Customer Capital and Dynamic Barriers to Entry*

Wouter Dessein   Jin Li   Chang Sun†

July 15, 2025

Abstract

Many industries are dominated by large and very profitable firms. We develop a theory of firm dynamics, where competing firms operate a fixed-cost technology but due to customer inertia can only slowly build up a customer base. We show how the interaction between scale economies and customer inertia creates dynamic entry barriers and persistent performance differences. Our model also resolves the 'paradox of entry barriers': markets with higher fixed operating costs have higher long-run profits, but are unambiguously less attractive to enter, except when firms can also invest in product quality. In the latter case, early entrants can create persistent performance differences through upgrading and may want to enter high-cost markets.

*We thank Bruno Cassiman, Alfonso Gambardella and seminar audiences at Bocconi, Columbia Business School, Duke Fuqua, Harvard Business School, HEC Lausanne, KU Leuven, Peking University, University of Melbourne, Wisconsin-Madison as well as various conferences for helpful comments.

†Dessein: Columbia University, Graduate School of Business, wd2179@columbia.edu. Li: HKU Business School, the University of Hong Kong, jli1@hku.hk. Sun: HKU Business School, the University of Hong Kong, sunc@hku.hk.

1 Introduction

Many industries are increasingly dominated by large and very profitable firms. America's 500 largest companies by revenue account for roughly a fifth of employment, half of sales, and two-thirds of profits. At the same time, large firms are also increasingly entrenched: only 52 of the 500 were born after 1990, and their average age has crept up from 75 to 90 years since then.[1] Consistent with these observations, Autor et al. (2020) document the growing dominance of "superstar firms" — highly productive companies that capture disproportionate market shares. As they show, superstar firms are not only larger and more profitable, they also tend to be more productive and invest more in R&D.

This raises several classic questions: Why does entry not limit profitability in these markets (Bain, 1956)? What allows firms to build up a persistent advantage over competitors (Penrose, 2009; Porter, 1980)? How does firm concentration affect investment and product quality (Bain, 1956; Sutton, 1991)? It also creates a strategic puzzle for start-ups looking to enter new or growing markets that are not yet saturated. Industries with high entry barriers tend to be more profitable in the long run, making them more attractive to entrants (Porter, 1980). However, high barriers should also make entry harder and less profitable in the short run, thus reducing their attractiveness for new firms. How should an entrepreneur resolve this 'barriers to entry paradox' when facing multiple entry opportunities?

To address these questions, we provide a theory of dynamic entry barriers, where competing firms operate a fixed-cost technology but can only slowly build up a customer base. A key distinction between established firms and new or recent entrants is that incumbents already have customers. We follow a literature that emphasizes the role of such an established customer base as a source of incumbency advantage (Dubé et al., 2010; Bronnenberg et al., 2009, 2012). Customer inertia is pervasive: fewer than 10% of Americans switch banks in a given year,[2] and the market shares of Colgate or Crest toothpaste today still reflect historical differences in when and where those brands entered (Bronnenberg et al., 2009). The same result holds for consumer packaged goods more generally. Across manufacturing, new plants sell barely 41% percent of the output of otherwise identical incumbent plants, and a 27% percent demand shortfall persists even after a decade (Foster et al., 2016). In the presence of such customer inertia or other product market frictions (search costs, imperfect information), a firm's customer base - its 'customer capital' - then becomes a state variable that only slowly changes over time (Gourio and Rudanko, 2014; Foster et al., 2016). As such,

[1] America's corporate giants are getting harder to topple, The Economist, August 2023.

[2] The real value of consumers switching banks: it's more than you might think, Kearny, June 2023.

it is an important type of 'intangible capital'.

Relative to the literature on customer capital, we emphasize that scale economies are essential for creating a persistent incumbency advantage. Without scale economies, new entrants can slowly build up market share in dynamic environments where customers constantly arrive and depart - their small size poses no disadvantage. Similarly, scale economies without customer inertia do not favor large incumbents, as new entrants or smaller competitors can quickly build up a large customer base. As we argue, it is the interaction between scale economies and customer inertia that creates barriers to entry, high profitability and persistent performance differences. Importantly, we also analyze the impact of scale economies on the attractiveness of entry.

Formally, we consider a dynamic environment where both new customers and new firms arrive each period. Firms sell an experience good and need to grow their customer base by accumulating repeat customers. Production is characterized by exogenous fixed costs, operational expenses that are necessary to operate in the market, as well as endogenous fixed costs (Sutton, 1991): investments that firms can make to upgrade the quality of their product or services. In this environment, we characterize the transition path from a nascent market, where new firms enter, grow a customer base and invest, to a mature, steady state, in which incumbents earn large profits yet entry ceases.

We derive the following insights:

Dynamic Barriers to Entry: The interaction between customer inertia and fixed costs creates a powerful barrier to entry. New firms can only enter profitably if they build a sufficiently large base of repeat customers quickly enough to cover their fixed operational expenses. However, as an industry gets more crowded, it becomes increasingly hard for new entrants to accumulate such a customer base. New firms then eventually stop entering, even while existing firms are profitable, both in the steady state and from an NPV perspective at entry. As we show, when there are only exogenous fixed costs, a firm's value at any point in time (its discounted profits) equals the life-time value of its customer base plus the value of entry in the market (equal to zero in the steady state). While higher fixed operating costs lower short-run profits, they also reduce entry and therefore increase the long-run customer base of incumbents. Long-term profits, therefore, are strictly increasing in fixed operating costs.

Persistent performance differences. When fixed, up-front investments can raise product quality—as in Sutton (1991)—only firms with a sufficiently large or rapidly expanding customer base find the expenditure worthwhile. Because the product is an experience good, quality upgrades pay off chiefly by increasing repeat purchases, so firms wait until their installed base crosses a critical threshold before investing. Late movers often never attain that scale and therefore settle for more basic offerings, remaining smaller and less profitable than early entrants. The smartphone industry illustrates the logic: Apple's vast user base justifies heavy upfront capital spending on custom chips, advanced cameras, and premium displays, whereas most smaller manufacturers skip such upgrades because the incremental returns are too low, further entrenching Apple's lead. A parallel can also be seen in agricultural equipment: Deere leverages an installed base built over more than a century—and the deep loyalty of multigenerational farming families—to invest heavily in precision-agriculture technologies such as GPS-guided automation, telematics, and AI-enabled sprayers (Malnight and Buche, 2022). These costly innovations raise yields for existing customers and lock them into Deere's ecosystem, advantages that smaller rivals, facing a thinner customer base and lower marginal returns, cannot replicate.

In the long run, incumbents who invest in this way are more valuable than late entrants who do not, for three reinforcing reasons: (i) they retain a larger customer base, thanks to lower past attrition; (ii) the lifetime value of their existing customers is higher because those customers are less likely to churn; and (iii) a superior product lowers future attrition, boosting the expected value of yet-to-be-acquired customers. Consequently, an upgraded incumbent's net present value equals not only the lifetime value of its current customers plus the value of entry (as with purely exogenous scale economies) but also the extra surplus it can extract from future customers—surplus a non-upgraded entrant can never match.

Competition and Investment: Customers benefit from investments in quality. Barriers to entry and competition, however, affect the incentives for upgrading. When fixed operating costs are low, or the arrival rate of new firms is high, current or future competition may be so fierce that no firm manages to accumulate a sufficiently large customer base that justifies product quality investments. The industry may then be trapped in a low firm concentration, low product quality equilibrium. Even when some firms eventually upgrade their products, an increase in the barriers to entry in the industry may still benefit consumers by increasing the fraction of firms that offer high-quality products and by speeding up the timing of these investments.

The Paradox of Entry Barriers: While high fixed operating costs lead to higher long-run profits and increase investments in quality, it is unclear if they also make entry more attractive. To resolve this, we consider an entrepreneur who can be an early entrant in one of two otherwise identical markets—one with low fixed costs and another with higher costs. The following trade-off arises. On the one hand, at any time before the last entrant in the high-cost market, the two markets will generate the same flow of new customers, yet a firm in the low-cost market bears a smaller cost burden. On the other hand, once entry stops in the high-cost market, a firm operating in the latter market will see a higher in-flow of customers.

With only exogenous fixed costs, the zero-profit condition for entry implies that those higher customer inflows are more than off-set by the higher operating costs. Indeed, once entry stops in the high-cost market, the value of a firm equals the life-time value of its customer base (identical in both markets) plus the value of entry, which is zero in the high-cost market but still positive in the low-cost market. As a result, entry in the low-cost market is always more attractive.

The calculus changes once endogenous fixed costs—optional investments that permanently raise product quality and customer retention—enter the picture. Early entrants who upgrade can monetize future buyers better than late entrants who never upgrade. Because the high-cost market experiences less subsequent entry, it delivers a larger stream of future customers; for an upgrader, that larger stream now more than offsets the heavier cost burden once entry stops. Hence, provided entry stops quickly enough, an early entrant who intends to upgrade, optimally chooses the high-cost market. Conversely, late entrants or firms that never upgrade still favor the low-cost market, because they cannot capture the extra value embodied in the larger customer flow. As we show, the above results generalize to a general equilibrium setting in which not just one entrepreneur, but all entrepreneurs, at any time, can enter either market.

Taken together, the analysis highlights a nuanced "barriers-to-entry paradox": high fixed costs make industries more profitable in the long run, but whether they raise or lower the attractiveness of entering depends on the entrant's timing and strategy. Early movers planning to invest in quality may rationally seek out the tougher, high-cost markets, whereas latecomers or firms that see no opportunity to generate persistent performance differences through upgrading, rationally flee to low-cost markets that offer quicker but ultimately smaller payoffs.

Outline. After discussing the related literature, Section 2 introduces our continuous-time model of firm dynamics. Section 3 solves for the equilibrium, both transition paths and steady state, when there are only exogenous fixed operational costs and firms make only one decision: whether or not to enter. Section 4 introduces endogenous fixed costs-investments in product quality, where, in addition to the entry decision, firms also need to decide whether and when to upgrade their product quality. Section 5 analyzes the paradox of entry barriers by allowing for entry in two markets that differ in their level of fixed operational costs (and, hence, future entry). We conclude in Section 6.

1.1 Literature

The persistence of dominant firms has been formally studied since at least Gilbert and Newbery (1982), who argued that incumbent firms have an incentive to maintain their monopoly power by pre-emptive innovation or patenting. We are further related to a classic literature in industrial organization that has shown how, in markets for experience goods, early entrants enjoy a competitive advantage relative to latecomers merely as a consequence of having entered sooner (Schmalensee, 1982; Bagwell, 1990; Grossman and Horn, 1988). We add to this literature the crucial interaction between demand advantages of early entrants, on the one hand, and scale economies in the form of fixed operational costs and investments in product upgrades, on the other hand. The presence of both ingredients is necessary for barriers to entry to exist in a fully dynamic model (where new firms and new consumers arrive in each period) and for long-term performance differences to survive in steady state. As we consider a continuum of firms, the strategic pricing incentives and pre-emptive innovation emphasized in this literature are also absent in our model.

Our paper further contributes to the literature on firm dynamics. In many of the classic papers, such as Jovanovic (1982), Hopenhayn (1992), Atkeson and Kehoe (2005) and Luttmer (2007), firm-specific productivity shocks determine entry and exit rates and, hence, firm dynamics. Much of this literature does not allow firms to make investments in quality or innovation, addressing other questions instead. The literature on customer capital (Gourio and Rudanko, 2014; Foster et al., 2016), for example, studies how the intangible capital embodied in a firm's customer base affects a firm's pricing strategy, and how convex costs of customer acquisition limit firm expansion.

