative with respect to time. The first and second lines on the right-hand side of (18) represent, respectively, flows into and out of state $(1, z)$. Inflows are given by those successful entrants that draw productivity z upon entry, and by the share of incumbents that are either of type z and size $n = 2$ and lose one customer, or that have one customer and transition into the productivity state z from some $\tilde{z} \neq z$. Outflows are given by firms in state $(1, z)$ that either die, lose their only customer, gain a second customer, or transition to a distinct productivity state, $\tilde{z} \neq z$. Note that the aggregate state enters the evolution of the measure of firms only implicitly through its influence on the jump dynamics of S_0 and θ_1 . Therefore, the dynamics of S_1 are smooth.

Similarly, for any $n \geq 2$:

$$\begin{aligned}
\partial_t S_{n,t}(z) &= \eta(\theta_{n,t}(z, \varphi)) S_{n-1,t}(z) + (n+1)\delta_c S_{n+1,t}(z) + \sum_{\tilde{z} \neq z} \lambda_z(z|\tilde{z}) S_{n,t}(\tilde{z}) \\
&\quad - \left(\delta_f + n\delta_c + \eta(\theta_{n+1,t}(z, \varphi)) + \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) \right) S_{n,t}(z)
\end{aligned} \tag{19}$$

The interpretation of this equation is similar to the previous one. Flows into state (n, z) , $n \geq 2$, are given by the share of firms of size $(n-1)$ that obtain their n^{th} customer, the share of firms of size $(n+1)$ that lose one customer, and the share of size- n firms that transition into productivity level z from some state $\tilde{z} \neq z$. Outflows are given by firms in state (n, z) that either die, lose or gain a customer, or suffer a productivity shock.

Finally, the measure of potential entrants, $S_{0,t}(\varphi)$, evolves according to the following ODE:

$$\partial_t S_{0,t}(\varphi) = \delta_f \mathcal{S}_t + \delta_c \sum_{z \in \mathcal{Z}} S_{1,t}(z) - \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \eta(\theta_{1,t}(z_0, \varphi)) S_{0,t}(\varphi) \tag{20}$$

where $\mathcal{S}_t := \sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} S_{n,t}(z)$ is the measure of *incumbent* firms (i.e. firms with one or more customers). Formula (20) shows inflows and outflows in the measure of potential entering firms. Flows into state $n = 0$ (first line) are given by the measure of dying incumbents, and those firms of size one (and any productivity level z) that lose their last remaining customer. Outflows (second line) are given by the measure of potential entrants who successfully attract their first customer, and thereby enter the product market with some idiosyncratic productivity level z_0 drawn from π_z .

Incidentally, equation (20) allows us to derive the *firm entry rate* in the economy, defined as the ratio of *actual* entrants to the total mass of incumbents: $\delta_t^e(\varphi) := [S_{0,t}(\varphi)/\mathcal{S}_t] \sum_{z_0} \pi_z(z_0) \eta(\theta_{1,t}(z_0, \varphi))$. In

particular, from equation (20) we obtain:

$$\delta_t^e(\varphi) + \frac{\partial_t S_{0,t}(\varphi)}{S_t} = \underbrace{\delta_f + \delta_c \sum_z g_{1,t}(z, \varphi)}_{\text{Firm exit rate}}$$

where $g_1 := S_1/S$ is the *share* of incumbent firms of size $n = 1$. The last equation states that the firm entry rate is equal to the firm exit rate, adjusted by the instantaneous relative change in the measure of potential entrants.

Equations (18)-(20) offer a full characterization of out-of-steady-state dynamics, and thanks to the property of block recursivity we can analyze the transition paths of the economy in response to aggregate shocks (Section 4). We can further specialize these equations to obtain the *time-invariant* distribution of firms by equating flows in and out of every possible state (n, z) , so that $\partial_t S_{n,t}(z, \varphi) = 0, \forall n, z, \varphi$.³⁰ In Appendix A.3.1, we derive the aggregate equilibrium measures in the stationary solution explicitly.

3.5 Aggregates and Equilibrium Definition

To close the equilibrium description, we compute economic aggregates. The total time-invariant mass of inactive buyers is $B_t^I = \sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} \theta_{n,t}(z, \varphi) S_{n-1,t}(z)$ (from equation (15)), and the total mass of actives is $B_t^A = 1 - B_t^I$ (from equation (17)). Since each seller produces one unit per customer, the industry's aggregate output is equal to the total mass of active buyers in the economy: $Y_t = B_t^A$. The industry's aggregate price index is given by the weighted sum of all prices, $P_t(\varphi) = \sum_{n=1}^{+\infty} \sum_{z \in \mathcal{Z}} S_{n,t}(z) p_n(z, \varphi)$, where the weight on each price is $S_{n,t}$, i.e. the measure of firms charging price p_n in equilibrium. In Appendix B we show how to embed this aggregate characterization of the equilibrium into a familiar CES environment.

Finally, in the above description, all equilibrium objects have been calculated for a given value of inactivity, $U^B(\varphi)$. To close the equilibrium characterization, we find U^B as the solution to a fixed-point problem. In particular, we must ensure that the equilibrium allocation satisfies the free entry condition, (8). Thus:

$$\kappa = \sum_{z_0 \in \mathcal{Z}} \pi_{z_0}(z_0) \eta(\theta_1(z_0, \varphi | U^B)) \left(W_1(z_0, \varphi | U^B) - x_1(z_0, \varphi | U^B) \right) \quad (21)$$

Note that we are using the notation $(\cdot | U^B)$ to make it explicit that the allocation found in Section 3.3

³⁰ In general, an analytical solution for the stationary distribution does not exist. One exception is when $\delta_f = 0$. In this case, the flow equations amount to a continuous-time Markov chain with reflection bound at zero and exponentially distributed transition times (sometimes called a *birth-death process*), where transition rates are endogenous and state-dependent. For this particular case, one can show (see Appendix A.3.2) that:

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

is the solution for $g_n := S_n/S$ if, and only if, $\{\theta_n\}$ satisfies $\sum_{n=1}^{+\infty} \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{\delta_c^n} < +\infty$.

depends upon the value of the buyer's outside option. Assumption 2, continuity of the value functions, and the Intermediate Value Theorem guarantee that a unique solution $U^B \geq 0$ to equation (21) exists.

We are now ready to define an equilibrium:

Definition 1 *A block-recursive equilibrium (BRE) is, for each aggregate state $\varphi \in \Phi$: (i) a set of value functions $V^S : \mathbb{N} \times \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}_+$ and $V^B : \mathbb{N} \times \Omega \times \mathcal{Z} \rightarrow \mathbb{R}_+$, a value of inactivity $U^B(\varphi) \geq 0$, and a joint surplus $W : \mathbb{N} \times \mathcal{Z} \rightarrow \mathbb{R}_+$; (ii) a contract $\omega_n(\mathbf{s}) = \{p_n(\mathbf{s}), \mathbf{x}_n(\mathbf{s}'), x_n^+(\mathbf{s}), x_n^-(\mathbf{s})\}$ for incumbent firms, and a contract $\omega_0(\varphi) = \{x_1(z_0, \varphi)\}_{z_0 \in \mathcal{Z}}$ for potential entrant firms; (iii) a decision rule $\hat{x}(\varphi)$ for inactive buyers and a promised utility $x_n(\mathbf{s})$ for active buyers of firms of type (n, z) ; (iv) a market tightness function $\theta : \mathcal{X} \rightarrow \mathbb{R}_+$; (v) aggregate measures of agents: $\{S_0(\varphi), \mathcal{S}(\varphi), B^A(\varphi), B^I(\varphi)\}$; and a distribution of sellers and buyers: $\{S_n(z, \varphi), B_n^A(z, \varphi), B_n^I(z, \varphi)\}$; such that:*

- a. *The value functions solve (4) and (5); $U^B(\varphi)$ is the fixed point of the free-entry condition (21); and the joint surplus $W_n(z, \varphi)$ solves (13).*
- b. *The entrant's contract $\omega_0(\varphi)$ solves problem (7), and the incumbent's contract $\omega_n(\mathbf{s})$ is such that: $x_n^+(\mathbf{s})$ satisfies (11), $x_n^-(\mathbf{s}) = x_{n-1}(\mathbf{s})$, and $p_n(\mathbf{s})$ is given by (14).*
- c. *$\hat{x}(\varphi)$ solves the inactive buyer's problem, (1a)-(1b), and sellers promise utility $x_n(\mathbf{s}) = x_{n-1}^+(\mathbf{s})$.*
- d. *Market tightness $\theta(x; \varphi)$ is consistent with the sorting behavior of inactive buyers, (2).*
- e. *Aggregates and the distribution of agents satisfy the flow equations in Section 3.4.*

We next establish that the equilibrium exists and is unique within the class of equilibria that we have restricted our attention to:

Proposition 2 (Existence) *Under Assumption 2 and some regularity conditions, there exists a unique BRE with non-negative entry in all states of the world, $\varphi \in \Phi$.*

Proof. See Appendix A.4 for an incomplete proof.

Proposition 3 (Efficiency) *When it exists, a BRE with non-negative entry is constrained-efficient.*

Proof. See Appendix A.5 for an incomplete proof.

4 Computation and Numerical Example

In this section, we discuss the computational implementation of the equilibrium and present a numerical example in order to illustrate the type of predictions that the model can generate in terms of pricing and firm dynamics.

4.1 Analytical solutions

For the remainder of the paper, we make use of the following meeting rates: $\eta(\theta) = \theta(1 + \theta^\gamma)^{-\frac{1}{\gamma}}$ and $\mu(\theta) = \eta(\theta)/\theta = (1 + \theta^\gamma)^{-\frac{1}{\gamma}}$. These meeting rates satisfy the regularity conditions stated in Assumption 2.³¹ In particular, to obtain concavity of $\eta \circ \mu^{-1}$ (Assumption 2(iii)) we need $\gamma > 1$, which we assume henceforth.

Equation (10) defines the following $\theta : \mathcal{X} \times \Phi \rightarrow \mathbb{R}_+$ mapping:

$$\theta(x; \varphi) = \left[\left(\frac{\Gamma^B(\varphi)}{x - U^B(\varphi)} \right)^{-\gamma} - 1 \right]^{\frac{1}{\gamma}} \quad (22)$$

Note, once again, that $\theta(x; \varphi)$ is a continuous, increasing and concave function of x . The joint surplus maximization rule (equation (11)) reads:

$$W_{n+1}(\mathbf{s}) - W_n(\mathbf{s}) - x_{n+1}(\mathbf{s}) = \left[\theta_{n+1}(\mathbf{s}) \right]^\gamma (x_{n+1}(\mathbf{s}) - U^B(\varphi)) \quad (23)$$

for any $n \in \mathbb{N}$ and $\mathbf{s} \in \mathcal{Z} \times \Phi$, where $\theta_{n+1}(\mathbf{s}) := \theta(x_{n+1}(\mathbf{s}); \varphi)$. This equation reflects the relevant trade-offs in the equilibrium: when firms offer a higher value, they attract more buyers because their relative outside option worsens (right-hand side). Yet, the remaining value that accrues to the seller is also lower (left-hand side). The joint-surplus maximizing rule splits the rents so that, for the marginal customer, these payoffs are equalized. Finally, $W_n(\mathbf{s})$ is defined in equation (13), and therefore can be found as the solution to a second-order difference equation in n , with boundary conditions $W_0(\mathbf{s}) = 0$ and, via free entry of firms (equation (8)):

$$\forall \varphi \in \Phi : \quad \kappa = \sum_{z_0} \pi_z(z_0) \eta(\theta_1(z_0, \varphi)) \underbrace{(W_1(z_0, \varphi) - x_1(z_0, \varphi))}_{=V_1^S(z_0, \varphi)}$$

The solution of the model can thus be represented by a system of three equations, (13)-(22)-(23), and three unknowns, $\{x_n(\mathbf{s}), W_n(\mathbf{s}), \theta_n(\mathbf{s})\}$, with prices $\{p_n(\mathbf{s})\}$ and distributions determined independently through equation (14) and the flow equations of Section 3.4, respectively. The system of equations can be further collapsed into a single equation for $\Delta_n(\mathbf{s}) := W_n(\mathbf{s}) - W_{n-1}(\mathbf{s})$, which represents the marginal joint surplus of an additional customer. Particularly, we obtain the following result:

Proposition 4 (Solution to the Joint Surplus) *For each $(z, \varphi) \in \mathcal{Z} \times \Phi$, $\{\Delta_n(z, \varphi) : n = 1, 2, \dots\}$ solves the recursion:*

$$\Delta_{n+1}(z, \varphi) - U^B(\varphi) = \left(\Gamma^B(\varphi)^{\frac{\gamma}{1+\gamma}} + \Gamma_n^S(z, \varphi)^{\frac{\gamma}{1+\gamma}} \right)^{\frac{1+\gamma}{\gamma}}$$

³¹ An alternative matching function that would comply with our assumptions is the so-called urn-ball matching function, for which $\eta(\theta) = \theta(1 - e^{-1/\theta})$. Note, however, that the more common Cobb-Douglas matching function $\eta(\theta) = \theta^\gamma$ ($0 < \gamma < 1$) fails requirement (iii).

where:

$$\Gamma_n^S(z, \varphi) := \rho(z, \varphi) \sum_{j=1}^n \Delta_j(z, \varphi) + n \left(\delta_c \Delta_n(z, \varphi) - v - (\delta_f + \delta_c) U^B(\varphi) \right) + \mathcal{C}(n; z, \varphi) - \Xi_n(z, \varphi) \quad (24)$$

for all $n \geq 1$, with $\Gamma_0^S(z, \varphi)$ satisfying $\kappa = \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \Gamma_0^S(z_0, \varphi)$, and where we have used $\rho(z, \varphi) \equiv r + \delta_f + \sum_{z'} \lambda_z(z'|z) + \sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi)$ as the effective discount rate of the firm, and $\Xi_n(z, \varphi) := \sum_{z'} \lambda_z(z'|z) \sum_{j=1}^n \Delta_j(z', \varphi) \sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) \sum_{j=1}^n \Delta_j(z, \varphi')$.

Proof. [to be written]

The result exposed in Proposition 4 has a natural interpretation. First, recall from equation (2) that $\Gamma^B(\varphi)$ is the expected value from matching for a buyer, which is equalized across markets. Likewise, $\Gamma_n^S(z, \varphi)$ can be interpreted as the expected match value for seller (n, z) .³² Therefore, we can establish that, in equilibrium, the marginal net gain in joint surplus between the seller and the buyers for an n -to- $(n+1)$ transition is a CES combination of both of their ex-ante match surpluses. The equilibrium thus reduces to solving a recursion for Δ_n , with initial condition given by free entry.

Algorithm Since the model allows for an analytical representation of the stationary equilibrium, the computational implementation does not require the use of approximation methods of any kind. Moreover, by virtue of block recursivity, we can also compute the distribution of agents exactly, as well as transitional dynamics.

To find the equilibrium sequences described by the system of equations (13)-(22)-(23), we use a shooting algorithm in which we iterate over the value of inactivity $U^B(\varphi)$. The state space grid is $\mathcal{N} \times \mathcal{Z} \times \Phi$, where $\mathcal{N} = \{1, \dots, \bar{n}\}$, $\mathcal{Z} = \{z_i\}_{i=1}^{k_z}$, and $\Phi = \{\varphi_j\}_{j=1}^{k_\varphi}$. Firm size is truncated at a sufficiently large $\bar{n} \in \mathbb{N}$. The following describes the steps of the algorithm:

1. Set the iteration counter at $q = 0$ and make an initial guess, $U^B(\varphi|q = 0) > 0$ (a $k_\varphi \times 1$ vector).
2. For a given iteration $q \in \mathbb{N}$:
 - (a) Solve for the sequences $\{\theta_n(z, \varphi|q), x_n(z, \varphi|q), W_n(z, \varphi|q)\}_{(n,z) \in \mathcal{N} \times \mathcal{Z}}$ by iterating *forward*, for each $\varphi \in \Phi$.
 - (b) Compute the new $U^B(\varphi|q+1)$ that equates κ to the right-hand side of equation (21).
3. Stop if $\|U^B(\varphi|q+1) - U^B(\varphi|q)\| \leq \varepsilon$, for some small tolerance level $\varepsilon > 0$. Otherwise, go back to step 2. with $q \leftarrow q+1$.

After convergence, we have the full sequences $\{\mathcal{W}^*(\varphi), \mathcal{X}^*(\varphi), \Theta^*(\varphi)\}$ defined on every point of the grid, from which the price matrices $\mathcal{P}^*(\varphi)$ and distributions can be readily computed from our results above.

³² To see that $\Gamma^S(z, \varphi)$ is the expected match surplus for the seller, note that if we use definition (24) in equation (13) we obtain that $\Gamma_n^S = \eta(\theta_{n+1}) (W_{n+1} - W_n - x_{n+1})$.

Let us now present results on a numerical example of the economy. For this exercise, we assume constant returns to scale in technology in order to isolate our pricing mechanism giving rise to differential firm growth rates, and thus set $\mathcal{C}(n, \cdot) \propto n$. We normalize $v = 1$ so that the marginal utility from consumption serves as a numeraire. For illustrative purposes, we adopt a 2×2 specification for the exogenous shocks, with $\mathcal{Z} \times \Phi = \{z_L, z_H\} \times \{\varphi_L, \varphi_H\}$, where the subscript H (or L) marks the “boom” (or “bust”) state.

4.2 Stationary Policies

Figure 5 shows the seller’s and each buyer’s value, the pricing policy, and the resulting market tightness, across firm sizes, for each idiosyncratic and aggregate state. We see that the joint surplus is higher for larger matches. When small, the seller promises a high continuation value with a low price in order to attract a larger mass of inactive buyers. As the firm grows, the captive customer obtains decreasing values, and the growth rate of the firm shrinks.

Figure 5: Joint surplus, market tightness, promised utility and firm growth, for each exogenous state.

Figure 6 shows the stationary distribution of sellers by size (relative to the total mass of incumbents), on the left panel, the distribution of inactive buyers on the central panel, and the distribution of customers, on the right panel. The figure shows that the model is able to generate a firm size distribution with a fat right tail, as seen in the data. Because size and age are highly correlated in the model, the distribution of firms across age cohorts qualitatively resembles the data, too (Figure 1). Most firms in the economy have only few customers, and only a few sellers have many. Most active buyers are matched with small

or medium-sized sellers. Inactive buyers, on the other hand, are queuing mostly on markets populated by small sellers, for those markets offer high values and low prices, though low meeting rates due to long queues. Moreover, the bust state has an overall lower mass of active agents. The boom state, in contrast, has on average larger sellers, and exhibits a higher entry rate of firms and thus, on average, younger sellers. Since these firms typically price lower, the aggregate price level in a boom is lower than in a bust, and aggregate output is higher since more agents are matched to sellers.

Figure 6: Ergodic firm size distribution, and distribution of active buyers, under different aggregate states.

These observations in the stationary equilibrium align with certain aspects of the data. Depressed entry rates and the aging of the industry are phenomena that have been recently reported in the data by [Pugsley and Şahin \(2015\)](#) as a persistent symptom of the decline in business dynamism in the U.S. since the early 1980s. In the model, when costs are higher for all levels of the customer base, firms are deterred from entering into the product market and, as a result, sellers have on average more customers. On the other hand, the average markup of the economy is increasing in firm size/age, as reported by [Foster et al. \(2008\)](#) and [Peters \(2016\)](#), and is on average higher in recessions. Some of these moments could serve as targets for a quantitative evaluation of the model.³³

4.3 The Response to Aggregate Shocks

Figure 7 presents the impulse responses to an expansionary demand shock to the cost of search, c (e.g. a drop in the disutility of shopping). The first thing to note is that the response of economic aggregates is slow-moving: even though firm-level policies jump with the aggregate state, the underlying distribution of agents across states is slow-moving (obeying the transitions described in Section 3.4), and the sluggish adjustment is accounted for by the fact that firms take time to reach their stationary size.

The response of the economy to the demand shock reflects both level as well as compositional effects. Due to a boom in demand, a higher fraction of buyers search for sellers, and more matches occur. This causes a boom in both the average size of firms as well as the average firm growth rate, and the aggregate output of the economy goes up. The response on average prices is accounted for by shifts in the distribution of firms. Given a higher demand for products, outside firms find it easier to start their

³³ This is currently work in progress.

operations, and the increase in the entry of small and low-price firms increases competition and accounts for the countercyclical response of the average price level.

Figure 7: Impulse response functions to an expansionary aggregate demand shock, in deviations from steady-state. All diagrams plot cross-sectional averages using the theoretical distribution of firms over states (n, z) .

Thus, qualitatively, the model is able to generate countercyclicity in price markups. Not only this is in line with the empirical evidence stating that markups respond negatively to the business cycle (Rotemberg and Woodford (1999)), but countercyclical markups are also an important ingredient in sticky price models to get the correct comovement of economic aggregates in response to demand-sided shocks (Galí *et al.* (2007), Justiniano *et al.* (2010)). Here, we show that our model generates these observations endogenously, without having to rely on exogenous sources of price stickiness such as infrequent pricing or menu costs. The model therefore offers potential to explain the cyclical behavior of economic aggregates in markets in which firms have customer retention and attraction concerns.

To further unpack the average price response, in Figure 8 we study the impulse responses of a decomposition based on the different price components (left panel), and a cross-sectional decomposition in firm sizes (right panel). The left panel shows the responses of each one of the different components from equation (14). We note that, in response to an aggregate demand shock, firms cut prices especially through the separation penalties. Intuitively, when new matches are formed, firms have to penalize customers less for the risk of separation in order for incentive-compatibility to be satisfied, as matches are now less constrained by destruction risk. On the right panel, we plot the response of the contribution of

each type of firm (by size) to the aggregate price level.³⁴ In this plot, we see that the price response is primarily driven by the smallest firms, whose contribution to the average price level booms in response to the demand shock. In contrast, the contribution by larger, higher-price firms decreases. This therefore reflects that the countercyclical price response on average is primarily driven by a powerful compositional shift in the firm size distribution, whereby the entry of new firms with low prices amplifies the response of the economy to aggregate shocks.

Figure 8: Impulse response functions to an expansionary aggregate demand shock, in deviations from steady-state. Decomposition of the price response into its different components from equation (14) [left panel], and the cross-sectional contribution to the average price by each firm type [right panel].