Papers that do endogenize innovation decisions typically focus on what happens in the steady state or solve the dynamic innovation problem numerically. In contrast, our focus is on the transitional dynamics of the industry, which we characterize analytically.[3] Closest related to us is Klepper (1996), who emphasizes the importance of firm size in appropriating the returns from process innovation. Unlike in our paper, however, firms are myopic in their investment decisions and it is simply assumed that larger firms have a demand advantage over smaller ones: at equal prices, firms capture a share of demand in proportion to their (historical) size. This gives them more incentives for process innovation, as the value of a unit cost reduction is proportional a firm's output. Also related is a recent literature on firm reputation (Atkeson et al., 2015; Cabral, 2016; Board and Meyer-ter Vehn, 2022; Vellodi,

[3] Jovanovic and MacDonald (1994) also focuses on transitional dynamics, in particular shake-outs. But they do not allow firms to make endogenous choices about investing in new technologies. A firm's technology evolves exogenously.

2022). In the latter papers, consumers learn about quality from public information, and a firm's reputation acts as a state variable governing firm dynamics.[4] In contrast, in our model, consumers are repeat buyers and learn about quality from using a product, making a firm's customer base a state variable.

Finally, we contribute to the debate on the correlation between firm concentration and investments in intangibles (R&D, Software, Brand value) (Crouzet and Eberly, 2019); and more broadly, the impact of firm concentration on consumer welfare and firm productivity (see Autor et al. (2020), and Syverson (2019) for an excellent overview). As argued informally by Crouzet and Eberly (2019), one mechanism through which the intangibles-concentration connection can occur is that market share shifts to firms that invest more in intangibles, as this allows them to deliver a higher-quality product at a lower price.[5] Compared to the latter paper, the relationship in our model goes two ways. First, larger firms are more likely to invest in product quality/innovation than smaller firms as the return of such investments depends on their (current and future) customer base. Secondly, because they invest more in product quality, they grow faster than smaller firms, gaining a permanent advantage.

2 Simple Model Setup

We consider a simple model of firm dynamics. Both entrepreneurs and customers arrive at a constant rate in a market. Entrepreneurs decide whether to start a firm and enter. Firms that enter operate a fixed cost technology and slowly grow their customer base. New customers are randomly matched with firms, whereas existing customers decide whether to stick with their match or leave. The retention rate of customers is assumed to be fixed in our baseline model. In Section 4, we allow firms to invest and upgrade the quality of their product, which increases customer retention.

[4] Vellodi (2022), for example, studies how consumer reviews and rating design by a platform affect barriers to entry and incumbency advantages. Atkeson et al. (2015) shows that entry restrictions (e.g. entry taxes) may increase incentives for quality. They only focus on steady-state outcomes, however, and quality investments must be made upon entry.

[5] Autor et al. (2020) present a formal model in which they link the rise of superstar firms to an increase in the toughness of competition that shifts sales to the most productive firms. They also note that: "Similar results could occur from any force that makes the industry more concentrated—more "winner takes most"—such as an increased importance of network effects or scale-biased technological change from information technology advances." (p.656)

2.1 Model Primitives

Formally, we consider a continuous-time model over the interval $[0, \infty)$. At any time $t > 0$, μ new entrepreneurs arrive and decide whether to start a firm and enter. Firms can choose to permanently leave the market at any time point. Let $G(t)$ represent the total mass of firms in the market at time t.

At any time t, N new customers arrive and are distributed equally across the existing firms. Hence, at time t, each existing firm has an inflow of $N/G(t)$ new customers. To prevent firms from receiving an infinite number of customers at $t = 0$, we assume an initial mass $G_0 > 0$ of entrepreneurs, where G_0 is sufficiently small so that all time 0 entrepreneurs enter: $G(0) = G_0$.

For each firm in the market, its size is given by the number of customers it serves. We denote by $m(t; t_e)$ the customer base at time t of a firm that entered at time $t_e \leq t$. Each firm starts with a zero customer base upon entry, $m(t_e; t_e) = 0$. Firms then slowly grow their customer base by accumulating repeat customers. We posit that $m(t; t_e)$ evolves as

$$\frac{\mathrm{d}m(t; t_e)}{\mathrm{d}t} = -\alpha m(t; t_e) + \frac{N}{G(t)}, \tag{1}$$

where $\alpha \in (0, \infty)$ is the attrition rate of existing customers. Solving the ordinary differential equation (1), we obtain

$$m(t; t_e) = \int_{t_e}^{t} \frac{N}{G(s)} e^{-\alpha(t-s)} \mathrm{d}s. \tag{2}$$

Note that as long as the customer inflow $N/G(s)$ is bounded away from zero and $t > t_e$, the customer base $m(t, t_e)$ will be strictly positive for any attrition rate $\alpha \in (0, \infty)$.

In Section 3, we set the customer attrition rate α as an exogenous parameter. Section 4 endogenizes α by allowing firms to make a one-time investment in product quality, $I > 0$, to reduce α from α_L to $\alpha_H < \alpha_L$. Both the choice to invest and its timing are optimally chosen by the firm.

Firms that enter incur fixed operating costs, modeled as a flow cost c, regardless of the amount of customers they have. Firms also earn one unit of revenue per customer. Hence, the instantaneous profit of each firm is $m(t; t_e) - c$. If a firm never exits, its discounted profits at entry (gross of any investment) are given by

$$V^e(t_e) = \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left(m(t; t_e) - c\right) \mathrm{d}t, \tag{3}$$

where $\rho > 0$ is the discount factor. More generally, the discounted payoffs from time t_1 to t_2 (gross of any investment) are $\int_{t_1}^{t_2} e^{-\rho(t-t_e)} \left(m(t;t_e) - c\right) \mathrm{d}t$.

2.2 Market equilibrium

Our goal is to describe the industry dynamics in which firms make their entry and exit decisions optimally. For ease of exposition, again assume that firms that enter the market never exit, then we say that $\{G(t), \mu(t)\}_{t=0}^{\infty}$ is an industry equilibrium if two conditions are satisfied. First, the mass of firms in the market at time t, $G(t)$, is equal to the total mass of firms that entered the market up to time t. Second, the discounted profits of the firms are non-negative whenever a positive mass of firms enter. Formally, for each time t,

$$G(t) = \int_0^t \mu(\tau) d\tau, \quad 0 \leq \mu(t) \leq \mu, \tag{4}$$

and $V^e(t) \geq 0$ whenever $\mu(t) > 0$. The industry equilibrium where firms can exit can be defined accordingly.

2.3 Discussion of the Model

We now discuss some key assumptions of our model.

First, firms incur a flow cost c rather than a one-time sunk entry cost C. These setups are equivalent in our context since firms never exit (their flow payoffs remain non-negative), allowing us to interpret c as an entry barrier.

Second, we posit an exogenous customer attrition rate α and assume that firms charge a constant price (normalized to one). In Appendix B, we outline a simple continuous-time dynamic discrete choice model of consumers following Arcidiacono et al. (2016), which allows us to microfound α and endogenize firms' prices. Section 4.3 provides a detailed discussion of these micro-foundations.

Third, we assume that new customers are randomly matched to firms. In Online Appendix OA.3, we extend our model to allow higher-quality firms to not only have lower customer attrition, but also have an advantage in attracting new customers. This extension is only relevant in Section 4 when firms can invest in product quality.

Finally, our baseline model considers entry into a single market and assumes a constant arrival rate μ of new firms. Section 5 shows that our framework naturally extends to multimarket entry decisions. This also allows us to endogenize the entry rate μ in a given market, and let it be time-varying, as we do in Section 5.3.

3 Industry Equilibrium with Exogenous Fixed Costs

We first consider the simplest version of our model in which entrepreneurs only decide whether to enter/stay in the market (and incur the fixed operational cost c if they do so). The customer attrition rate α is exogenously given. In Section 4, we will endogenize α by letting firms invest in product quality.

3.1 Dynamic Barriers to Entry

Note first that new entrepreneurs cannot enter the market forever. If they did, the mass of firms $G(t)$ would grow without bound and the inflow of new customers for a new entrant would converge to 0. Given the fixed operating cost c, this makes entry unprofitable. In Appendix A, we also rule out cycles of entry and exit. Hence, there must exist a "last entrant".

Let T be the last time any firm enters the market. At T, the last entrant's net present value must equal 0. If not, from (3), an entrepreneur who arrives just after T would also find it profitable to enter. Since no new firms enter after T, the last entrant experiences either a constant inflow of new customers (if no firms exit) or an increasing inflow (if there is exit). A direct consequence is that the last entrant's customer base and net present value increase over time and, hence, it never exits. But then also no other firms exit after T:[6] the mass of firms is constant after T.

We conclude that for $t \geq T$, $G(t) = G(T)$, and the last entrant has a constant inflow of new customers $N/G(T)$. In turn, this implies that its customer base at $t \geq T$ is given by,

$$m(t, T) = \int_T^t \frac{N}{G(T)} e^{-\alpha(t-s)} \mathrm{d}s = \frac{N}{G(T)} \cdot \frac{1 - e^{-\alpha(t-T)}}{\alpha},$$

[6]Indeed, the customer base and hence, flow profits, of firms that entered before T are always larger than that of the last entrant for any given time.

and its net present value at time of entry T equal:

$$\begin{aligned} V^e(T) &= \int_T^{\infty} e^{-\rho(t-T)} \left(m(t,T) - c\right) \, \mathrm{d}t \\ &= \frac{1}{\rho} \cdot \frac{1}{(\rho+\alpha)} \frac{N}{G(T)} - \frac{c}{\rho}. \end{aligned}$$

Since $1/(\rho+\alpha)$ is the life-time value of a new customer at the time of its arrival, the first term in the above expression is the life-time value of all future new customers, discounted to the time of entry. This life-time value depends both on the customer attrition rate α and the firm's discount factor ρ. The second term equals all future operational costs discounted to the time of entry.

As the net present value of the last entrant's entry $V^e(T)$ must be equal to 0, it follows that the steady state mass of firms is given by

$$G(T) = \frac{N}{(\rho+\alpha)\cdot c}.$$

For $t < T$, the mass of firms is given by $G(t) < G(T)$. Consequently, the entrepreneurs who enter before T will obtain more customers, resulting in a net present value that strictly exceeds 0. Every entrepreneur who has the chance to enter the market will then do so, leading to an entry rate of μ. It follows that the last entrant enters at

$$T = \frac{G(T) - G_0}{\mu}$$

The following proposition characterizes the equilibrium path of $G(t)$:

Proposition 1 *Dynamic Barriers to Entry There is a unique industry equilibrium: Entrepreneurs enter at rate μ until $t = T$ and the mass of firms in the market equals*

$$G(T) = \frac{N}{(\rho+\alpha)c}, \tag{5}$$

There is no exit; hence $G(T)$ is also the steady-state market size.

Even though entry is strictly profitable prior to T and even though all firms will be making strictly positive flow profits in the long run, there is no entry after T. Intuitively, given that firms incur a fixed operating cost c regardless of their customer base, new entrepreneurs only enter if they can build a sufficiently large customer base, sufficiently fast. This is increasingly difficult as the market becomes more crowded and new customers need to be shared among a larger mass of firms. As Proposition 1 shows, when $G(t) = G(T)$, the market is "full" and there is no more entry.

The steady-state market size has a straightforward interpretation: It is directly proportional to the life-time value of all new customers that arrive in the market, $N/(\rho+\alpha)$, and inversely proportional to fixed operational costs (c). Thus, for a given fixed operational cost c, customer arrival rate N and attrition rate α, the total number of firms increases as firms are more patient. The latter finding arises because firms grow their customer base gradually. New entrants start with few customers and operate at a loss, only becoming profitable as their customer base expands over time. Since profits are backloaded, more patient firms value these delayed returns more highly, leading to a larger equilibrium number of firms in the market.