For completeness, in Figure 9 we plot the response of the economy to an aggregate expansionary supply shock (e.g. a reduction in the firm’s costs of production for all size levels). Unlike before, because costs are proportional to size, the shock has a stronger impact on the larger, higher-price firms, and these sellers can now increase their prices on average and cut down on their promises. While output still booms do to the increase in productive efficiency, less buyers now become matched, so the average growth rate of firms decreases. Depending on the parametrization, however, the entry margin could once again play an important role, and put competitive downward pressure on prices. A proper quantitative exercise would then establish the exact effects that customer market considerations may have in the responses to aggregate shocks of a calibrated economy.

Simulation Results Finally, for the purpose of illustration, we present results from a simulation exercise. The simulation allows us to compute size as firm revenue while disentangling firm output from prices, an important distinction to correctly understand the sources of firm heterogeneity in the data as emphasized in the Introduction and in the work of Foster *et al.* (2008, 2016).³⁵

³⁴ That is, where the average price level is computed by $P_t = \sum_n p_n g_{n,t}$ for stationary price policies $\{p_n\}$ and equilibrium shares $\{g_{n,t}\}$ (with $g_{n,t} = S_{n,t}/S_t$), the figure plots deviations of the variable $\hat{p}_{n,t} = p_n g_{n,t}/P_t$ from steady-state.

³⁵ We simulate the model for $K = 5,000$ firms during $T = 500$ periods of length $\Delta = 0.01$. We assume that all firms start with a size drawn from the ergodic size distribution at $t = 0$, although we have tried other initial distributions and in all cases the distribution of sizes converges to the stationary S_n .

Figure 9: Impulse response functions to an expansionary aggregate supply shock, in deviations from steady-state. All diagrams plot cross-sectional averages using the theoretical distribution of firms over states (n, z) .

Figures 10-11 show the results. We see that the simulation generates a negative correlation between prices and sizes, and in particular between sizes and size growth (a negative deviation from Gibrat’s law). Moreover, the sales distribution (lower-left panel of Figure 10) resembles the theoretical one, seen in the left-most panel of Figure 6. The reason is that there is a high correlation between sales (our empirical measure of size) and number of customers (our theoretical measure of size), as seen in the lower-right panel of Figure 10. Finally, Figure 11 depicts histograms for price changes, non-zero price changes and sales growth, and shows that the simulation obtains substantial dispersion in all of these dimensions.

Our next step is to quantitatively match these effects by calibrating the model to firm-level data, and study the implications for pricing and firm dynamics of shocks to firm productivity or firm demand.

5 Final Remarks and Next Steps

We have presented a dynamic search model of customer acquisition through firm pricing with aggregate and idiosyncratic shocks and a relevant scope for firm dynamics. In the model, firms of different customer base sizes use a menu of prices and continuation promises differently in order to trade off two conflicting concerns: attracting new customers to increase future profits, and extracting surplus from incumbent customers to increase current profits. We have shown that smaller sellers are more concerned with the former, so they choose to make higher utility promises in order to grow more rapidly. The model therefore exhibits cross-sectional price dispersion on the basis of cross-sectional size differences, and offers

Figure 10: Simulation results. Sales and price correlations, sales distribution, and correlation between number of customers and normalized sales (i.e. sales divided by their mean).

Figure 11: Simulation results. Distribution of price changes, non-zero price changes, and sales growth.

a micro-founded interpretation for sluggish price dynamics at both the firm and the aggregate level.

We have analyzed a number of predictions on both pricing and firm dynamics dimensions, and found them to be qualitatively in line with the data. For example, we find asymmetries in the rate and frequency of price adjustments, observations prevalent in the data that the literature tends to address with some form of exogenous menu cost (e.g. [Nakamura and Steinsson \(2010\)](#)). Moreover, markups vary with firm size and age in the way that has been documented in the micro data, and seem to be countercyclical and thus a potentially strong source of macroeconomic transmission. Future versions of our work will provide a quantitative analysis of how aggregate supply and demand shocks can account for salient features of the U.S. economy, especially regarding the secular decline in business dynamism experienced since the early 1980s ([Pugsley and Şahin \(2015\)](#)). In addition, our next step is to quantify the theory by using a calibration to discipline parameters, and use the parametrized model to conduct policy analysis. One interesting avenue would be, for example, to analyze the implications for monetary policy in customer markets.

References

- APEL, M., FRIBERG, R. and HALLSTEN, K. (2005). Micro foundations of macroeconomic price adjustment: Survey evidence from Swedish firms. *Journal of Money, Credit and Banking*, **37** (2), 313 – 338.
- ARGENTE, D. and YEH, C. (2016). Price-setting under uncertain demand: A menu cost model with price experimentation. *Working Paper, University of Chicago*.
- ARKOLAKIS, C. (2010). Market penetration costs and the new consumer margin in international trade. *Journal of Political Economy*, **118** (6), 1151 – 1199.
- (2016). A unified theory of firm selection and growth. *The Quarterly Journal of Economics*, **131** (1), 89–155.
- ATKESON, A. and KEHOE, P. J. (2005). Modeling and measuring organization capital. *Journal of Political Economy*, **113** (5), 1026 – 1053.
- AUDRETSCH, D. B., KLUMP, L., SANTARELLI, E. and THURIK, A. R. (2004). Gibrat’s law: Are the services different? *Review of Industrial Organization*, **24**, 301 – 324.
- BAI, Y., RIOS-RULL, V. and STORESLETTEN, K. (2012). Demand shocks as productivity shocks. *Working Paper, University of Pennsylvania*.
- BILS, M. (1989). Pricing in a customer market. *The Quarterly Journal of Economics*, **104** (4), 699 – 718.
- BLINDER, A. S., CANETTI, E. R. D., LEBOW, D. E. and RUDD, J. B. (1998). *Asking About Prices: A New Approach to Understanding Price Stickiness*. Russell Sage Foundation.
- 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.
- BUTTERS, G. R. (1977). Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, **44** (3), 465 – 491.
- 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. (2014). Firm-level shocks and labor adjustments. *IFAU Working Paper*.
- CASON, T. N. and FRIEDMAN, D. (2003). Buyer search and price dispersion: A laboratory study. *Journal of Economic Theory*, **112** (2), 232 – 260.

- CAVENAILE, L. and ROLDAN, P. (2016). Advertising, innovation, and economic growth. *Working Paper, New York University*.
- CAVES, R. E. (1998). Industrial organization and new findings on the turnover and mobility of firms. *Journal of Economic Literature*, **36** (4), 1947 – 1982.
- DELOECKER, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, **79** (5), 1407 – 1451.
- DIAMOND, P. A. (1971). A model of price adjustment. *Journal of Economic Theory*, **3**, 156 – 168.
- DIXIT, A. K. and STIGLITZ, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, **67** (3), 297 – 308.
- DROZD, L. A. and NOSAL, J. B. (2012). Understanding international prices: Customers as capital. *American Economic Review*, **102** (1), 364 – 395.
- 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.
- EATON, J., ESLAVA, M., JINKINS, D., KRIZAN, C. J. and TYBOUT, J. (2014). A search and learning model of export dynamics. *Working Paper, Pennsylvania State University*.
- 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)*.
- FISHMAN, A. and ROB, R. (2003). Consumer inertia, firm growth and industry dynamics. *Journal of Economic Theory*, **109**, 24 – 38.
- 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.
- GALENIANOS, M. and KIRCHER, P. (2009). Directed search with multiple job applications. *Journal of Economic Theory*, **144** (2), 445 – 471.
- and — (2012). On the game-theoretic foundations of competitive search equilibrium. *International Economic Review*, **53** (1), 1 – 21.

- GALÍ, J., GERTLER, M. and LÓPEZ-SALIDO, D. (2007). Markups, gaps and the welfare costs of business fluctuations. *Review of Economics and Statistics*, **89** (1), 44 – 59.
- GEROSKI, P. A. (1998). An applied econometrician’s view of large company performance. *Review of Industrial Organization*, **13** (3), 271 – 294.
- GIBRAT, R. (1931). *Les Inégalités Économiques*. Paris, Librairie du Recueil Sirey.
- GILCHRIST, S., SCHOENLE, R., SIM, J. and ZAKRAJŠEK, E. (2016). Inflation dynamics during the financial crisis. *Working Paper, Boston University*.
- GOURIO, F. and RUDANKO, L. (2014a). Can intangible capital explain cyclical movements in the labor wedge? *American Economic Review*, **104** (5), 183–88.
- and — (2014b). Customer capital. *Review of Economic Studies*, **81** (3), 1102 – 1136.
- HALL, B. H. (1987). The relationship between firm size and firm growth in the U.S. manufacturing sector. *Journal of Industrial Economics*, **35** (4), 583 – 606.
- HALL, R. E. (2008). General equilibrium with customer relationships: A dynamic analysis of rent-seeking. *Working Paper, Stanford University*.
- HALL, S., WALSH, M. and YATES, A. (1997). How do UK companies set prices? *Bank of England Quarterly Bulletin*, pp. 180 – 192.
- HALTIWANGER, J., JARMIN, R. S. and MIRANDA, J. (2013). Who creates jobs? small versus large versus young. *The Review of Economics and Statistics*, **95** (2), 347 – 361.
- HOPENHAYN, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, **60** (5), 1127 – 1150.
- HOTTMAN, C., 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. (2014). The life cycle of plants in India and Mexico. *The Quarterly Journal of Economics*, **129** (3), 1035 – 1084.
- JOVANOVIĆ, B. (1982). Selection and the evolution of the industry. *Econometrica*, **50** (3), 649 – 670.
- JUSTINIANO, A., PRIMICERI, G. E. and TAMBALOTTI, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, **57** (2), 132 – 145.
- KAAS, L. and KIMASA, B. (2016). Firm dynamics with frictional product and labor markets. *Working Paper, University of Konstanz*.

- 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.
- , —, RUDANKO, L. and TRACHTER, N. (2016). Relative price dispersion: Evidence and theory. *Working Paper, University of Chicago*.
- 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.
- KLETTE, T. J. and KORTUM, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, **112** (5), 986 – 1018.
- KRUSELL, P. and SMITH, A. A. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, **106** (5), 867 – 896.
- 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*.
- (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, **122** (3), 1103 – 1144.
- MCGRATTAN, E. R. (2015). Intangible capital and measured productivity. *Working Paper, University of Minnesota*.
- and PRESCOTT, E. C. (2014). A reassessment of real business cycle theory. *American Economic Review*, **104** (5), 177–82.
- MENZIO, G. (2007). A search theory of rigid prices. *Working Paper, University of Pennsylvania*.
- 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. (2015). Equilibrium price dispersion across and within stores. *Working Paper, University of Pennsylvania*.
- 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. (2010). Monetary non-neutrality in a multisector menu cost model. *The Quarterly Journal of Economics*, **125** (3), 961 – 1013.
- and — (2011). Price setting in forward-looking customer markets. *Journal of Monetary Economics*, **58** (3), 220 – 233.
- and — (2013). Price rigidity: Microeconomic evidence and macroeconomic implications. *Annual Review of Economics*, **5**, 133 – 163.
- PACIELLO, L., POZZI, A. and TRACHTER, N. (2016). Price dynamics with customer markets. *Working Paper, Federal Reserve Bank of Richmond*.
- PERLA, J. (2016). Product awareness, industry life cycles, and aggregate profits. *Working Paper, University of British Columbia*.
- PETERS, M. (2016). 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.
- PUGSLEY, B. and ŞAHİN, A. (2015). Grown-up business cycles. *Federal Reserve Bank of New York Staff Reports* 707.
- 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.
- ROSSI-HANSBERG, E. and WRIGHT, M. L. J. (2007). Establishment size dynamics in the aggregate economy. *American Economic Review*, **97** (5), 1639 – 1664.

- 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 No. 6909*.
- RUDANKO, L. (2017). Pricing in a frictional product market. *Working Paper, Federal Reserve Bank of Philadelphia*.
- SCHAAL, E. (2017). Uncertainty and unemployment. *Econometrica*, **Forthcoming**.
- SHI, S. (2009). Directed search for equilibrium wage-tenure contracts. *Econometrica*, **77** (2), 561 – 584.
- (2016). Customer relationship and sales. *Journal of Economic Theory*, **166**, 483–516.
- SORENSEN, A. T. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy*, **108** (4), 833 – 850.
- SUTTON, J. (1997). Gibrat’s legacy. *Journal of Economic Literature*, **35** (1), 40 – 59.
- VARIAN, H. R. (1980). A model of sales. *The American Economic Review*, **70** (4), 651 – 659.

A Derivations and Proofs

A.1 Derivation of the HJB Equations

In this appendix, we show how to derive the HJB equations for inactive buyers (equation (1b)), active buyers (equation (4)), incumbent sellers (equation (5)), and potential entrant sellers (equation (7)), from their discrete-time counterparts. Throughout, we assume a discrete-time stationary environment in which time intervals are equidistant and of some (short) length $\Delta > 0$.

Inactive Buyers The value of an inactive buyers in state $\varphi \in \Phi$ is:

$$U^B(\varphi) = \max_{\hat{x}(\varphi) \in \mathcal{X}} u^B(\hat{x}(\varphi); \varphi)$$

where $u^B(\hat{x}(\varphi); \varphi)$ is given by the Bellman equation:

$$\begin{aligned} u^B(\hat{x}(\varphi); \varphi) = & -c(\varphi)\Delta + e^{-r\Delta} \left(\left[\eta(\theta(\hat{x}(\varphi); \varphi))\Delta + o(\Delta) \right] \max \left\{ \hat{x}(\varphi), u^B(\hat{x}(\varphi); \varphi) \right\} \right. \\ & + \sum_{\varphi' \in \Phi} \left[\lambda_{\varphi}(\varphi' | \varphi)\Delta + o(\Delta) \right] u^B(\hat{x}(\varphi); \varphi') \\ & \left. + \left[1 - \eta(\theta(\hat{x}(\varphi); \varphi))\Delta - \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi' | \varphi)\Delta + o(\Delta) \right] u^B(\hat{x}(\varphi); \varphi) \right) + o(\Delta) \end{aligned} \quad (\text{A.1})$$

where $o(\Delta)$ has the property $\lim_{\Delta \rightarrow 0} \frac{o(\Delta)}{\Delta} = 0$. Here, and in every value function to follow, we use that, for a given Poisson arrival rate $k \geq 0$, the term $k\Delta + o(\Delta)$ (respectively, $1 - k\Delta + o(\Delta)$) approximates the probability of exactly one arrival (respectively, no arrivals) of the Poisson shock within some time interval $[t, t + \Delta]$. This approximation is valid for $\Delta > 0$ small enough. The probability of two or more arrivals is approximately equal to $o(\Delta)$, and accounted for in the very last term of equation (A.1). Moreover, we have used a discount factor of $e^{-r\Delta}$, which approximates the usual discrete-time discounting of $\frac{1}{1+r\Delta}$ when $\Delta > 0$ is small enough.

The right-hand side of equation (A.1) has a straightforward interpretation: when the buyer is inactive, she pays the cost $c(\varphi)$ to draw a market $\hat{x}(\varphi)$. In this market, there is a probability $\left[\eta(\theta(\hat{x}(\varphi); \varphi))\Delta + o(\Delta) \right]$ of matching with exactly one firm in period $t + \Delta$, in which case the buyer would choose between (i) joining and receiving the promised value $\hat{x}(\varphi)$, or (ii) remaining inactive and obtaining the continuation value $u^B(\hat{x}(\varphi); \varphi)$.³⁶ To obtain equation (1b), we subtract $e^{-r\Delta}u^B(\hat{x}(\varphi); \varphi)$ from both sides, divide every term by Δ , and take the continuous-time limit as $\Delta \rightarrow 0$ (using that $\lim_{\Delta \rightarrow 0} \frac{1-e^{-r\Delta}}{\Delta} = r$).

Active Buyers Now consider the value at time t of being the customer of a firm of size n that promises value x in state $\mathbf{s} = (z, \varphi) \in \mathcal{Z} \times \Phi$, and offers contract $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$. In the discrete-time approximation, we have:

³⁶ We are implicitly assuming that the buyers searches in the same market in period $t + \Delta$ if unsuccessful in period t . This is without loss of generality because, in equilibrium, the buyer is indifferent among all active markets and across all aggregate states of nature.

$$\begin{aligned}
V^B(n, \omega; \mathbf{s}) = & (v - p)\Delta + e^{-r\Delta} \left([(\delta_f + \delta_c)\Delta + o(\Delta)] U^B(\varphi) \right. \\
& + [(n-1)\delta_c\Delta + o(\Delta)] \max \{x'(n-1; \mathbf{s}), U^B(\varphi)\} \\
& + [\eta(\theta(x'(n+1; \mathbf{s}); \varphi))\Delta + o(\Delta)] \max \{x'(n+1; \mathbf{s}), U^B(\varphi)\} \\
& + \sum_{z' \in \mathcal{Z}} [\lambda_z(z'|z)\Delta + o(\Delta)] \max \{x'(n; z', \varphi), U^B(\varphi)\} + \sum_{\varphi' \in \Phi} [\lambda_{\varphi}(\varphi'|\varphi)\Delta + o(\Delta)] \max \{x'(n; z, \varphi'), U^B(\varphi')\} \\
& + \left[1 - \delta_f\Delta - n\delta_c\Delta - \eta(\theta(x'(n+1; \mathbf{s}); \varphi))\Delta - \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z)\Delta - \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi'|\varphi)\Delta + o(\Delta) \right] \times \\
& \left. \times \max \{V^B(n, \omega; \mathbf{s}), U^B(\varphi)\} \right) + o(\Delta)
\end{aligned} \tag{A.2}$$

Equation (A.2) is the buyer's Bellman equation in full generality. Particularly, we recognize the possibility that, as indicated by the max operators, the customer is free to opt out of the match at the beginning of period $t + \Delta$, in which case she stops buying from the firm and must become inactive for at least one period. The buyer may also transition to inactivity if either the firm dies or the match is destroyed exogenously. If she remains within the firm, she acknowledges that the match might change in value once the firm shrinks (second line), grows (third line), or is hit by an exogenous shock (fourth line). The fifth line is the complementary case in which none of the Poisson shocks hit, in which case the customer is still given the chance to transition to inactivity.

To obtain our final HJB equation, we first recall that transiting to inactivity is not optimal under any contingency. Namely, $U^B(\varphi) \leq \underline{x}$, $\forall \varphi \in \Phi$. Since, in equilibrium, the seller offers continuation values in the set $\mathcal{X}^*(\varphi)$ of equilibrium markets, we can eliminate the max operators above. In this case, we then proceed by subtracting $e^{-r\Delta}V^B(n, \omega; \mathbf{s})$ from both sides, dividing every term by Δ , and finally taking the continuous-time limit as $\Delta \rightarrow 0$ to obtain equation (4).

Incumbent Sellers Consider now a seller with $n \in \mathbb{N}$ customers who is currently promising a value x in state $\mathbf{s} = (z, \varphi)$. The seller must choose contract $\omega = \{p, \mathbf{x}'(n'; \mathbf{s}')\}$ to maximize:

$$\begin{aligned}
V^S(n, x; \mathbf{s}) = \max_{\omega \in \Omega} \left\{ (np - \mathcal{C}(n; \mathbf{s}))\Delta + e^{-r\Delta} \left([\delta_f\Delta + o(\Delta)] V_0^S(\varphi) \right. \right. \\
& + [n\delta_c\Delta + o(\Delta)] \max \{V^S(n-1, x'(n-1; \mathbf{s}); \mathbf{s}), V_0^S(\varphi)\} \\
& + [\eta(\theta(x'(n+1; \mathbf{s}); \varphi))\Delta + o(\Delta)] \max \{V^S(n+1, x'(n+1; \mathbf{s}); \mathbf{s}), V_0^S(\varphi)\} \\
& + \sum_{z' \in \mathcal{Z}} [\lambda_z(z'|z)\Delta + o(\Delta)] \max \{V^S(n, x'(n; z', \varphi); z', \varphi), V_0^S(\varphi)\} \\
& + \sum_{\varphi' \in \Phi} [\lambda_{\varphi}(\varphi'|\varphi)\Delta + o(\Delta)] \max \{V^S(n, x'(n; z, \varphi'); z, \varphi'), V_0^S(\varphi')\} \\
& + \left[1 - \delta_f\Delta - n\delta_c\Delta - \eta(\theta(x'(n+1; \mathbf{s}); \varphi))\Delta - \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z)\Delta - \sum_{\varphi' \in \Phi} \lambda_{\varphi}(\varphi'|\varphi)\Delta + o(\Delta) \right] \times \\
& \left. \left. \times \max \{V^S(n, x; \mathbf{s}), V_0^S(\varphi)\} \right) \right\} + o(\Delta)
\end{aligned}$$

subject to $V^B(n, \omega; \mathbf{s}) \geq x$. As in the case of buyers above, we write the seller's Bellman equation in full generality. In particular, at the beginning of period $t + \Delta$, the seller can choose to cease operations and transition to $n = 0$, as indicated by the max operators. Otherwise, the interpretation of the above equation is very similar to the one given in the main text. To arrive at equation (5), we anticipate that $V_0^S(\varphi) = 0, \forall t$, by free entry, and therefore $V^S(n, x; \mathbf{s}) \geq V_0^S(\varphi) = 0, \forall (n, x; \mathbf{s})$, without loss of generality. Then, we can proceed by taking the usual continuous-time limit, as described above.