3.2 Temporary Performance Differences and Long-run Profits

Next, we compare the value of firms across cohorts. To do so, it is instructive to first decompose the value of a firm as the value of its customer base plus the value of entry. Denoting by $V(t;t_e)$ the continuation value at time t of a firm that entered at time t_e, we have that:

$$V(t;t_e) = \frac{m(t;t_e)}{\rho+\alpha} + \int_t^{\infty} e^{-\rho(t'-t)} \left(\frac{1}{\rho+\alpha} \cdot \frac{N}{G(t')} \right) \mathrm{d}t' - \frac{c}{\rho}, \tag{6}$$

where the first term is the value of its current customer base $m_i(t;t_e)$, the second term is the value of all its future customers, who will arive at a rate $N/G(t')$ at time $t' > t$, and the third term is the total discounted cost of operating in the market.[7] In turn, the last two terms equal exactly the value of entry at time t, $V^e(t)$. It follows that at time t, the value of a firm who entered a time t_e is given by

$$V(t;t_e) = \frac{m(t;t_e)}{\rho+\alpha} + V^e(t) \tag{7}$$

where $V^e(t) = 0$ for $t \geq T$.

Consider now two entrepreneurs who entered at time t_e and $t'_e > t_e$. At any time $t \geq t'_e$, the earlier entrant has a larger customer base (and thus higher profits/continuation value)

[7] The second term calculates the lifetime value of customers acquired at each instant $t' \geq t$ and discounts it back to the current time t. Alternatively, consistent with the calculation of the value of entry in equation (3), we can calculate this term using the firm's sales to new customers after t at each instant t', i.e., $\int_t^{\infty} e^{-\rho(t'-t)} \left[\int_t^{t'} e^{-\alpha(t'-s)} \frac{N}{G(s)} \mathrm{d}s\right] \mathrm{d}t'$. Switching the order of integration (Fubini's theorem), we can verify that the two integrals are identical.

than the later entrant. Indeed, based on the expression in (2) of the customer base at time t of a firm that entered at t_e, we have:

$$m(t; t_e) = m(t; t'_e) + m(t'_e; t_e)e^{-\alpha(t-t'_e)} > m(t; t'_e). \tag{8}$$

The early entrant has an advantage over the late entrant because the former has accumulated customer capital from t_e to t'_e, i.e., $m(t'_e; t_e) > 0$. The extra customer capital depreciates at a rate α. By time t, it shrinks to $m(t'_e; t_e)e^{-\alpha(t-t'_e)}$ but is still positive. Therefore, the early entrant always has higher customer capital. It is also clear that this difference is decreasing in α and converges to zero when $t \to \infty$. In the long run, as t goes to infinity, the customer base of all firms converges to

$$m_{lr} = \int_0^\infty \frac{N}{G(T)} e^{-\alpha s} \mathrm{d}s = \frac{1}{\alpha} \cdot \frac{N}{G(T)} = (1 + \frac{\rho}{\alpha}) \cdot c,$$

and flow profits converge to

$$\pi_{lr} = m_{lr} - c = \frac{\rho \cdot c}{\alpha}.$$

Proposition 2 (Temporary Performance Differences and Long-run Profits) *Consider two entrepreneurs who entered at time t_e and t'_e ($t_e < t'_e$). At any time $t \geq t'_e$, the earlier entrant has a larger customer base and profits than the later entrant. In the long run, when $t \to \infty$, this difference disappears and the profits of all firms converge to $\pi_{lr} = \frac{\rho \cdot c}{\alpha}$, and their continuation value to c/α.*

Proposition 2 shows that early entrants have a competitive advantage that slowly disappears. Firms that enter earlier have accumulated a positive customer base when late entrants enter. However, this advantage erodes over time: customers acquired earlier will eventually leave. The profits of all firms, therefore, converge to the same level in the long run.

The formula for long-run profits in Proposition 2 highlights three necessary and complementary ingredients for firms to be strictly profitable in the long run: scale economies (fixed operating costs $c > 0$), a positive level of customer inertia ($\alpha < \infty$), and impatience ($\rho > 0$).

Higher fixed operating costs c result in higher long-term profits as they reduce entry and therefore result in a higher long-run customer base for firms that do enter. The formula further indicates that long-run profits are higher when firms are less patient. This is due to the backloaded nature of profits, which necessitates a higher long-run profit to compensate less patient entrants for their incurred costs. As firms become infinitely patient, i.e. $\rho \to 0$, entry continues until $m_{lr} = c$ and long-run profits are 0. Note that impatience does not matter as far as the continuation value of a firm is concerned.

Finally, while more repeat customers (a lower attrition α) increase long-run profitability, more new customers (a higher N) do not: π_{lr} is independent of N. Intuitively, new customers do not provide an advantage to incumbents – they are equally distributed among all firms in the market. As a result, they simply attract more entrants and do not increase long-term profits. In contrast, more repeat customers do favor incumbents because they already have a customer base and new entrants not yet. Incumbents therefore benefit more from a reduction in customer attrition than new entrants.

4 Industry Equilibrium with Endogenous Fixed Costs

We now consider a more general scenario in which companies can make a one-time investment in product quality. By making an investment $I > 0$ and upgrading its product, we assume a firm can reduce its customer attrition rate from $\alpha = \alpha_L$ to $\alpha = \alpha_H < \alpha_L$. The subscripts L and H refer to a low quality (L-type) and a high quality (H-type) firm. The random utility model in Appendix B, discussed in Section 4.3, provides a micro-foundation.

4.1 Investment in Quality and Persistent Performance Differences

What determines whether a firm invests in quality? And at what time is it optimal to do so? The following lemma shows that if a firm ever upgrades, it always does so when its customer base first reaches a critical size.

Lemma 1 *If a firm ever invests in quality, it happens when its customer base $m(t; t_e)$ first reaches a cutoff*

$$m^* \equiv \frac{\rho(\rho + \alpha_H)}{\alpha_L - \alpha_H} I. \tag{9}$$

Lemma 1 shows that the customer base is a sufficient statistic for the timing of upgrade. In particular, the timing does not depend on how many other firms are in the market. This might seem counterintuitive, as the number of competitors affects future customer acquisition and should theoretically influence upgrade decisions. Indeed, the decision of whether to upgrade or not does depend on the number of firms, as we will discuss in detail below, but the timing does not.

The reason is that because the firm chooses the timing to maximize its profit, the marginal benefit of updating earlier must equal its marginal cost. The marginal benefit of earlier upgrading comes from increasing the retention of existing customers–customers the firm already has. Since the magnitude of this benefit is completely determined by the size of the existing customer base, competitive factors that affect future customer acquisition are irrelevant to the timing calculation.

To calculate the cut-off level m^*, we use a perturbation argument that compares the firm's profits between two adjacent investment times, t_u and $t_u + \Delta t$, for some small Δt. By investing earlier at t_u, the firm benefits by having higher customer retention in the interval $[t_u, t_u + \Delta t]$. There are $m(t_u; t_e)$ such customers, and the total value of retaining them is

$$(\alpha_L - \alpha_H) \cdot \frac{m(t_u; t_e)}{(\rho + \alpha_H)} \cdot \Delta t,$$

where recall that $1/(\rho + \alpha_H)$ is the life-time value of a customer with attrition rate α_H. The cost of investing early is $\rho I \Delta t$. It then follows that there is gain in upgrading earlier as long as

$$(\alpha_L - \alpha_H) \cdot \frac{m(t_u; t_e)}{(\rho + \alpha_H)} \geq \rho I. \tag{10}$$

Setting (10) to an equality, we obtain the expression for the cut-off customer base at which the firm should upgrade (if it ever does so).

The formula for m^* yields several intuitive comparative static results. When the reduction in the attrition rate α is larger or the investment cost I is lower, firms require a smaller customer base to justify the upgrade. The discount rate ρ also plays a role: a lower discount rate reduces the required customer base because it increases the present value of keeping a customer.

As a final observation, note that Lemma 1 states that the upgrade takes place when the customer base first reaches the cutoff. It is possible that the cutoff is reached twice. This occurs when customer numbers follow a hump-shaped pattern over time—a pattern typical for early market entrants. These firms initially attract many customers due to limited competition, but their customer base may later decline when high customer outflow (driven by their large customer base) exceeds the diminishing inflow (caused by increasing market competition).

The following proposition shows how, in equilibrium, endogenous investments in quality can create asymmetries between early movers who invest (upgrade), and late movers who do not find it worthwhile to do so:

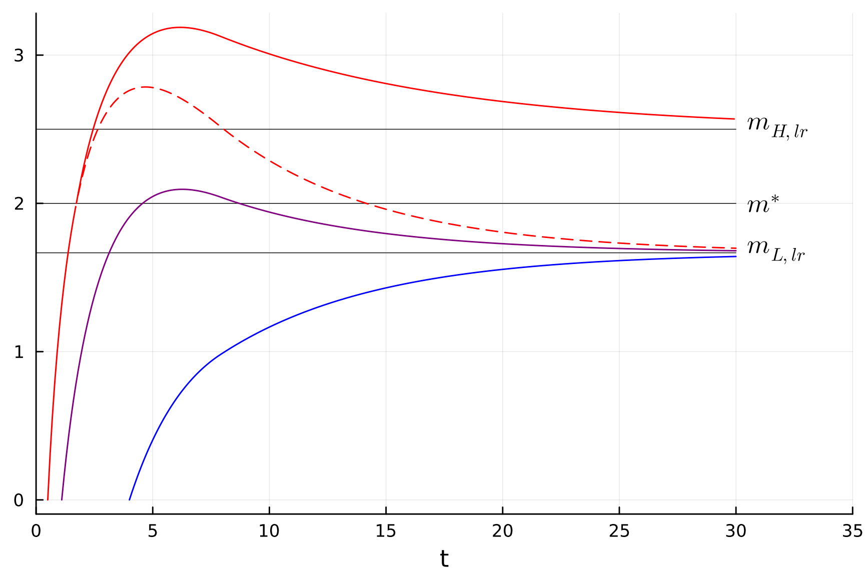
Proposition 3 (Investment and Persistent Performance Differences) *In equilibrium,* *firms enter at a constant rate μ until T and there is no exit. When G_0 is sufficiently small, there exist two cutoff investment costs $\underline{I} \leq \bar{I}$ such that*

1. *If $I < \underline{I}$, all entrants upgrade.*
2. *If $I > \bar{I}$, no entrants upgrade.*
3. *If $I \in [\underline{I}, \bar{I}]$, there exists a cutoff $T_u \in [0, T)$, such that firms that enter at $t \in [0, T_u]$ upgrade and later entrants do not.*

When firms upgrade, they do so at the earliest time t for which $m(t, t_e) = m^*$. When $I \geq \underline{I}$, firms enter until the total mass of firms equals $G(T) = \frac{N}{(\rho+\alpha_L)c}$

Proposition 3 shows how only early entrants choose to upgrade for intermediate investment costs. Intuitively, an earlier entry enables firms to build a larger customer base, making upgrading more valuable. Later entrants, despite having access to the same technology and financial resources, rationally choose not to upgrade because their smaller customer base makes the fixed investment cost unprofitable.

Figure 1: Example paths of customer capital in the case of partial upgrading



Notes. This figure plots the customer base $m(t; t_e)$ for three firms entering at different time t_e. The earliest entrant (red curve) upgrades, while the second and third entrants (purple and blue curves) do not. m^* denotes the threshold for upgrading in the necessary condition. $m_{H,lr}$ and $m_{L,lr}$ denote the long-run firm sizes for H- and L-type firms. Parameters: $\rho = 0.1, N = 1, c = 1, \alpha_H = 0.1, \alpha_L = 0.15, \mu = 0.5, I = 5, G(0) = 0.05.$

Figure 1 depicts customer dynamics for three firms entering the market at different times. The red curve represents a firm entering before T_u. Upon reaching a customer base of m^*, this firm upgrades, achieving a larger customer base $m_{H,lr}$ than its non-upgrading counterfactual (shown by the dotted line). The other two curves represent firms entering after T_u. These firms do not upgrade, and their customer sizes eventually converge to $m_{L,lr}$, where

$$m_{L,lr} = \frac{N}{G(T)\alpha_L} < \frac{N}{G(T)\alpha_H} = m_{H,lr}.$$

Notably, even when the customer size of one of these firms reaches m^*, it doesn't upgrade, because the net present value of doing so is negative. This illustrates how the upgrade timing and the upgrade decision are distinct considerations: the decision of whether to upgrade depends also on the competitive conditions the firm faces, which depend on the time it enters the market.