Entrant Sellers Finally, we characterize the problem of potential entrants. In the discrete-time approximation, potential entrants draw their initial productivity z_0 from the distribution π_z , after which they pay the set-up cost $\kappa > 0$ allowing them to post a contract. Let $v_0^S(\mathbf{s}_0)$ denote the value of drawing productivity z_0 , where $\mathbf{s}_0 := (z_0, \varphi)$. Then, the entrant chooses the initial contract $\omega_0 := \mathbf{x}'(1, \mathbf{s}_0) \in \mathcal{X}$, specifying the promised utility for the first customer of the firm, to maximize:

$$\begin{aligned} v_0^S(\mathbf{s}_0) = \max_{\omega_0} & \left\{ e^{-r\Delta} \left(\left[\eta \left(\theta(\mathbf{x}'(1; \mathbf{s}_0); \varphi) \right) \Delta + o(\Delta) \right] \max \left\{ V^S(1, x'(1; \mathbf{s}_0); \mathbf{s}_0), V_0^S(\varphi) \right\} \right. \right. \\ & + \sum_{\varphi' \in \Phi} \left[\lambda_\varphi(\varphi' | \varphi) \Delta + o(\Delta) \right] V_0^S(\varphi') \\ & \left. \left. + \left[1 - \eta \left(\theta(x'(1; \mathbf{s}); \varphi) \right) \Delta - \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi' | \varphi) \Delta + o(\Delta) \right] V_0^S(\varphi) \right) \right\} + o(\Delta) \end{aligned} \quad (\text{A.3})$$

where $V_0^S(\varphi)$ is the expected value of entry, defined by:

$$V_0^S(\varphi) = -\kappa\Delta + \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) v_0^S(z_0, \varphi) \quad (\text{A.4})$$

To obtain equation (7), subtract $e^{-r\Delta} V_0^S(\varphi)$ from both sides of equation (A.4) making use of the fact that $\sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) = 1$, and take the continuous-time limit in the usual fashion, noting that, by free entry, we can drop the possibility of voluntary separation, and thereby eliminate the max operator from equation (A.3).

A.2 Proof of Proposition 1: Joint Surplus Problem

Proof. Denote by $\bar{\omega} = \{p, \mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\}$ a generic policy of the typical seller, where $\mathbf{x}(\mathbf{s}') = \{x(z', \varphi), x(z, \varphi')\}$ stands for the promised utility under size n and a new state $\mathbf{s}' \neq \mathbf{s}$ (recall that $x(\mathbf{s}) = x$ by stationarity), whereas $x^+(\mathbf{s})$ and $x^-(\mathbf{s})$ are the upsizing and downsizing choices, respectively. Then, the value of the seller in equilibrium $V^S(n, x; \mathbf{s})$, can be written as the maximand on the right-hand side of (5), evaluated at $\bar{\omega}$. This means that $V^S(n, x; \mathbf{s}) = \max_{\bar{\omega}} \tilde{V}^S(n, x; \mathbf{s} | \bar{\omega})$, where $\tilde{V}^S(n, x; \mathbf{s} | \bar{\omega})$ is given by:³⁷

³⁷ Here we are arguing by free-entry that $V_0^S(\varphi) = 0, \forall \varphi \in \Phi$, to simplify the expression for $\tilde{V}^S$.

$$\begin{aligned} \tilde{V}^S(n, x; \mathbf{s}|\bar{\omega}) = & \frac{1}{\rho_n(\mathbf{s})} \left[\underbrace{pn - \mathcal{C}(n; \mathbf{s})}_{\text{Flow profit}} + \underbrace{\eta\left(\theta(x^+(\mathbf{s}); \varphi)\right) V^S(n+1, x^+(\mathbf{s}); \mathbf{s})}_{\text{Seller grows}} + \underbrace{n\delta_c V^S(n-1, x^-(\mathbf{s}); \mathbf{s})}_{\text{Seller shrinks}} \right. \\ & \left. + \underbrace{\sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) V^S(n, x(z', \varphi); z', \varphi)}_{\text{Idiosyncratic shock}} + \underbrace{\sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) V^S(n, x(z, \varphi'); z, \varphi')}_{\text{Aggregate shock}} \right] \end{aligned} \quad (\text{A.5})$$

and we have defined:

$$\rho_n(\mathbf{s}) := r + \delta_f + n\delta_c + \eta\left(\theta(x^+(\mathbf{s}); \varphi)\right) + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi)$$

as the *effective* discount rate of the firm. The effective discount rate adjusts the actual discount rate (r) for the death, separation, growth, and all state-transition rates faced by the agent.

On the other hand, since the PK constraint is binding, we can solve for p in the buyer's value (equation (4)) to obtain:

$$\begin{aligned} p^{IC}(\mathbf{x}'(n'; \mathbf{s}')) = & v - \rho_n(\mathbf{s})x + \delta_f U^B(\varphi) + \eta\left(\theta(x'(n+1; \mathbf{s}); \varphi)\right) x'(n+1; \mathbf{s}) \\ & + \delta_c \left(U^B(\varphi) + (n-1)x'(n-1; \mathbf{s}) \right) + \sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) x(z', \varphi) + \sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) x(z, \varphi') \end{aligned} \quad (\text{A.6})$$

Equation (A.6) is nothing but the value of the buyer expressed in terms of the price level, and it provides the incentive-compatible price, i.e. the price that is consistent with continuation utilities $\mathbf{x}'(n'; \mathbf{s}')$. Intuitively, other things equal, the price level is higher when promised future utilities are higher and when the discounted value of the buyer is lower, since extracting surplus from the buyer today must be compensated in the future because of incentive-compatibility. Notice that the price appears only linearly in the firm's problem, and it has no bearing on the agents' incentives within the search market. This shows that the firm offers the highest possible price such that the PK constraint is still honored.

Using the above notation, we can now plug in the incentive-compatible price (A.6) into the seller's value (A.5) and, by rearranging terms, we obtain:

$$\begin{aligned} \tilde{W}(n, x; \mathbf{s}|\bar{\omega}) = & \frac{1}{\rho_n(\mathbf{s})} \left[\underbrace{n \left(v + (\delta_f + \delta_c) U^B(\varphi) \right)}_{\text{Customers' total surplus}} - \underbrace{\left(\mathcal{C}(n; \mathbf{s}) + \eta\left(\theta(x^+(\mathbf{s}); \varphi)\right) x^+(\mathbf{s}) \right)}_{\text{Seller's total costs}} \right. \\ & + \underbrace{\eta\left(\theta(x^+(\mathbf{s}); \varphi)\right) W(n+1, x^+(\mathbf{s}); \mathbf{s})}_{\text{Seller grows}} + \underbrace{n\delta_c W(n-1, x^-(\mathbf{s}); \mathbf{s})}_{\text{Seller shrinks}} \\ & \left. + \underbrace{\sum_{z' \in \mathcal{Z}} \lambda_z(z'|z) W(n, x(z', \varphi); z', \varphi)}_{\text{Idiosyncratic shock}} + \underbrace{\sum_{\varphi' \in \Phi} \lambda_\varphi(\varphi'|\varphi) W(n, x(z, \varphi'); z, \varphi')}_{\text{Aggregate shock}} \right] \end{aligned} \quad (\text{A.7})$$

where we have defined:

$$\widetilde{W}(n, x; \mathbf{s}|\bar{\omega}) := \widetilde{V}^S(n, x; \mathbf{s}|\bar{\omega}) + nx$$

as the joint surplus under contract $\bar{\omega}$. The object $\widetilde{W}$ on the right-hand side is the *maximized* joint surplus as defined in the main text, satisfying $W(n, x; \mathbf{s}) := \max_{\bar{\omega}} \widetilde{W}(n, x; \mathbf{s}|\bar{\omega})$. Note additionally that the right-hand side of (A.7) does not depend on x , which proves Part 2 of the proposition.

Part 1 is now almost clear. Since the joint surplus is invariant to the price level by construction, the optimal contract ω^* can be found by splitting the program into two separate stages. In the first stage, the seller chooses the continuation values $\{\mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\}$ that maximize (A.7). In the second stage, once the surplus has been maximized, the seller chooses the incentive-compatible price level (via equation (A.6)) that is consistent with the promised continuation utilities. Formally:

$$\omega^* = \left\{ p, \{\mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\} \right\} \quad s.t.$$

$$\{\mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\} = \arg \max \widetilde{W}(n, x; \mathbf{s}|\bar{\omega}) \quad (\text{A.8a})$$

$$p = p^{IC}(\{\mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\}) \quad (\text{A.8b})$$

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, ω^* , must maximize the joint surplus. Conversely, for any set $\{\mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\}$ of continuation values that maximizes the joint surplus, there is a price level, given by $p = p^{IC}(\{\mathbf{x}(\mathbf{s}'), x^+(\mathbf{s}), x^-(\mathbf{s})\})$, that maximizes the seller's value subject to the PK constraint. Therefore, the seller's problem (equation (5)) and the joint surplus problem (equations (A.8a)-(A.8b)) are *equivalent*. $\square$

A.3 Proofs of Section 3.4

A.3.1 Derivation of KF equations and Aggregate Measures of Agents

In this section, we work with the equilibrium *shares* of agent types, defined by $g_{n,t}(z) := \frac{S_{n,t}(z)}{\mathcal{S}_t}$, for each $n \in \{0, 1, 2, \dots, \bar{n}\}$, for some large $\bar{n} \in \mathbb{N}$ and any z , where $\mathcal{S}_t := \sum_{n \geq 1} \sum_z S_{n,t}(z)$ is the total measure of incumbents. Since we have a discrete support, g is a probability mass function (p.m.f.). Working with shares as opposed to measures thus allows us to use the following restrictions to check for convergence: $g_{n,t}(z) \leq 1$ and $\sum_{n \geq 1} \sum_z g_{n,t}(z) = 1$, $\forall t \geq 0$, for each given aggregate state φ .³⁸

After a period of size $\Delta > 0$, the share of firms of type (n, z) when $n = 1$ becomes:

$$\begin{aligned} g_{1,t+\Delta}(z) &= \left[\pi_z(z) \eta(\theta_{1,t+\Delta}(z, \varphi)) \Delta + o(\Delta) \right] \frac{S_{0,t}(\varphi)}{\mathcal{S}_t} + 2 \left[\delta_c \Delta + o(\Delta) \right] g_{2,t}(z) + \sum_{\tilde{z} \neq z} \left[\lambda_z(z|\tilde{z}) \Delta + o(\Delta) \right] g_{1,t}(\tilde{z}) \\ &\quad + \left[1 - \delta_f \Delta - \delta_c \Delta - \eta(\theta_{2,t+\Delta}(z, \varphi)) \Delta - \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) \Delta + o(\Delta) \right] g_{1,t}(z) \end{aligned} \quad (\text{A.9})$$

Similarly, for $n \geq 2$, we have:

³⁸In the main text, instead, we show KF equations for *measures* instead of shares. The derivations are however nearly identical.

$$\begin{aligned}
g_{n,t+\Delta}(z) &= \left[\eta(\theta_{n,t+\Delta}(z, \varphi)) \Delta + o(\Delta) \right] g_{n-1,t}(z) + (n+1) \left[\delta_c \Delta + o(\Delta) \right] g_{n+1,t}(z) + \sum_{\tilde{z} \neq z} \left[\lambda_z(z|\tilde{z}) \Delta + o(\Delta) \right] g_{n,t}(\tilde{z}) \\
&+ \left[1 - \delta_f \Delta - n \delta_c \Delta - \eta(\theta_{n+1,t+\Delta}(z, \varphi)) \Delta - \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) \Delta + o(\Delta) \right] g_{n,t}(z)
\end{aligned} \tag{A.10}$$

Once again, for any Poisson rate k , we use the fact that $k\Delta + o(\Delta)$ approximates the probability of one arrival, $1 - k\Delta + o(\Delta)$ approximates the probability of no arrivals, and the probability of two or more arrivals is the higher-order term $o(\Delta)$, where $\lim_{\Delta \rightarrow 0} \frac{o(\Delta)}{\Delta} = 0$. These equations describe the law of motion for $g_{n,t}(\cdot)$: due to customer acquisition, attrition, or an exogenous shock, the first line shows the shares of firms transitioning into state (n, z) ; the second line shows the share of firms of type (n, z) that are not hit by any shock, and thereby remain type (n, z) .

Subtracting $g_{n,t}(z)$ from both sides of equation (A.10) and dividing by Δ gives:

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

Taking the limit as $\Delta \rightarrow 0$,

$$\begin{aligned}
\partial_t g_{n,t}(z) &= \eta(\theta_{n,t}(z, \varphi)) g_{n-1,t}(z) + (n+1) \delta_c g_{n+1,t}(z) + \sum_{\tilde{z} \neq z} \lambda_z(z|\tilde{z}) g_{n,t}(\tilde{z}) \\
&- \left(\delta_f + n \delta_c + \eta(\theta_{n+1,t}(z, \varphi)) + \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) \right) g_{n,t}(z)
\end{aligned}$$

A similar derivation on (A.9) shows that, for $n = 1$,

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

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:

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

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

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

In the stationary solution, inflows and outflows are equated for every pair of idiosyncratic states (n, z) , so we impose $\partial_t g_{n,t}(z) = 0$ and $\partial_t S_{0,t}(\varphi) = 0$ on the above equations. The system of KF equations then becomes a system of second-order equations which can be easily solved numerically on the state-space grid. In particular, we find a solution for the matrix $\{g_n(z)\}_{n,z}$, and the share of potential entrants per incumbent firm, $h_0 := S_0/\mathcal{S}$.

Once we have found a solution, we can compute aggregate measures as follows. Firstly, using (16) we obtain the object $b_n^A(z) := \frac{B_n^A(z)}{\mathcal{S}}$ by:

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

for any $n = 1, 2, \dots$, and $b^A := B^A/\mathcal{S} = \sum_{n \geq 1} \sum_z n g_n(z)$. On the other hand, from equation (15) we know that $B_n^I(z) = \mathcal{S} \theta_n(z) g_{n-1}(z)$. Therefore, adding across states (from $n = 2$ onward) yields:

$$\begin{aligned} \mathcal{S} \sum_{n \geq 2} \sum_z \theta_n(z) g_{n-1}(z) &= \sum_{n \geq 2} \sum_z B_n^I(z) \\ &= B^I - \sum_z B_1^I(z) \\ &= 1 - B^A - \sum_z \theta_1(z) S_0 \end{aligned}$$

Using the above definitions, we can write:

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

Finally, solving for $\mathcal{S}$, we obtain:

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

Once the total measure of incumbents is known, all other aggregate measures can be easily obtained. For instance, the mass of potential entrants is $S_0 = \mathcal{S} h_0$, and the different measures of incumbent sellers can be obtained by $S_n = \mathcal{S} g_n$. The measure of active buyers is $B^A = \mathcal{S} b^A$, and the measure of inactive buyers is $B^I = 1 - B^A$. $\square$

A.3.2 The Invariant Size Distribution when $\delta_f = 0$

In an environment without exogenous (z, φ) shocks, and let us introduce the object $\sigma_n = S_n/(S_0 + \mathcal{S})$, defined for any $n = 0, 1, 2, \dots$. Then, when $\delta_f = 0$, we can re-write the flow equations of Section 3.4 at steady state as:

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

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 $n \rightarrow (n+1)$ and $n \rightarrow (n-1)$, respectively. If we solve the difference equation on $n \geq 1$ recursively, we find:

$$\sigma_n = \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{\delta_c^n} \sigma_0 \tag{A.11}$$

Imposing that $\sum_{n=0}^{+\infty} \sigma_n = 1$ in equation (A.11) yields:

$$\sigma_0 = \left(1 + \sum_{n=1}^{+\infty} \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{\delta_c^n} \right)^{-1} \quad (\text{A.12})$$

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

$$\sum_{n=1}^{+\infty} \frac{1}{n!} \frac{\prod_{i=0}^{n-1} \eta(\theta_{i+1})}{\delta_c^n} < +\infty \quad (\text{A.13})$$

Under necessary condition (A.13), the stationary solution of the birth-death process $\{\sigma_n\}$ is given by (A.11)-(A.12). Using that $g_n = \sigma_n(1 + S_0/S)$ for $n \geq 1$, and $S_0/S = \sigma_0/(1 - \sigma_0)$, we then arrive at the expressions for g_n and S_0 in the main text. $\square$

A.4 Proof of Proposition 2: Existence *[INCOMPLETE]*

A.4.1 Model without shocks

We start with the model without shocks. Provided that $\frac{c+rU^B}{x_n-U^B} < 1$ and $\gamma > 1$, the equilibrium is characterized by:

$$W_{n+1} - W_n - x_{n+1} = \theta_{n+1}^\gamma (x_{n+1} - U^B) \quad (\text{A.14})$$

$$(r + \delta_f)W_n = n \left(v - \delta_c(W_n - W_{n-1}) + (\delta_f + \delta_c)U^B \right) - \mathcal{C}(n) + \eta(\theta_{n+1})(W_{n+1} - W_n - x_{n+1}) \quad (\text{A.15})$$

$$\theta_{n+1} = \left[\left(\frac{c + rU^B}{x_{n+1} - U^B} \right)^{-\gamma} - 1 \right]^{\frac{1}{\gamma}} \quad (\text{A.16})$$

These equations are, respectively, the first-order condition of the joint surplus problem, the definition of the joint surplus, and the indifference condition among inactive buyers. This system can be further simplified as follows. First, define

$$\Gamma_n^S \equiv (r + \delta_f)W_n - n \left(v - \delta_c(W_n - W_{n-1}) + (\delta_f + \delta_c)U^B \right) + \mathcal{C}(n)$$

and note that $\Gamma_n^S = \eta(\theta_{n+1})(W_{n+1} - W_n - x_{n+1})$ by (A.15). Next, we see that $\eta(\theta_{n+1})(W_{n+1} - W_n - x_{n+1}) = \eta(\theta_{n+1})\theta_{n+1}^\gamma (x_{n+1} - U^B) = \eta(\theta_{n+1})\theta_{n+1}^\gamma \frac{\Gamma_n^B}{\mu(\theta_{n+1})} = \theta_{n+1}^{1+\gamma} \Gamma_n^B$, where the first equality uses (A.14), the second one uses (A.16), and the last one uses $\eta(\theta) = \theta\mu(\theta)$. Thus, we have found:

$$\theta_{n+1} = \left(\frac{\Gamma_n^S}{\Gamma_n^B} \right)^{\frac{1}{1+\gamma}} \quad (\text{A.17})$$

Next, we can see from (A.14) that $W_{n+1} - W_n = x_{n+1} + \theta_{n+1}^\gamma (x_{n+1} - U^B) = x_{n+1}(1 + \theta_{n+1}^\gamma) - \theta_{n+1}^\gamma U^B$, so:

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

where the second equality uses (A.16). Now, plugging in (A.17), we find $W_{n+1} - W_n - U^B = \Gamma_n^B \left(1 + \left(\frac{\Gamma_n^S}{\Gamma_n^B} \right)^{\frac{\gamma}{1+\gamma}} \right)^{\frac{1+\gamma}{\gamma}}$, that is:

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

This is a recursive formula for W_n (a second-order difference equation). The boundary conditions are $W_0 = 0$, and $\kappa = \eta(\theta_1)(W_1 - x_1)$ (the latter from $\kappa = \eta(\theta_1)V_1^S$, and $W_1 \equiv V_1^S + x_1$). But recall that $\eta(\theta_1)(W_1 - x_1) = \theta_1^{1+\gamma} \Gamma_1^B$, so then:

$$\theta_1 = \left(\frac{\kappa}{\Gamma_1^B} \right)^{\frac{1}{1+\gamma}}$$

That is, using (A.17) at $n = 1$, we have $\Gamma_0^S = \kappa$, and thus

$$W_1 - U^B = \left((\Gamma_1^B)^{\frac{\gamma}{1+\gamma}} + \kappa^{\frac{\gamma}{1+\gamma}} \right)^{\frac{1+\gamma}{\gamma}}$$

To make progress, say that the cost function is $\mathcal{C}(n) = wn^\psi$, where $w, \psi > 0$. Define $\Delta_n \equiv W_n - W_{n-1}$,

$\beta \equiv v + (\delta_f + \delta_c)U^B$, and $w_n \equiv \mathcal{C}(n)/n$. Note $W_n = \sum_{j=1}^n \Delta_j$. Define the mapping:

$$\mathcal{T} : x \mapsto \left(\Gamma^{B \frac{\gamma}{1+\gamma}} + x^{\frac{\gamma}{1+\gamma}} \right)^{\frac{1+\gamma}{\gamma}}$$

Then, above we found that:

$$\Delta_{n+1} = U^B + \mathcal{T}(\Gamma_n^S)$$

with initial condition $\Delta_1 = U^B + \mathcal{T}(\kappa)$, and

$$\Gamma_n^S = (r + \delta_f) \sum_{j=1}^n \Delta_j - n \left(v + (\delta_f + \delta_c)U^B - w_n - \delta_c \Delta_{n-1} \right)$$

where we have defined $w_n \equiv \mathcal{C}(n)/n$.

We want to show that $\Delta_n > 0$ (so that the joint surplus increases with the size of the firm) and $\Delta_{n+1} < \Delta_n$ (so that the joint surplus is concave). Note that the latter condition will ensure that both θ_n and x_n is decreasing (since these functions inherit monotonicity of Δ_n through the IC+FOC). In turn, x_n being decreasing is a necessary (though not sufficient) condition for p_n to be increasing.

Need to show: $\forall n \geq 1$: $\Delta_n > 0$ (i.e. W_n is increasing) and $\Delta_{n+1} < \Delta_n$ (i.e. W_n is concave).

1. **Part 1: W increasing** $\Delta_n > 0, \forall n \geq 1$.

- $\Delta_1 = U^B + \mathcal{T}(\kappa) > 0$, which is true because $\kappa > 0$.
- $\Delta_2 > 0$ iff $0 < U^B + \mathcal{T}(\Gamma_1^S) = U^B + \mathcal{T}((r + \delta_f + \delta_c)\mathcal{T}(\kappa) + rU^B - v + w_1)$. A necessary condition (else imaginary parts) is:

$$\mathcal{T}(\kappa) > \frac{v - w_1 - rU^B}{r + \delta_f + \delta_c}$$

This condition is also sufficient because it immediately implies $\Delta_2 > 0$.

- $\Delta_3 > 0$ iff $0 < U^B + \mathcal{T}(\Gamma_2^S) = \mathcal{T}((r + \delta_f)\mathcal{T}(\kappa) + (r + \delta_f + 2\delta_c)\mathcal{T}((r + \delta_f + \delta_c)\mathcal{T}(\kappa) + rU^B - v + w_1) + 2(rU^B - v + w_1))$. A necessary condition (or else imaginary parts) is:

$$(r + \delta_f)\mathcal{T}(\kappa) + (r + \delta_f + 2\delta_c)\mathcal{T}((r + \delta_f + \delta_c)\mathcal{T}(\kappa) + rU^B - v + w_1) + 2(rU^B - v + w_1) > 0$$

which again is also a sufficient condition for it implies $\Delta_3 > 0$.