A direct implication from Proposition 3 is that incumbents who invest in quality (early entrants) have a persistent performance difference over late movers who do not: they are more profitable not only in the short-run but also in the long-run, and this for three complementary reasons. First, incumbents that upgrade retain a larger long-term customer base than late movers who do not upgrade, thanks to lower past attrition: $m_{H,lr} > m_{L,lr}$. Second, the lifetime value of this customer base is higher because those customers are less likely to churn ($\alpha_H < \alpha_L$). Finally, a superior product lowers future attrition, boosting the expected value of yet-to-be-acquired customers. Consequently, as we show next, an upgraded incumbent's continuation value includes not only the lifetime value of its current customers plus the value of entry (as for a firm who does not upgrade), but also the extra surplus it can extract from future customers.

Formally, compare an incumbent who entered at time t_e and upgraded at time t^u with a late mover who entered at time $t'_e > t_e$ and will never upgrade. From (7), the continuation value of the non-upgraded late mover at time $t > t^u$ can again be decomposed as the value of its customer base plus the value of entry:

$$V(t; t'_e) = \frac{m(t; t'_e)}{\rho + \alpha_L} + V^e(t), \tag{11}$$

In contrast, the continuation value of the upgraded incumbent can now be decomposed as

$$V(t; t_e) = \frac{m(t; t_e)}{\rho + \alpha_H} + \int_t^\infty e^{-\rho(t'-t)} \left(\frac{1}{\rho + \alpha_H} \cdot \frac{N}{G(t')} \right) \mathrm{d}t' - \frac{c}{\rho}.$$

As in (7), the first term is still the value of the firm's customer base. Unlike in (7), however, the last two terms do not equal the value of entry at time $t > t^u$, as late entrants now have a higher customer attrition rate $\alpha_L > \alpha_H$. In particular, we now have that

$$V(t;t_e) = \frac{m(t;t_e)}{\rho + \alpha_H} + V^e(t) + \int_t^\infty e^{-\rho(t'-t)} \left(\frac{\alpha_L - \alpha_H}{(\rho + \alpha_H)(\rho + \alpha_L)} \cdot \frac{N}{G(t')} \right) \mathrm{d}t', \tag{12}$$

The last term is new and represents the added value that new customers generate to the incumbent compared to late entrants. Due to its investment in upgrading the quality of its product, the incumbent retains customers at a higher rate than new entrants at time t, increasing their life-time value.

Our theory illuminates why early movers with established customer bases can sustain competitive advantages through strategic quality investments, as exemplified by John Deere's dominance in agricultural equipment. With a customer base built over 150 years—including multigenerational farming families exhibiting extraordinary loyalty—Deere has sufficient scale to justify massive R&D investments in precision agriculture. The company pours billions into GPS-guided automation, AI-enabled sprayers, and integrated data platforms that can exceed \$100,000 per tractor. These investments yield high returns precisely because Deere's vast installed base ensures widespread adoption: existing Deere owners readily purchase compatible upgrades, while enhanced productivity further reduces churn. Smaller competitors like AGCO cannot match this investment intensity—their thinner customer bases mean identical R&D expenditures generate far lower returns.

Similarly, Apple demonstrates how massive scale transforms customer capital into permanent technological advantages through component innovation. With over 220 million iPhones shipped annually, Apple amortizes enormous fixed costs across its base, justifying investments no competitor can match: \$450 million for Ceramic Shield glass with Corning, exclusive access to TSMC's first-year 3nm production, and \$3 billion in micro-LED development. These investments generate exceptional returns not simply because Apple can afford them, but because each enhancement—from 5× tetraprism zoom lenses to custom Photonic Engine processing—increases retention rates already exceeding 90%. When Apple develops exclusive features, it captures value across hundreds of millions of users for years, while Android manufacturers with lower retention and smaller individual bases cannot justify similar investments. This scale advantage—where higher retention rates and larger customer bases enable increasingly ambitious quality upgrades—perpetuates Apple's technological leadership and strengthens its dynamic competitive position over time.

4.2 Competition and Investments in Quality

In this section, we examine how competition affects firms' incentives to upgrade product quality–a key policy consideration for understanding when market concentration may benefit or harm consumers. In our model, more competition can be due to a lower fixed cost c and thus lower barriers to entry (a larger T) or due to an increase in the rate μ of new firms entering the market at any time $t < T$.

To analyze these effects, we examine the investment threshold $\underline{I}$ below which all entrants upgrade. This threshold serves as a measure for the firms' incentive to upgrade. The higher is $\underline{I}$,the more likely are firms to upgrade. As shown in Appendix, this threshold has a closed-form expression given by

$$\underline{I} = \frac{(\alpha_L - \alpha_H)(\rho + \alpha_L)c}{\rho\alpha_L(\rho + \alpha_H)}. \tag{13}$$

Not surprisingly, this investment threshold is decreasing in ρ, as more impatient firms are less likely to invest. Notably, this threshold is also increasing in the fixed operating cost c. The reason for this is that as barriers-to-entry increase, firms anticipate less future competition and a larger future customer base. This makes upgrading more valuable.

Let us denote by $G_H(t)$ the mass of H-type firms at time t. Define $g_H(t) \equiv G_H(t)/G(t)$, the fraction of firms that have upgraded at any point of time t. Note that $g_H(t)$ is a proxy of the consumer welfare. At any time t, new customers are randomly matched to firms in the market. According to the random utility model in Appendix B, consumers must enjoy higher value when matched to a H-type firm. Consumer welfare, measured by the total number of customers served at each instant, then increases with $g_H(t)$.

The next proposition studies the impact of competition on $g_H(t)$, the share of firms that has upgraded at time t.

Proposition 4 (Competition and Investment) .

1. *A higher market entry rate μ or a lower flow fixed cost c reduces the share of H-type firms in the steady state. The decrease is strict when $I \geq \underline{I}$.*

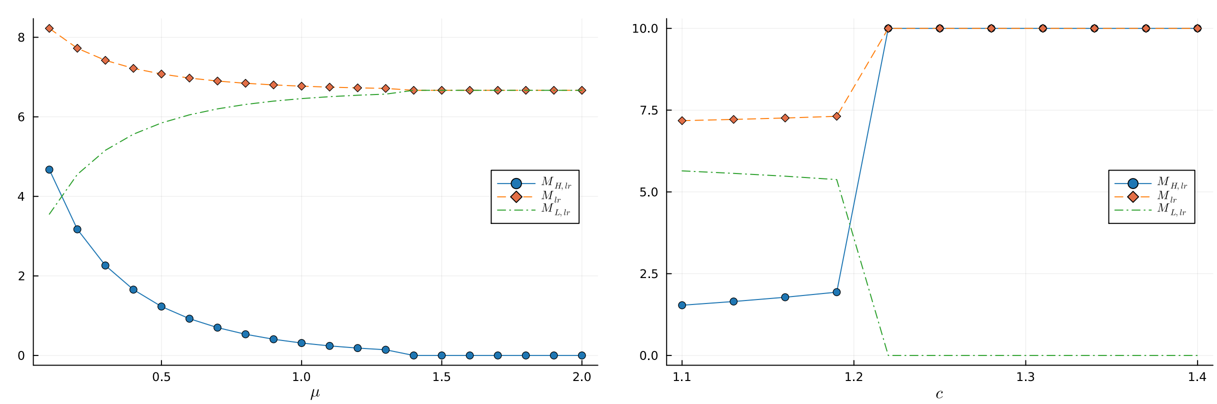
2. *Assume $I \geq \underline{I}$. The share of H-type firms at any time t, $g_H(t)$, decreases in μ and increases in c. The decrease/increase is strict for sufficiently large t.*

Proposition 4 demonstrates how intensified competition can harm investment and consumer welfare. The core mechanism is that excessive competition prevents firms from achieving the scale needed to justify technology upgrades. This can trap markets in equilibria characterized by many small firms offering low-quality products.

Our model identifies two distinct channels through which competition affects upgrade incentives. The first one is a more traditional competition effect, operated through the fixed cost c, and reflects the degree of competition in the long run. When the fixed costs decrease, more firms ultimately enter the market. This lower each firm's customer base in the long-run, reducing the returns to upgrading. This effect operates through the equilibrium number of firms—a standard result in industrial organization where lower barriers to entry lead to more fragmented markets.

The second channel operates through the speed of entry (μ). This parameter affects how quickly firms enter, not how many enter. When μ is low, early entrants enjoy a longer transient advantage: they have longer time to build substantial customer bases before facing full competition, making upgrades profitable. However, when μ is high and all firms enter nearly simultaneously, no firm gets this "first-mover" opportunity to grow large. Even though the same number of firms eventually enter regardless of μ (when $I > \underline{I}$), rapid entry eliminates the transient market power that makes upgrading worthwhile. This speed-of-entry effect represents a new insight: competition policies must consider not just how many firms compete, but how quickly competition materializes.

Figure 2: The impact of entry rate μ and flow fixed cost c on the long-run market sizes



(a) Market Size and μ (b) Market Size and c

Notes. The two panels plot the number of customers served in the long run under different values of μ and c. Each dot represents a long-run equilibrium. M_{lr} represents the total number of customers served, and $M_{i,lr}, i = H, L$ represents the number of customers served by the H- and L-type firms, respectively. Baseline parameter values: $\rho = 0.1, N = 1, \mu = 0.5, c = 1, \alpha_H = 0.1, \alpha_L = 0.15, I = 5, G(0) = 0.05$. The baseline values of μ and c are irrelevant when the corresponding parameters are varied in the comparative statics.

The numerical examples in Panels (a) and (b) of Figure 2 illustrate the impact of μ and c on the total number of customers, respectively. As shown in the proof of Proposition 4, the total number of customers served is increasing in the steady-state fraction of H-type firms, because customer attrition is lower when customers are served by H-type firms. With stronger competition for customers either during the transition (higher μ) or after the steady state (lower c), the fraction of H-type firms declines, and the total number of customers served drops.

4.3 Microfoundations and Extensions

4.3.1 Consumer Choice

In Appendix B, we outline a simple continuous-time dynamic discrete choice model of consumers following Arcidiacono et al. (2016). It provides a microfoundation of the customer attrition rates, α_H and α_L, and a measure of consumer welfare. In particular, after customers are matched with firms, they have the opportunity to choose whether to remain with their current firm or opt for the outside option (leaving the market) permanently. A Poisson process governs the arrival of these opportunities. The values of staying or leaving are affected by choice-specific payoff shocks. The customer attrition rate decreases with the flow utility of consuming the current firm's product and increases with the Poisson arrival rate. If we assume that the H-type firm provides a higher-quality product and consumers obtain higher flow utility, we must have $\alpha_H < \alpha_L$. We can also show that the value of being matched with H-type firms is higher than that of being matched with L-type firms. For a new consumer, the probability of being matched with H-type firms equals the current fraction of H-type firms, $g_H(t)$. Therefore, consumers have a higher expected value when $g_H(t)$ is higher.

4.3.2 Pricing

In our baseline model, we assume that firms charge a constant price (normalized to one) and abstract away from optimal pricing. The consumer choice model allows us to discuss pricing. We assume that firms cannot credibly commit to future prices as in many earlier works on consumer lock-in (Farrell and Klemperer, 2007). The literature also assumes that firms cannot commit to future quality, and we make a slightly different assumption: consumers do not expect quality upgrades. With these assumptions, firms choose an optimal price to maximize the value of a matched customer. They face a trade-off between extracting higher profits from the customer now and a higher chance of losing him/her later. Under certain parameter restrictions, a unique optimal price exists for H-type (L-type) firms, denoted as p_H (p_L), with $p_H > p_L$. Despite the higher price, the value of being matched with an H-type firm remains higher than that of an L-type firm, and consumer welfare continues to increase strictly with $g_H(t)$.

When H-type firms charge a higher price than L-type firms, upgrading brings additional benefits to the firm due to higher profits from existing customers. In Online Appendix OA.2, we show that all our theoretical results are qualitatively unchanged in an extension with two arbitrary prices $p_H > p_L$. For example, firms upgrade only when their customer base reaches a critical size, with both prices affecting this threshold. Intuitively, other than slower depreciation of the current customer capital, upgrading earlier also brings more sales, proportional to $(p_H - p_L)m(t; t_e)$. We can obtain a similar cutoff as in Lemma 1.

4.3.3 Discerning Customers

In the baseline model, we have assumed that the only benefit of upgrading is a lower customer attrition rate. In contrast, we assumed that the customer arrival rate is the same between H- and L-type firms. In Online Appendix OA.3, we present an extension in which H-type firms may have a higher customer arrival rate. This gives firms more incentives to upgrade. More firms upgrade, and they upgrade earlier compared to the baseline case here. We show that under certain parameter restrictions, we can still prove that less competition due to lower entry rates μ and/or higher fixed costs c leads to a larger share of H-type firms in the steady state.

5 Entry Choice with Multiple Markets.

We now apply our framework to a fundamental question in corporate strategy: which markets should firms enter? Porter's (1979) Five Forces framework has long guided firms in assessing market attractiveness.[8] The basic insight of Porter was that the competitive intensity of an industry, such as current rivalry and the threat of future entrants, determines its profitability. Consistent with this perspective, Section 3 has shown that long-run profits of firms are increasing in fixed operational costs, as shown by $\pi_{lr} = \frac{\rho}{\alpha} \cdot c$. While higher fixed costs reduce profitability for a given customer base, they also act as a barrier-to-entry, reducing the number of future entrants, a key force in Porter's framework.

[8] According to HBS's Institute for Strategy & Competitiveness, "A Five Forces analysis can help companies assess which industries to compete in and how to position themselves for success." (see `https://www.isc.hbs.edu/strategy/business-strategy/Pages/the-five-forces.aspx`)

This long-run perspective suggests that markets with higher fixed costs are more attractive, at least for firms that enter sufficiently early. But while high barriers promise greater eventual profitability, they also impose immediate costs that reduce short-run profits. Hence the paradox of entry barriers.

In what follows, we formally analyze the impact of barriers-to-entry, in the form of high (exogenous) fixed costs, on the attractiveness of entry. To do so, we consider a setting with two markets, 1 or 2, where $c_1 < c_2$ but all other features are identical. We ask which market is more attractive to an entrepreneur arriving at some time t. Is it market 1 with lower fixed operating costs or market 2 with higher barriers-to-entry and less future competition? We also ask how this choice depends on the timing of entry, and whether there are also endogenous fixed costs (opportunities to make investments to upgrade product quality).

5.1 Market choice with (only) exogenous fixed costs.

We start with analyzing the case where upgrading is not possible (as in Section 3) or equivalently, $I > \bar{I}$. We denote the customer attrition rate by α_L. Note that in this case, firms only differ in their timing of entry. There is no quality difference between firms.

Consider two markets, 1 and 2, with $c_1 < c_2$ so that there is entry over a longer period in market 1: $T_1 > T_2$, where T_i denotes the arrival time of the last entrant in market i. All other market characteristics are identical. To build intuition, we first consider the market choice of a single entrepreneur who arrives at some time t. Because of capacity constraints (e.g. limited attention or financial resources), the entrepreneur can enter at most one market. The arrival rate of other entrepreneurs is μ in each market: they only have expertise in one specific market. In Section 5.3, we generalize our insights to the case where all entrepreneurs optimize which market to enter.

Recall from (7) that the continuation value at time t of a firm that entered market $i \in \{1, 2\}$ at time t_e is given by the value of its customer base at time t plus the value of entry in market i at time t:

$$V_i(t; t_e) = \frac{m_i(t; t_e)}{\rho + \alpha_L} + V_i^e(t) \tag{14}$$

Note first that once there is no more entry in either market, $t > T_1$, the continuation value of a firm in market 2 with higher fixed costs is always higher than that of a firm in market 1. Indeed, since there is no more entry in market 2 for $t \in [T_2, T_1]$, while there is in market 1, firms in market 2 will then have accumulated a larger customer base: $m_2(t; t_e) > m_1(t; t_e)$ for $t > T_1$. Since the value of entry is 0 in both markets, it then follows immediately from

(14) that $V_2(t;t_e) > V_1(t,t_e)$: firms in market 2 are more valuable going forward.

But we are interested in the value of entry, not in the continuation value of firms in the mature phase of an industry. Consider therefore the value of entry at time $t_e < T_2$ when entry has a positive value in both markets.

Observe first that at any time $t \in [t_e, T_2]$, the firm has the same inflow of new customers in each market, but incurs lower fixed costs in market 1. Indeed, as long as $t < T_2$, both markets experience the same rate of entry and, hence, competition. Next, observe that at any time $t \leq T_2$, we have that

$$m_1(t;t_e) = m_2(t;t_e) = m(t;t_e)$$

since both markets have the same rate of entry up to T_2. But since there is no more entry in market 2 after T_2, the value of entry must be zero at T_2: $V_2^e(T_2) = 0$. It then follows directly from (14) that at T_2, the continuation value is larger in market 1 than market 2:

$$V_1(T_2;t_e) = \frac{m(T_2;t_e)}{\rho + \alpha_L} + V_1^e(T_2) > \frac{m(T_2;t_e)}{\rho + \alpha_L} = V_2(T_2;t_e).$$

In sum, as long as there is entry in both markets, $t < T_2$, the firm accumulates the same customer base in both markets, but incurs lower operating costs in market 1: $c_1 < c_2$. Once entry stops in the high-cost market, at T_2, the continuation value equals the value of the customer base at that time, identical in both markets, plus the value of entry, which is 0 in the high-cost market. Both factors make entry more attractive in the low-cost market. We obtain the following result:

Proposition 5 *Assume no upgrades are possible. Whenever entry occurs in at least one market, the Entrepreneur strictly prefers to enter market 1 with low fixed costs.*

Intuitively, higher fixed costs reduce future competition and hence increase the inflow of future customers. The zero-profit condition for entry, however, ensures that this higher customer inflow is either exactly offset by higher operating costs (when there is no more entry) or more than offset (when there is only entry in the low-cost market).[9]

[9] Appendix A provides a more succinct but less instructive proof by taking the derivative of the value of entry with respect to c.

5.2 Market entry choice with endogenous fixed costs

Consider now the possibility for firms to make fixed investments to upgrade their product quality. As shown in Section 4, this may create permanent performance differences between early entrants, whose large customer base gives them an incentive to make such investments, and late entrants, who are smaller and do not find this profitable.

To develop some intuitions, we consider the market entry choice of an early entrant who upgrades at time $t_i^u < T_2$ in market i when there is still entry in both markets. This implies that this firm will upgrade at the same time in both markets, $t_1^u = t_2^u = t^u$. Indeed, the timing of upgrading only depends on $m_i(t; t_e)$ and $m_1(t; t_e) = m_2(t; t_e)$ for $t \leq T_2$. Assume further that late entrants, who enter after t^u, never upgrade in either market. While our main result does not rely on these two restrictions, this simplifies the exposition.

As shown in Section 4, the upgrader's continuation value in market i at $t > t^u$ can be decomposed as

$$V_i(t; t_e) = \frac{m_i(t; t_e)}{\rho + \alpha_H} + V_i^e(t) + \int_t^\infty e^{-\rho(t'-t)} \left(\frac{\alpha_L - \alpha_H}{(\rho + \alpha_H)(\rho + \alpha_L)} \cdot \frac{N}{G_i(t')} \right) \mathrm{d}t', \tag{15}$$

where the first term is the value of the firm's customer base, with $m_1(t; t_e) = m_2(t, t_e)$ for $t \leq T_2$, and the second term is the value of entry in market i at time t. The third term, which is absent in a setting without investments, is the extra value that future customers will generate to the upgrader compared to what they yield to late entrants who do not upgrade and therefore have a lower retention rate.[10]

Importantly, the third term in the upgrader's continuation value (15) is larger when it has entered the high-cost market 2, as it then has a higher inflow of new customers for $t > T_2$ (when there is no more entry in market 2):

$$N/G_2(t') > N/G_1(t'). \tag{16}$$

In the absence of upgrading, this higher customer inflow in market 2 was offset by higher operating costs because of the no-entry condition. Because of upgrading, however, the early entrant can now better monetize new customers than a late entrant who does not upgrade.

This creates a trade-off between entering market 1 versus 2: lower fixed costs in the short run, but a higher inflow of new customers, which more than outweighs the higher fixed costs

[10] In our extension with endogenous prices, upgrading also allows the early entrant to charge higher prices. For simplicity, we keep prices exogenous in this section, but this would further increase the additional value future customers yield to early entrants.

needed to serve them in the long run.

The following proposition shows more generally, that for high μ, an entrepreneur who enters early, will prefer to enter the high-cost market as that market experiences less entry and hence will deliver a higher customer flow. Because of his investment in product quality, the entrepreneur will value those customers more than new entrants who never upgrade. Indeed, an entrepreneur who never upgrades always prefers to enter the low-cost market:[11]

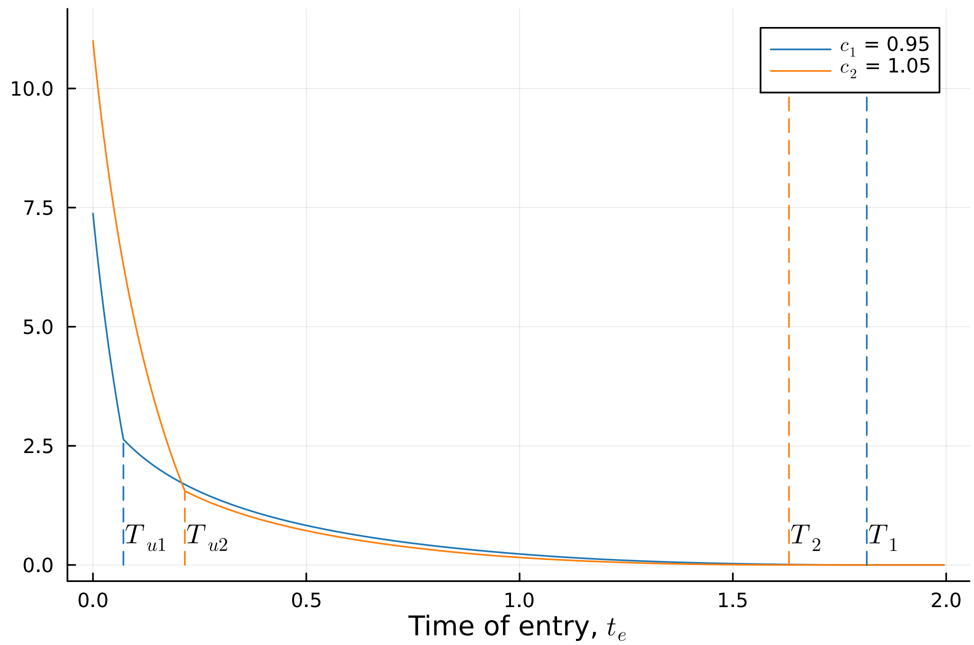
Proposition 6 *Assume $I \geq \underline{I}$ so that the last entrant does not upgrade. For sufficiently large μ and small $G(0)$, an early entrant (when t is sufficiently small) prefers to enter market 2 with high barriers-to-entry (and upgrades his product when his customer base reaches the critical threshold). A late entrant who never upgrades, prefers to enter market 1 with low barriers-to-entry.*

A large μ does not affect the steady-state mass of firms that enter markets 1 and 2, nor does it affect long-term profits. On the other hand, it shortens the time to reach market saturation. It therefore reduces the short-run cost benefits of market 1 and amplifies the long-term revenue advantage of market 2 as future profits are less discounted. Therefore, a sufficiently high μ makes market 2 more attractive for early entrants.[12] A small initial mass of entrants $G(0)$, finally, ensures that it will be optimal for an early entrant to upgrade.