- ...
- Since we are looking for a decreasing Δ_n sequence, then it is enough to check that $\Delta_{\bar{n}} > 0$ (i.e. at the last n). Generally, $\Delta_n > 0$ iff $\Gamma_{n-1}^S > 0$ iff $(r + \delta_f) \sum_{j=1}^n \Delta_j + n(\delta_c \Delta_n - \beta + w_n) > 0$ iff $(r + \delta_f) \sum_{j=1}^n \Delta_j > n(\beta - w_n - \delta_c \Delta_n)$.³⁹ From here, note it helps that w_n is increasing in n (i.e. that $\mathcal{C}(n)$ is convex in n).
- Induction proof:

– For $n = 1$, $\Delta_1 > 0$ is guaranteed. For $n = 2$, $\Delta_2 > 0$ iff $\mathcal{T}(\kappa) > \frac{v - w_1 - rU^B}{r + \delta_f + \delta_c}$, which is assumed.

³⁹ A sufficient (though not necessary) condition is $\beta - w_n - \delta_c \Delta_n < 0$ iff $\Delta_n > (\beta - w_n)/\delta_c$. But this is generally too restrictive with the restrictions of Part 2, so we will not consider this case.

- Now suppose $\Delta_n > 0$ for arbitrary $n \geq 2$. WTS $\Delta_{n+1} > 0$. Note $\Delta_{n+1} = U^B + \mathcal{T}\left((r + \delta_f) \sum_{j=1}^n \Delta_j + n(\delta_c \Delta_n - \beta + w_n)\right)$, where $\Delta_j > 0$ for all $j = 1, \dots, n$ by induction. A necessary and sufficient condition is $(r + \delta_f) \sum_{j=1}^n \Delta_j + n\delta_c \Delta_n > n(\beta - w_n)$. This is more likely to be true when $\beta \approx w_n$

2. **Part 2: W concave** $\Delta_{n+1} < \Delta_n, \forall n \geq 1$.

- Since $\mathcal{T}$ is an increasing mapping, $\Delta_{n+1} < \Delta_n$ iff $\Gamma_n^S < \Gamma_{n-1}^S$. Note:

$$\Gamma_n^S - \Gamma_{n-1}^S = (r + \delta_f + n\delta_c)\Delta_n - (n-1)\delta_c\Delta_{n-1} + \mathcal{M}_n - \beta$$

where $\mathcal{M}_n \equiv \mathcal{C}(n) - \mathcal{C}(n-1)$ is the marginal cost for the n^{th} customer [Note: if $\mathcal{C}$ is convex in n , then $\mathcal{M}_1 = w_1$ and $\mathcal{M}_n > w_n, \forall n \geq 2$]. Thus, $\Gamma_n^S - \Gamma_{n-1}^S < 0$ iff:

$$\Delta_n < \frac{\beta - \mathcal{M}_n + (n-1)\delta_c\Delta_{n-1}}{r + \delta_f + n\delta_c}$$

- Induction:

- To start, we need $\Delta_2 < \Delta_1$ iff $\mathcal{T}((r + \delta_f + \delta_c)\mathcal{T}(\kappa) - \beta + w_1) < \mathcal{T}(\kappa)$ iff $\mathcal{T}(\kappa) < \frac{\kappa + \beta - w_1}{r + \delta_f + \delta_c}$.
- By induction, suppose $\Delta_n < \Delta_{n-1}$ for arbitrary $n \geq 2$. WTS $\Delta_{n+1} < \Delta_n$, i.e. $\Delta_n < \frac{\beta - \mathcal{M}_n + (n-1)\delta_c\Delta_{n-1}}{r + \delta_f + n\delta_c}$ (as derived above). A sufficient (though not necessary) condition is to “squeeze Δ_{n-1} in”, i.e.

$$\Delta_n < \Delta_{n-1} < \frac{\beta - \mathcal{M}_n + (n-1)\delta_c\Delta_{n-1}}{r + \delta_f + n\delta_c} \quad \text{iff} \quad \Delta_{n-1} < \frac{\beta - \mathcal{M}_n}{r + \delta_f + \delta_c}$$

Since Δ_n needs to be decreasing, then if $\mathcal{M}_n$ is decreasing (i.e. $\psi < 1$, so that $\mathcal{C}$ is concave) it suffices to check this for $n = 2$, i.e. $\Delta_1 = T(\kappa) < \frac{\beta - \mathcal{M}_2}{r + \delta_f + \delta_c}$. Of course, a necessary condition is then that $\beta - \mathcal{M}_2 > 0$, i.e. $v + (\delta_f + \delta_c)U^B > \mathcal{C}(2) - \mathcal{C}(1) = w(2^\psi - 1)$ iff

$$v + (\delta_f + \delta_c)U^B > w(2^\psi - 1)$$

- For $n = 2$, the condition says $\Delta_1 = T(\kappa) < \frac{\beta - \mathcal{M}_2}{r + \delta_f + \delta_c}$. To avoid a contradiction with $\Delta_2 > 0$, we need $\beta - \mathcal{M}_2 > \Lambda_1$, but $\Lambda_1 = \beta - \mathcal{C}(1)$, so we need $\beta - (\mathcal{C}(2) - \mathcal{C}(1)) > \beta - \mathcal{C}(1)$ iff $\mathcal{C}(1) > \mathcal{C}(2) - \mathcal{C}(1)$, which means $\mathcal{C}(n)$ must be concave in the first step.

[TO BE COMPLETED]

A.4.2 Model with shocks

For each $(z, \varphi) \in \mathcal{Z} \times \Phi$, the set of sequences $\Delta \equiv \{\Delta_n(z, \varphi) : n = 1, 2, \dots\}$ solves the recursion:

$$\Delta_{n+1}(z, \varphi) - U^B(\varphi) = \left(\Gamma^B(\varphi)^{\frac{\gamma}{1+\gamma}} + \Gamma_n^S(z, \varphi)^{\frac{\gamma}{1+\gamma}} \right)^{\frac{1+\gamma}{\gamma}}$$

where $\Delta_n(z, \varphi) \equiv W_n(z, \varphi) - W_{n-1}(z, \varphi)$, and:

$$\Gamma_n^S(z, \varphi) := \left(r + \delta_f + \hat{\lambda}(z, \varphi) \right) \sum_{j=1}^n \Delta_j(z, \varphi) + n \left(\delta_c \Delta_n(z, \varphi) - \beta(\varphi) \right) + \mathcal{C}(n; z, \varphi) - \Xi_n(z, \varphi)$$

with $\beta(\varphi) \equiv v + (\delta_f + \delta_c)U^B(\varphi)$; $\hat{\lambda}(z, \varphi) \equiv \sum_{z'} \lambda_z(z'|z) + \sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi)$; and:

$$\Xi_n(z, \varphi) \equiv \sum_{j=1}^n \left(\sum_{z'} \lambda_z(z'|z) \Delta_j(z', \varphi) + \sum_{\varphi'} \lambda_\varphi(\varphi'|\varphi) \Delta_j(z, \varphi') \right)$$

The initial condition of the recursion (at $n = 1$) uses $\Gamma_0^S(z, \varphi)$, which satisfies the free entry condition:

$$\kappa = \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \Gamma_0^S(z_0, \varphi)$$

Again, we need to show:

1. $\Delta_n(z, \varphi) > 0, \forall (n, z, \varphi) \in \mathbb{N} \times \mathcal{Z} \times \Phi$ (i.e. $W_n(z, \varphi)$ is increasing in n).
 - **n = 1** $\Delta_1(z, \varphi) = U^B(\varphi) + \mathcal{T}(\Gamma_0^S(z, \varphi)) > 0$. A necessary condition (else imaginary parts) is $\Gamma_0^S(z, \varphi) > 0$. Since Γ_0^S is increasing in z and φ (**proof?**), then we need to impose $\Gamma_0^S(\underline{z}, \underline{\varphi}) > 0$.
 - **n = 2** $0 < \Delta_2(z, \varphi) = U^B(\varphi) + \mathcal{T}(\Gamma_1^S(z, \varphi)) = U^B(\varphi) + \mathcal{T}\left((r + \delta_f + \delta_c + \hat{\lambda}(z, \varphi)) [U^B(\varphi) + \mathcal{T}(\Gamma_0^S(z, \varphi))] + \mathcal{C}(1; z, \varphi) - v - (\delta_f + \delta_c)U^B(\varphi) - \Xi_1(z, \varphi)\right)$. A necessary condition (else imaginary parts) is:

$$(r + \hat{\lambda}(z, \varphi)) \underbrace{\left(U^B(\varphi) + \mathcal{T}(\Gamma_0^S(z, \varphi)) \right)}_{=\Delta_1(z, \varphi)} + (\delta_f + \delta_c) \mathcal{T}(\Gamma_0^S(z, \varphi)) + \mathcal{C}(1; z, \varphi) - v - \Xi_1(z, \varphi) > 0$$

[TO BE COMPLETED]

2. $\Delta_{n+1}(z, \varphi) < \Delta_n(z, \varphi), \forall n \geq 2, \forall (z, \varphi) \in \mathcal{Z} \times \Phi$ (i.e. $W_n(z, \varphi)$ is concave in n).

[TO BE COMPLETED]

A.5 Proof of Proposition 3: Efficiency [INCOMPLETE]

Proof. Consider a benevolent planner that is constrained by the search frictions of the economy and wants to maximize aggregate welfare subject to the resource constraints of the economy. As usual, the planner can allocate resources freely, so the problem does not feature contracts nor prices. Instead, we label each submarket 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 $\varphi^t = (\varphi_j : j \leq t)$.

The planner's choice variables are time series for the tightness in each market segment, $\Theta_t \equiv \{\theta_{n_t, t}(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\}$; distributions of inactive and active buyers across markets, $\mathbf{B}_t^I \equiv \{B_{n_t, t}^I(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\}$ and $\mathbf{B}_t^A \equiv \{B_{n_t, t}^A(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\}$ (respectively); a measure potential entrants $S_{0, t}$; and a distribution of firms across states, $\mathbf{S}_t \equiv \{S_{n_t, t}(z_t) : (n_t, z_t) \in \mathbb{N} \times \mathcal{Z}\}$.

Aggregate welfare maximization requires:

$$\left(\max_{\substack{\Theta_t, \mathbf{B}_t^I, \mathbf{B}_t^A \\ S_{0, t}, \mathbf{S}_t}} \right) \mathbb{E}_0 \int_0^{+\infty} e^{-rt} \left\{ -\kappa S_{0, t} + \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} \left[v B_{n_t, t}^A(z_t) - S_{n_t, t}(z_t) \mathcal{C}(n_t; z_t, \varphi_t) - c(\varphi_t) B_{n_t, t}^I(z_t) \right] \right\} dt \quad (\text{A.18})$$

Total welfare in this economy is thus equal to aggregate consumption gains for active buyers, net of production operating costs for active sellers,⁴⁰ and net of the search and entry costs for inactive buyers and inactive sellers.

⁴⁰ The sum of these first two terms thus equals the total gains from trade.