[11] While we state the result for $I \geq \underline{I}$, entry in the high-cost market may be more attractive even when $I < \underline{I}$, and all entrepreneurs eventually upgrade. Intuitively, what matters for our result is that new entrants do not immediately upgrade. Hence, their value of new customers will be smaller than that of early entrants who upgrade before them, and are better at monetizing/retaining new customers.

[12] We provide a lower bound of μ for early entrants to choose market 2 over market 1 in the proof of Proposition 6. See Appendix A.

Figure 3: Value of entry in two separate markets



Notes. We plot the value of entry as a function of the time of entry for two markets with different fixed costs, $c_1 = 0.95$ (blue lines) and $c_2 = 1.05$ (orange lines). T_j, $j \in \{1, 2\}$, indicates the time when entry stops, and T_{uj} indicates the last upgrader's time of entry. The other model parameters are: $\rho = 0.1, N = 1, \alpha_H = 0.1, \alpha_L = 1.0, \mu = 0.5, G(0) = 0.05, I = 55$.

Figure 3 illustrates Proposition 6 with a numerical example. We plot the value of entry as a function of time of entry in two markets that differ only in their fixed operating costs. Market 1 has a lower cost $c_1 = 0.95$, while market 2's cost is higher at $c_2 = 1.05$. The values of entry of the two markets are in orange and blue colors, respectively. Note that both curves have kinks, which coincide with the time of entry of the last upgrader. The possibility of upgrading changes the attractiveness of the two markets. It is clear that for entrants who are sufficiently close to T_1, the value of entry is higher in low-cost market 1, and an entrepreneur would prefer market 1 over market 2. This is reversed for early entrants – upgrading amplifies the benefit of higher entry barriers and less future competition, and it becomes more attractive to enter market 2.

5.3 Market Entry Choice: General equilibrium analysis

So far, we have assumed that only one entrepreneur has flexibility in his market choice. What if all entrepreneurs could choose?

Suppose again there are two markets with different entry barriers, $c_1 < c_2$. At $t = 0$, a mass $G(0)$ of entrepreneurs can enter either market. At any instant after $t = 0$, there are μ new entrepreneurs who can enter one of the two markets. The following results hold:

Proposition 7 1. *Exogenous fixed costs: If no firms upgrade in either market, there is a unique equilibrium in which entrepreneurs are indifferent between entering either market at any time and entry stops in both markets at the same time T. There are always strictly more firms in market 1 than in market 2, and the entry rate in market 1 is strictly higher before T.*

2. *Endogenous fixed costs: Given any parameter combinations $N, \rho, \alpha_L, \alpha_H$, $c_1 < c_2$ and $I > \underline{I}(\rho, \alpha_L, \alpha_H, c_2)$, entry stops in both markets at the same time $t = T$ and the equilibrium number of firms is higher in market 1 for sufficiently large t. However, when μ is sufficiently large and $G(0)$ is sufficiently small, the equilibrium number of firms is higher in market 2 at an earlier time.*

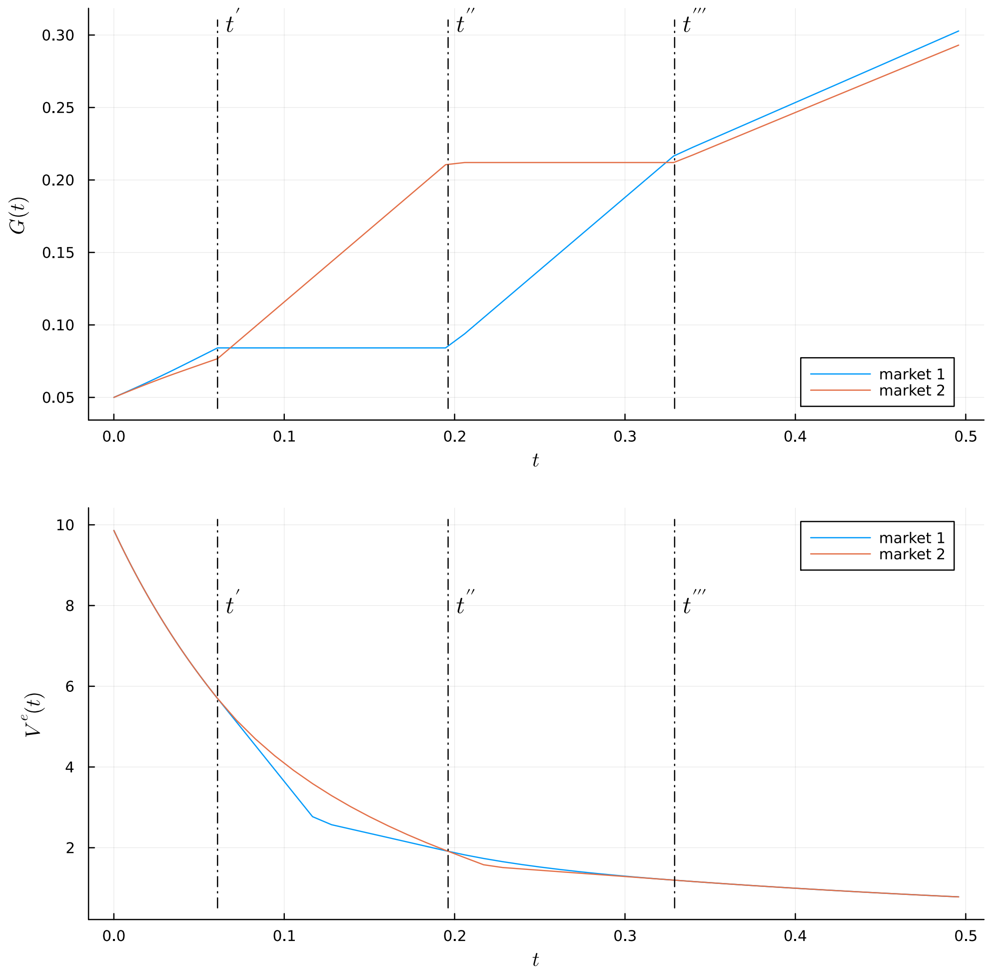
Proof. See Appendix Section A. ■

Our general equilibrium analysis generalizes the key findings of our partial equilibrium analysis in Sections 5.1 and 5.2, with some key differences.

With only exogenous scale economies, there is more entry in the low-cost market 1 at any time t. Both markets, however, yield the same NPV at entry. Intuitively, in equilibrium, at any time t, the excess entry in market 1 from time 0 to T and, hence, the lower future customer inflows in market 1, exactly offset its lower operating costs. Unlike in partial equilibrium, both markets also reach saturation at the same time $T_1 = T_2 = T$. To see the intuition for this, note that prior to T_i, when entry ends in market i, the value of entry must be strictly positive in market i. The value entry at $t = T_i$, however, must be 0. In turn, this implies that we must have $T_1 = T_2$.

With endogenous scale economies, dynamics can be quite rich, but some general results can be obtained. As with exogenous scale economies, over the full period in which entry occurs, there is cumulatively more entry in the low cost market: $G_1(T) > G_2(T)$. As in section 5.2, however, there is a sense in which the high-cost market is more attractive for early entrants who upgrade: given the same conditions, for some $t' < T$, there will have been more cumulative entry in the high cost market 2: $G_2(t') > G_2(t')$.

Figure 4: General equilibrium with entry in two markets: an example



Notes. This figure illustrates an equilibrium when entrepreneurs can choose to enter one of the two markets at any instant. Market 1 has lower fixed costs than market 2. The upper panel illustrates the equilibrium paths $G(t)$ in both markets. The lower panel plots the value of entry as a function of time. Parameter values: $c_1 = 0.95, c_2 = 1.05, \rho = 0.1, N = 1, \alpha_H = 0.1, \alpha_L = 0, I = 55, G_1(0) + G_2(0) = 0.1, \mu_1 + \mu_2 = 0.1$.

Figure 4 illustrates some of the rich dynamics that can occur with endogenous fixed costs. Even though we consider a general equilibrium setting with a continuum of entrepreneurs optimally choosing which market to enter, this example illustrates how it may be strictly optimal to enter the high-cost market early on, and subsequently strictly optimal to enter the low-cost market at a later time. Thus, unlike with exogenous fixed costs, entrants are not necessarily indifferent in equilibrium.

Concretely, in the equilibrium of Figure 4, there is initially entry in both markets: for $t < t'$, $V_1^e(t) = V_2^e(t)$. Note that this must be the case in any example, as otherwise profits would be unlimited in the market with no entry, a contradiction. After an initial period of simultaneous entry, however, there is subsequently only entry in the high-cost market 2: for

$t \in (t', t'')$, $V_1^e(t) < V_2^e(t)$. As more firms enter market 2, we eventually have again that $V_1^e(t) > V_2^e(t)$ at which point (at t'') it becomes strictly optimal to enter the low-cost market 1. Finally, for $t > t'''$, there is again entry in both markets. At this point, entrants never upgrade in any market, so this mirrors the setting with exogenous scale economies.

5.4 Patient Entrepreneurs

In this section, we consider a scenario in which entrepreneurs have different discount rates. Many factors can lead to such differences. For example, some entrepreneurs aim to build a long-lasting and successful business, prioritizing long-term profits over short-term gains; others may have more financial resources or better access to external financing. Should they enter the high-barrier market?

Consider a single 'patient entrepreneur' with a discount rate ρ who arrives at a time $t < T$, where T is decreasing in the fixed operating costs c. The discount rate of all future entrants is $\rho' \geq \rho$. The following proposition states that the patient entrepreneur may prefer markets with higher barriers-to-entry:

Proposition 8 (Entrepreneur is more patient) *Assume $I \geq \underline{I}$ so that a late entrant does not upgrade. The patient entrepreneur's value of entry in a market increases with the fixed operating cost c of that market, provided ρ is sufficiently small or ρ' is sufficiently large.*

To see this result formally, suppose the patient entrepreneur does not upgrade and entered at time t_e. Applying equation (7), the entrepreneur's continuation value at time T, when entry stops, is given by

$$V(T; t_e) = \frac{m(T; t_e)}{\rho + \alpha_L} + \int_T^\infty e^{-\rho(t'-T)} \left(\frac{1}{\rho + \alpha_L} \cdot \frac{N}{G(T)} \right) \mathrm{d}t' - \frac{c}{\rho}, \tag{17}$$

where the second term can be rewritten as:

$$\int_T^\infty e^{-\rho(t'-T)} \left(\frac{\rho' + \alpha_L}{\rho + \alpha_L} \cdot c \right) \mathrm{d}t' = \frac{\rho' + \alpha_L}{\rho + \alpha_L} \cdot \frac{c}{\rho}.$$

As one can see, the marginal benefit of a higher c can be arbitrarily large with $\rho' > \rho$. In contrast, with $\rho' = \rho$, the second and third terms in the value function cancel out. Intuitively, for a typical (impatient) entrant, when c increases, the value of higher future customer inflows (the second term in (17)) is exactly offset by the higher operating costs (the third term). Since a patient entrepreneur values future customers more than a typical entrepreneur, however, she then strictly benefits from an increase in c.