The planner is subject to three sets of constraints. The first set concerns the evolution of the firm distribution (which we described in Section 3.4):

$$\partial_t S_{0,t} = \delta_f \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} S_{n_t,t}(z_t) + \delta_c \sum_{z_t \in \mathcal{Z}} S_{1,t}(z_t) - \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \eta(\theta_{1,t}(z_0)) S_{0,t} \quad (\text{A.19a})$$

$$\begin{aligned} \partial_t S_{1,t}(z_t) &= \pi_z(z_t) \eta(\theta_{1,t}(z_t)) S_{0,t} + 2\delta_c S_{2,t}(z_t) + \sum_{\tilde{z} \neq z_t} \lambda_z(z_t|\tilde{z}) S_{1,t}(\tilde{z}) \\ &\quad - \left(\delta_f + \delta_c + \eta(\theta_{2,t}(z_t)) + \sum_{\tilde{z} \neq z_t} \lambda_z(\tilde{z}|z_t) \right) S_{1,t}(z_t) \end{aligned} \quad (\text{A.19b})$$

$$\begin{aligned} \forall n_t \geq 2: \partial_t S_{n_t,t}(z_t) &= \eta(\theta_{n_t,t}(z_t)) S_{n_t-1,t}(z_t) + (n_t + 1)\delta_c S_{n_t+1,t}(z_t) + \sum_{\tilde{z} \neq z_t} \lambda_z(z_t|\tilde{z}) S_{n_t,t}(\tilde{z}) \\ &\quad - \left(\delta_f + n_t \delta_c + \eta(\theta_{n_t+1,t}(z_t)) + \sum_{\tilde{z} \neq z_t} \lambda_z(\tilde{z}|z_t) \right) S_{n_t,t}(z_t); \end{aligned} \quad (\text{A.19c})$$

for all $z_t \in \mathcal{Z}$. Note that implicit in these equations lies the evolution of any firm's size (i.e. the law of motion for n_t). The second set of constraints describes the distribution of buyers across states:

$$\forall z_t \in \mathcal{Z}: B_{n_t,t}^A(z_t) = n_t S_{n_t,t}(z_t) \quad (\text{A.20a})$$

$$\forall z_t \in \mathcal{Z}: B_{n_t,t}^I(z_t) = \theta_{n_t,t}(z_t) S_{n_t-1,t}(z_t) \quad (\text{A.20b})$$

$$1 = \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} \left(B_{n_t,t}^A(z_t) + B_{n_t,t}^I(z_t) \right) \quad (\text{A.20c})$$

These equations state that each customer consumes a single unit (equation (A.20a)); 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 (A.20b)); and that every buyer in the economy is in either the active or the inactive state (equation (A.20c)).

Finally, the mass of potential entering firms needs to be non-negative to ensure that there is no negative entry:

$$S_{0,t} \geq 0 \quad (\text{A.21})$$

Firstly, we can simplify the problem by substituting constraints (A.20a)-(A.20b)-(A.20c) into the objective function. To solve the simplified planner's problem, we use standard tools from Optimal Control theory. The state vector has been reduced to $\mathbb{S}_t \equiv \{S_{0,t}, \mathbf{s}_t\}$, whereas Θ_t is the control variable. The current-value Hamiltonian of the planning problem reads:

$$\begin{aligned}
\mathcal{H}_t(\Theta_t; \mathbb{S}_t) = & -c(\varphi_t) - \kappa S_{0,t} + \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} S_{n_t,t}(z_t) \left(n_t(v + c(\varphi_t)) - \mathcal{C}(n_t; z_t, \varphi_t) \right) \\
& + \psi_{0,t} \left[\delta_f \sum_{n_t=1}^{+\infty} \sum_{z_t \in \mathcal{Z}} S_{n_t,t}(z_t) + \delta_c \sum_{z_t \in \mathcal{Z}} S_{1,t}(z_t) - \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \eta(\theta_{1,t}(z_0)) S_{0,t} \right] \\
& + \sum_{z_t \in \mathcal{Z}} \left\{ \psi_{1,t}(z_t) \left[\pi_z(z_t) \eta(\theta_{1,t}(z_t)) S_{0,t} + 2\delta_c S_{2,t}(z_t) + \sum_{\tilde{z} \neq z_t} \lambda_z(z_t|\tilde{z}) S_{1,t}(\tilde{z}) - \left(\delta_f + \delta_c + \eta(\theta_{2,t}(z_t)) + \sum_{\tilde{z} \neq z_t} \lambda_z(\tilde{z}|z_t) \right) S_{1,t}(z_t) \right] \right. \\
& + \sum_{n_t=2}^{+\infty} \psi_{n_t,t}(z_t) \left[\eta(\theta_{n_t,t}(z_t)) S_{n_t-1,t}(z_t) + (n_t+1)\delta_c S_{n_t+1,t}(z_t) + \sum_{\tilde{z} \neq z_t} \lambda_z(z_t|\tilde{z}) S_{n_t,t}(\tilde{z}) \right. \\
& \left. \left. - \left(\delta_f + n_t\delta_c + \eta(\theta_{n_t+1,t}(z_t)) + \sum_{\tilde{z} \neq z_t} \lambda_z(\tilde{z}|z_t) \right) S_{n_t,t}(z_t) \right] \right\} + \vartheta_t S_{0,t}
\end{aligned}$$

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, $S_{0,t}$), and $\vartheta_t \geq 0$ is the multiplier on the non-negative entry condition, with $\vartheta_t S_{0,t} = 0$ by complementary slackness.

In vector notation, the necessary conditions for optimality are:

$$\nabla_{\Theta} \mathcal{H}_t(\Theta_t; \mathbb{S}_t) = \mathbf{0} \quad (\text{A.22a})$$

$$\nabla_{\mathbb{S}} \mathcal{H}_t(\Theta_t; \mathbb{S}_t) = -\nabla_t \psi_t + r \psi_t \quad (\text{A.22b})$$

where ∇ denotes the gradient, and ψ_t is the vector of co-state variables. Throughout, we impose $\nabla_t \psi_t = \mathbf{0}$ by stationarity. 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.

For the set of conditions (A.22a), given $z_t \in \mathbb{Z}$ we have:

$$\begin{aligned}
n = 1 : \quad 0 &= (\psi_{1,t}(z_t) - \psi_{0,t}) \pi_z(z_t) \eta'(\theta_{1,t}(z_t)) S_{0,t} \\
n \geq 2 : \quad 0 &= (\psi_{n,t}(z_t) - \psi_{n-1,t}(z_t)) \eta'(\theta_{n,t}(z_t)) S_{n-1,t}(z_t)
\end{aligned}$$

As for the set of conditions (A.22b), we have:

$$\begin{aligned}
n = 0 : \quad & -\partial_t \psi_{0,t} + r \psi_{0,t} = -\kappa - \psi_{0,t} \sum_{z_0 \in \mathcal{Z}} \pi_z(z_0) \eta(\theta_{1,t}(z_0)) + \sum_{z_t \in \mathcal{Z}} \psi_{1,t}(z_t) \pi_z(z_t) \eta(\theta_{1,t}(z_t)) + \vartheta_t \\
n \geq 1 : \quad & -\partial_t \psi_{n,t}(z) + r \psi_{n,t}(z) = n(v + c(\varphi_t)) - \mathcal{C}(n, z; \varphi_t) + \delta_f (\psi_{0,t} - \psi_{n,t}(z)) \\
& + n\delta_c (\psi_{n-1,t}(z) - \psi_{n,t}(z)) - \sum_{\tilde{z} \neq z} \lambda_z(\tilde{z}|z) \psi_{n,t}(\tilde{z}) \\
& + \eta(\theta_{n+1,t}(z)) (\psi_{n+1,t}(z) - \psi_{n,t}(z))
\end{aligned}$$

where for simplicity we call $\psi_{0,t}$ by $\psi_{n-1,t}(z)$ when $n = 1$. We can now show that a block-recursive equilibrium

with non-negative entry of firms satisfies the optimality conditions on the planner by making an appropriate guess for the co-state variables of the planning problem. For this, we pick a decentralized equilibrium allocation $\{\mathcal{W}^*(\varphi_t), \mathcal{X}^*(\varphi_t), \Theta^*(\varphi_t), U^B(\varphi_t)\}$, and guess:

$$\begin{aligned} \psi_{0,t}(\varphi^t) &= 0 \\ \forall n \geq 1 : \quad \psi_{n,t}(z_t, \varphi^t) &= W_n(z_t, \varphi_t) \end{aligned}$$

[TO BE COMPLETED]

B Model with a CES Industry Structure

The stationary equilibrium described in Section 3.2 can be embedded into a familiar [Dixit and Stiglitz \(1977\)](#) environment in which the single homogenous product is only one of a continuum of measure one of different product varieties, over which the buyers have CES preferences. Nesting the above economy into a collection of independent industries, producing imperfectly substitutable varieties of the same final good, is a parsimonious way of obtaining a generalized model of cross-industry dynamics.

Let $Y(\nu)$ be the aggregate output of variety ν of the final good, and

$$\bar{Y} = \left(\int_0^1 Y(\nu)^{\frac{\sigma-1}{\sigma}} d\nu \right)^{\frac{\sigma}{\sigma-1}}$$

be the aggregate output of the economy, where $\sigma > 1$ is the elasticity of substitution between varieties. Each variety can be thought of as being produced in an industry whose functioning is identical to that of our baseline model. From our analysis above, we know that aggregate industry output is given by $Y(\nu) = \sum_{n=1}^{+\infty} n S_n(\nu)$, where the right side of the equality is the total mass of buyers shopping in industry $\nu \in [0, 1]$, and $S_n(\nu)$ is the total mass of firms of size n producing good ν in equilibrium. Moreover, $B^A := \bar{Y}$ is now the total mass of active buyers in the economy, $B^I = 1 - \bar{Y}$ is the total mass of inactive buyers (of which a mass $B^I(\nu)$ searches in industry ν). We let buyers value each variety differently by indexing their utility accordingly, so v is now a mapping $[0, 1] \rightarrow \mathbb{R}_+$ as opposed to a real number.

Such an environment would lead to the following iso-elastic inverse demand function:

$$Y(\nu) = \left(\frac{P(\nu)}{\bar{P}} \right)^{-\sigma} \bar{Y} \tag{B.1}$$

where $\bar{P} := \left(\int_0^1 P(\nu)^{1-\sigma} d\nu \right)^{\frac{1}{1-\sigma}}$ is the price index of the final good, and $P(\nu)$ is the average price of variety ν . As we have seen, each variety is sold by uncountably many firms of countably many different sizes, each of which sets a different price. Letting $p_n(\nu)$ be the price set by firm of size n , we then have:

$$P(\nu) = \sum_{n=1}^{+\infty} S_n(\nu) p_n(\nu)$$

The aggregate price level $\bar{P}$ is well-defined and finite: since $p_n(\nu) < v(\nu)$ for each $n \in \mathbb{N}$, then $P(\nu) < v(\nu) \mathcal{S}(\nu)$ and therefore $\bar{P} < \left(\int_0^1 [v(\nu) \mathcal{S}(\nu)]^{1-\sigma} d\nu \right)^{\frac{1}{1-\sigma}} < +\infty$. Moreover, along the stationary equilibrium, the demand function (B.1) can be written:

$$\frac{B^A(\nu)}{B^A} = \left(\frac{P(\nu)}{\bar{P}} \right)^{-\sigma}$$

which is a simple accounting equation establishing that the share of active buyers who are consuming variety ν depends on the relative average price of the industry. For instance, for any two $(\nu, \nu') \in [0, 1]^2$, we have $\frac{B^A(\nu)}{B^A(\nu')} = \left(\frac{P(\nu)}{P(\nu')} \right)^{-\sigma}$, i.e. industries whose prices are lower on average are relatively more crowded by consumers.