This result resonates with the discussions on capital requirements in Porter (1980). He suggests that large corporations with financial resources have an advantage in entering high-barrier industries, where such barriers limit the pool of other entrants. Here we model financial resources as a lower discount rate rather than the capability of paying out a large up-front cost. We find that entrepreneurs with a lower discount rate have an advantage in entering the high-barrier market only when their discount rates are lower than the prevailing ones.

5.5 Other factors that affect market choice

The purpose of this section was to formally study the paradox of entry barriers. As such, our analysis has focused on the case where market 1 has higher barriers (operating costs) than market 2, keeping all other parameters fixed. Markets may differ on many other dimensions, however. Not surprisingly, one can show that, ceteris paribus, an entrepreneur will find entry in market 1 more attractive than entry in market 2 when (i) demand is higher in market 1 ($N_1 > N_2$), (ii) the entry rate is lower in market 1 ($\mu_1 < \mu_2$), and/or (iii) the entrepreneur is an earlier entrant in market 1 ($t_{e,1} < t_{e,2}$).

6 Conclusion

This paper develops a theory of dynamic entry barriers and persistent performance differences arising from the interaction between customer inertia and scale economies. Exogenous scale economies create barriers to entry: New firms eventually stop entering markets even as incumbents remain profitable, because entrants cannot accumulate customers quickly enough to cover their fixed operational costs. Endogenous fixed costs create persistent performance differences: Early entrants with large customer bases invest in quality upgrades, while late movers, lacking such scale, cannot justify these investments. Ultimately, a firm's value is determined by the life-time value of its customer base (its customer capital which a firm slowly builds up over time), the value of entry (zero after an initial phase of entry) and the firm's ability to generate superior value from future customer inflows (which requires asymmetric investments in quality between early and late movers).

Our analysis resolves the paradox of entry barriers. High-cost markets have higher future customer inflows as they see less entry in the long run. Entrepreneurs generally prefer to enter low-cost markets, however, because the zero-profit condition ensures that higher future customer inflows in high-cost markets are more than offset by higher operating expenses. Endogenous investments in quality may change this calculus, as we show. Early entrants that can grow a customer base quickly enough to justify large investments in quality may rationally choose high-cost markets: since upgraded firms retain customers at higher rates, they can better monetize the larger stream of future customers in high-cost markets, making higher barriers-to-entry worthwhile. This explains why some entrepreneurs seek out difficult markets while others flee to easier ones—timing and upgrade strategy determine which barriers help versus hinder success.

Our insights may shed light on conglomerate expansion strategies. In many markets, large diversified firms have leveraged their existing customer base when entering new markets, effectively lowering customer acquisition costs relative to de novo entrants. Classic examples are Disney, Microsoft, Apple, Google, Netfix and Amazon– firms with very sticky customer bases that have expanded the range of products and services offered to their customers over time. Amazon, for example, is a gigantic e-commerce platform that started as an online book store. Its prime membership program generates tremendous customer loyalty and now also offers services such as video streaming, music streaming and at-home grocery delivery. Netflix famously started as a DVD-by-mail company, focused on logistics but with an online-subscription model. Its main asset, however, was its loyal customer base who it gradually converted to video streaming and which justified the large investments required to transform Netflix into a technology and entertainment company.[13] More generally, pre-existing large customer bases allow conglomerates to make investments that standalone entrants cannot justify, potentially accelerating their dominance across multiple markets. However, this dynamic may also discourage innovative standalone entrants who, despite potentially superior products, cannot overcome the customer accumulation disadvantage—resulting in markets dominated by large incumbents expanding laterally rather than by specialized innovators. We leave the study of this trade-off, and other applications of our framework, for future work.

[13]See, e.g., Shih, Willy, and Stephen Kaufman. "Netflix in 2011." Harvard Business School Case (2014): 615-007.

References

Arcidiacono, Peter, Patrick Bayer, Jason R. Blevins, and Paul B. Ellickson, "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition," *The Review of Economic Studies*, 2016, *83* (3 (296)), 889–931.

Atkeson, Andrew and Patrick J. Kehoe, "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 2005, *113* (5), 1026–1053.

_, Christian Hellwig, and Guillermo Ordoñez, "Optimal regulation in the presence of reputation concerns," *The Quarterly Journal of Economics*, 2015, *130* (1), 415–464.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Bagwell, Kyle, "Informational product differentiation as a barrier to entry," *International Journal of Industrial Organization*, 1990, *8* (2), 207–223.

Bain, Joe S, *Barriers to new competition: their character and consequences in manufacturing industries*, Harvard University Press, 1956.

Board, Simon and Moritz Meyer ter Vehn, "A reputational theory of firm dynamics," *American Economic Journal: Microeconomics*, 2022, *14* (2), 44–80.

Bronnenberg, Bart J, Jean-Pierre H Dubé, and Matthew Gentzkow, "The evolution of brand preferences: Evidence from consumer migration," *American Economic Review*, 2012, *102* (6), 2472–2508.

_, Sanjay K Dhar, and Jean-Pierre H Dubé, "Brand history, geography, and the persistence of brand shares," *Journal of political Economy*, 2009, *117* (1), 87–115.

Cabral, Luís, "Living up to expectations: Corporate reputation and persistence of firm performance," *Strategy Science*, 2016, *1* (1), 2–11.

Crouzet, Nicolas and Janice C Eberly, "Understanding weak capital investment: The role of market concentration and intangibles," Technical Report, National Bureau of Economic Research 2019.

Dubé, Jean-Pierre, Günter J Hitsch, and Peter E Rossi, "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 2010, *41* (3), 417–445.

Farrell, Joseph and Paul Klemperer, "Chapter 31 Coordination and Lock-In: Competition with Switching Costs and Network Effects," in M. Armstrong and R. Porter, eds., *Handbook of Industrial Organization*, Vol. 3, Elsevier, January 2007, pp. 1967–2072.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "The slow growth of new plants: Learning about demand?," *Economica*, 2016, *83* (329), 91–129.

Gilbert, Richard J and David MG Newbery, "Preemptive patenting and the persistence of monopoly," *The American Economic Review*, 1982, pp. 514–526.

Gourio, François and Leena Rudanko, "Customer Capital," *The Review of Economic Studies*, July 2014, *81* (3), 1102–1136.

Grossman, Gene M and Henrik Horn, "Infant-industry protection reconsidered: the case of informational barriers to entry," *The Quarterly Journal of Economics*, 1988, *103* (4), 767–787.

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, 1982, *50* (3), 649–670.

— **and Glenn M. MacDonald**, "The Life Cycle of a Competitive Industry," *Journal of Political Economy*, 1994, *102* (2), 322–347.

Klepper, Steven, "Entry, Exit, Growth, and Innovation over the Product Life Cycle," *The American Economic Review*, 1996, *86* (3), 562–583.

Luttmer, Erzo GJ, "Selection, growth, and the size distribution of firms," *The Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

Malnight, Thomas W and Ivy Buche, "The strategic advantage of incumbency," *Harvard Business Review*, 2022, *100* (1-2), 43–48.

Penrose, Edith Tilton, *The Theory of the Growth of the Firm*, Oxford university press, 2009.

Porter, Michael E., "How Competitive Forces Shape Strategy," *Harvard Business Review*, 1979, *57* (2).

— , *Competitive strategy: techniques for analyzing industries and competitors*, New York: Free Press, 1980.

Schmalensee, Richard, "Product differentiation advantages of pioneering brands," *The American Economic Review*, 1982, *72* (3), 349–365.

Sutton, John, *Sunk costs and market structure: price competition, advertising, and the evolution of concentration*, Cambridge, Mass.: MIT Press, 1991.

Syverson, Chad, "Macroeconomics and Market Power: Context, Implications, and Open Questions," *Journal of Economic Perspectives*, August 2019, *33* (3), 23–43.

Vellodi, Nikhil, "Ratings Design and Barriers to Entry," Working Paper 2022.

Appendix

A Proofs

In this appendix, we provide proofs of the main propositions in the paper. We do not cover Propositions 2 and 5 because we provide the relevant proofs in the paper. We omit detailed derivations of some useful expressions and relegate the complete proofs, as well as auxiliary lemmas and corollaries, to the online appendix.

Proof of Proposition 1.

We first show that incumbents never exit in equilibrium. To see this, assume the first exit happens at time t_x for a firm that entered at $t_e < t_x$ (the firm would not have entered if $t_e = t_x$). We argue that there must be new entries between t_e and t_x, i.e., $G(t_x) > G(t_e)$.

If not, $G(t) = G(t_e), \forall t \in [t_e, t_x]$. We can derive a closed-form expression of $m(t; t_e)$ on this segment

$$m(t;t_e) = \int_{t_e}^{t} \frac{N}{G(t_e)} e^{-\alpha(t-s)} \mathrm{d}s = \frac{N}{\alpha G(t_e)} \left(1 - e^{-\alpha(t-t_e)}\right).$$

Therefore, $m(t; t_e)$ strictly increases in $[t_e, t_x]$. At the time of exit, the firm must make a negative flow profit, i.e., $m(t_x; t_e) - c \leq 0$. Together with the monotonicity of $m(t; t_e)$, the firm makes negative profits at every instant between $[t_e, t_x]$, so it would not have entered. Therefore, there must be new entries between t_e and t_x.

Consider a new entrant that entered at $t'_e \in (t_e, t_x)$ and would optimally exit at t'_x, where t'_x can potentially equal infinity (an incumbent that never exits). By construction, $t'_x \geq t_x$ because we have assumed the first exit happens at t_x. Denoting the continuation value of a firm with a customer base m at time t by $V_t(m)$, we examine the continuation value of the new entrant at t_x. In particular,

$$V_{t_x}(m(t_x; t'_e)) = \int_{t_x}^{t'_x} e^{-\rho(t-t_x)} \left(m(t; t'_e) - c\right) \mathrm{d}t.$$

Proposition 2 implies that $m(t; t'_e) < m(t; t_e), \forall t \geq t'_e$. Therefore, we must have

$$V_{t_x}(m(t_x; t'_e)) \leq \int_{t_x}^{t'_x} e^{-\rho(t-t_x)} \left(m(t; t_e) - c\right) \mathrm{d}t,$$

where equality is achieved only when $t'_x = t_x$. The right-hand side of the above inequality is the value of the earlier entrant (having entered at t_e) operating from t_x to t'_x. We know that this value cannot be positive; otherwise, this firm would not exit at t_x. This implies that $V_{t_x}(m(t_x; t'_e)) \leq 0$ and the inequality is strict if $t'_x > t_x$. We derive a contradiction: the later entrant at t'_e would have exited no later than t_x since $V_{t_x}(m(t_x; t'_e)) < 0$. This leaves us with the only possibility that $t'_x = t_x$.

We now show that $t'_x = t_x$ is also contradictory. First, since $G(t)$ is bounded by $G(0)$ and $G(0) + \mu G(t_x)$ on $t \in [0, t_x]$, $m(t; t_e)$ and $m(t; t'_e)$ must be continuous. We know that at the time of exit, $m(t_x; t_e) - c \leq 0$. Therefore,

$$m(t_x; t'_e) - c < m(t_x; t_e) - c \leq 0.$$

From Lemma OA.1.1, we know that $m(t; t'_e)$ can either be increasing or hump-shaped on $t \in [t'_e, t_x]$. Therefore, it is either always below c or it reaches a peak and then hits c from above at time $t''_x < t_x$. In the former case, the firm would not have entered because it always makes negative profits before t_x and the continuation value at t_x is non-positive. In the latter case, the firm would have exited before t_x because it makes negative profits on $t \in [t''_x, t_x]$.

Since we have shown that no firms exit in equilibrium, the mass of firms $G(t)$ is weakly increasing over time. The value of entry can be written as

$$V^e(t_e) = \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left[\int_{t_e}^{t} \frac{Ne^{-\alpha(t-s)}}{G(s)} \mathrm{d}s - c \right] \mathrm{d}t.$$

$G(t)$ is weakly increasing, so later entrants face more competition over their life cycles and $V^e(t_e)$ is weakly decreasing in t_e.

We now show that there exists a sufficiently large value of $\bar{G}$ such that the value of entry $V^e(t_e)$ must be negative if $G(t) \geq \bar{G}, \forall t \geq t_e$. We can calculate an upper bound

$$V^e(t_e) \leq \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left[\int_{t_e}^{t} \frac{Ne^{-\alpha(t-s)}}{\bar{G}} \mathrm{d}s - c \right] \mathrm{d}t = \frac{N}{\rho(\rho+\alpha)\bar{G}} - \frac{c}{\rho}$$

Therefore, $V^e(t_e)$ must be negative if $\bar{G} > \frac{N}{(\rho+\alpha)c}$. One implication is that entry must stop within a finite time. Otherwise, firms enter at a constant rate μ, and the mass of firms will exceed $\frac{N}{(\rho+\alpha)c}$ at some point. Since we know that $V^e(T) = 0$ and $G(t) = G(T), \forall t \geq T$, we can solve $G(T)$ as in equation (5). ∎

Proof of Proposition 3. To show that firms never exit, we apply similar arguments as in the proof of Proposition 1. The only complication now is that firms may upgrade, which changes the path $m(t; t_e)$. Consider the first exit that happens at t_x for a firm entered at t_e. We can show that, even if this firm upgrades between t_e and t_x, without new entrants in $[t_e, t_x]$, $m(t; t_e)$ monotonically increases. The firm's exit at t_x implies that it has never made positive profits in $[t_e, t_x]$, and it would not have entered. Therefore, there must be new entries between t_e and t_x.

Consider a firm that enters at $t'_e \in (t_e, t_x)$. Lemma OA.1.2 implies that the later entrant upgrades later. Therefore, at any time $t > t'_e$, we still have $m(t; t'_e) < m(t; t_e)$. This ensures that the continuation value of the later entrant at time t_x is non-positive and leaves us with the only possibility that $t'_x = t_x$. To rule out $t'_x = t_x$, we recognize that Lemma OA.1.1 still holds with upgrading because firms always upgrade when $m(t; t'_e)$ crosses m^* from below. Therefore, we have $m(t; t'_e) < c$ on $[t'_e, t_x]$ or $m(t; t'_e) < c$ on $[t''_x, t_x]$. In the former case, it would not have entered, and in the latter case, it would have exited before t_x. Therefore, firms never exit.

We first consider the case where all firms upgrade. The last entrant faces a constant number of competing firms, $G(T)$, after entry. Therefore, without upgrading,

$$m(t;T) = \frac{N}{G(T)\alpha_L}\left(1 - e^{-\alpha_L(t-T)}\right), \forall t \geq T,$$

we can solve for the optimal time for upgrading, $t_u(T)$, by setting $m(t;T) = m^*$

$$t_u(T) = T + \frac{-\log\left(1 - \frac{G(T)\alpha_L}{N}m^*\right)}{\alpha_L}.$$

Note that this also implies a necessary condition of this case, involving an endogenous variable $G(T)$: the steady-state customer capital of low-type firms, $m_{L,lr} \equiv \frac{N}{G(T)\alpha_L}$, is above the cutoff customer capital, m^*. This ensures that the last entrant's customer base will reach m^* at some point.

As derived in Online Appendix Section OA.1.1, the value of the last entrant can be written as

$$\begin{aligned} F(x) &\equiv \left(\max_{t_u} \Delta V(t_u;T)\right) + V_L^e(T) \\ &= \frac{\left(1 - \frac{\alpha_L m^*}{x}\right)^{\frac{\rho}{\alpha_L}}}{\rho + \alpha_L}\left[\frac{N(\alpha_L - \alpha_H)}{G(T)\rho(\rho + \alpha_H)} - \alpha_L I\right] + \frac{x}{\rho(\rho + \alpha_L)} - \frac{c}{\rho}, \end{aligned} \tag{A-1}$$

where $x = N/G(T) \in [\alpha_L m^*, \infty)$ is the steady-state customer arrival rate. It is clear that $F(x)$ is strictly increasing in x and

$$\lim_{x\to\infty} F(x) = \infty, \qquad \lim_{x\to\alpha_L m^*} F(x) = \frac{\alpha_L m^*}{\rho(\rho + \alpha_L)} - \frac{c}{\rho}.$$

Therefore, as long as $\frac{\alpha_L m^*}{\rho(\rho+\alpha_L)} - \frac{c}{\rho} < 0$, there exists a unique solution x (thus $G(T)$) such that $F(x) = 0$ in the corresponding range. This is equivalent to the inequality

$$I < \underline{I} = \frac{(\alpha_L - \alpha_H)(\rho + \alpha_L)c}{\rho\alpha_L(\rho + \alpha_L)}.$$

When $I \geq \underline{I}$, we know that we cannot find a solution x such that $F(x) = 0$. Therefore, at least some entrants (later entrants) will not upgrade. To ensure that at least some firms upgrade, we then need to make sure that the first entrant will upgrade, i.e., $\max_{t_u} \Delta V(t_u;0,I) \geq 0$. To derive the range of I such that this holds, we first

We first show that $\max_{t_u} \Delta V(t_u;0,I)$ strictly decreases in I. To see this, consider $I_1 < I_2$

and denote the optimal upgrading time under different investment costs as $t_u(0, I)$. We have the following relationship

$$\Delta V(t_u(0, I_2); 0, I_2) < \Delta V(t_u(0, I_2); 0, I_1) \leq \Delta V(t_u(0, I_1); 0, I_1),$$

where the first inequality can be obtained by observing equation (OA.1.1) and the relationship between ΔV and $\Delta \tilde{V}$, and the second inequality comes from the fact that $\Delta V(t_u; 0, I_1)$ is maximized under $t_u(0, I_1)$. Next, we show that $\lim_{I\to\infty} \max_{t_u} \Delta V(t_u; 0, I) \leq 0$ and $\max_{t_u} \Delta V(t_u; 0, I)|_{I=\underline{I}} > 0$ under sufficiently small $G(0)$. The first inequality results from the fact that $\Delta \tilde{V}(t_u; 0) + I$ does not depend on I, and it is bounded from above by the value of being a high type from $t = 0$. However, the value of being a high type is also bounded from above because $G(t) \geq G(0), \forall t$. Therefore, we can find sufficiently large I such that $\Delta V(t_u; 0, I) < 0$ for any t_u. We obtain $\lim_{I\to\infty} \max_{t_u} \Delta V(t_u; 0, I) \leq 0$. To prove the second inequality, note that when $I = 0$, upgrading always results in positive value $\Delta V(t_u; 0) > 0$ as long as $\alpha_H < \alpha_L$. The monotonicity in I and the two bounds imply that there is a unique $\bar{I}$ such that

$$\max_{t_u} \Delta V(t_u; 0, \bar{I}) = 0,$$

and for $I > \bar{I}$, firms that enter at $t = 0$ do not upgrade.

We now show that $\bar{I} > \underline{I}$ for sufficiently small $G(0)$. In particular, we show that the first entrant's value of upgrading is arbitrarily large when $G(0)$ is sufficiently small. Consider a firm that enters at $t = 0$ and immediately upgrades ($t_u = 0$). The net value of upgrading can be written as

$$\Delta V(0; 0, I) = \int_0^\infty e^{-\rho t} \left[\int_0^t \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(0) + \mu s} \mathrm{d}s \right] \mathrm{d}t - I.$$

In Online Appendix OA.1.1, we show that, for any t, the integral inside the first component is unbounded when $G(0) \to 0$. Therefore, we must have $\bar{I} > \underline{I}$.

Finally, the characterization of equilibrium when $I \in [\underline{I}, \bar{I}]$ follows from the construction of $\underline{I}$ and $\bar{I}$ as well as Lemma OA.1.2. ■

Proof of Proposition 4. In this proof, we denote the steady-state mass of firms $G(T)$ as G_{ss}. Correspondingly, we denote the steady-state mass of H-type firms as $G_{H,ss}$. In this proof, we only consider the case of partial upgrading, i.e., $I \in [\underline{I}, \bar{I})$, or equivalently, $G_{H,ss}/G_{ss} \in (0, 1)$, in both markets. The "corner" cases in which one market has complete or no upgrading, are generally easier to prove and we relegate them to Online Appendix OA.1.1.

Part 1: Consider two different market entry rates $\mu_1 < \mu_2$. First, note that there is entry at $t = T_u(\mu_2)$ under μ_1. This is because μ does not affect the firm's profit after the last firm enters therefore does not affect G_{ss}. With lower market entry rates, it takes longer for the economy to reach the steady state, $T(\mu_1) > T(\mu_2)$. Since $T_u(\mu_2) \leq T(\mu_2)$, we must have $T_u(\mu_2) < T(\mu_1)$.

Next, we show that the value of upgrading under μ_1 is non-negative for entrants at $T_u(\mu_2)$. To see this, note that the equilibrium under μ_2 has $G_{H,ss}(\mu_2)$ entrants at $t = T_u(\mu_2)$. Consider an entrant who enters at

$$t_e = \frac{G_{H,ss}(\mu_2) - G(0)}{\mu_1},$$

under the new entry rate μ_1. At this point, the economy has $G_{H,ss}(\mu_2)$ firms, the same as the number of firms up to $T_u(\mu_2)$ in the equilibrium under μ_2. In the equilibrium under μ_1, $G(t)$ rises from $G_{H,ss}(\mu_2)$ to G_{ss} at the rate μ_1 from t_e onward, while in the equilibrium under μ_2, $G(t)$ rises from $G_{H,ss}(\mu_2)$ to G_{ss} at the rate μ_2 from $T_u(\mu_2)$ onward. Therefore, from the perspective of the focal entrants at t_e, the future competition is stronger under μ_2 and net gains from upgrading are higher under μ_1. Therefore, the entrant at time t_e will upgrade in the equilibrium with μ_1. By Lemma OA.1.2, we must have

$$G_{H,ss}(\mu_1) \geq G_{H,ss}(\mu_2).$$

The inequality is strict if $G_{H,ss}(\mu_2) < G_{ss}$ because in this case, the future path under μ_2 has a strictly increasing segment, and the focal entrant under μ_1 has a strictly higher value of upgrading.

Consider two different flow fixed costs, $c_1 < c_2$. Denote the entry time of the last upgrader with $t_e \equiv T_u(c_1) < T(c_1)$. There are three sub-cases: (2-a) there is no entry at t_e under c_2, i.e., $t_e \geq T(c_2)$, (2-b) there is still entry at t_e under c_2, i.e., $t_e < T(c_2)$. In case (2-a), consider the last entrant under c_2. It faces constant $G(t) = G_{ss}(c_2)$ from $T(c_2)$ onward. Under both c_1 and c_2, the market entry rate is the same thus $G(t_e) \geq G_{ss}(c_2)$. Therefore, the firm that enters at t_e under c_1 faces tougher competition over its life cycle than the last entrant under c_2. Since the former upgrades, the latter must upgrade and $G_{H,ss}(c_2)/G_{ss}(c_2) = 1 > G_{H,ss}(c_1)/G_{ss}(c_1)$. In case (2-b), since $c_1 < c_2$, and the last entrant does not upgrade, by equation (5), we must have $G_{ss}(c_1) > G_{ss}(c_2)$. Therefore, the firm that enters at t_e under c_2 faces less competition over its life cycle (for some part strictly less) than the firm that enters at t_e under c_1. This gives stronger incentives for it to upgrade. Therefore, $G_{H,ss}(c_2) \geq G_{H,ss}(c_1)$, and $\frac{G_{H,ss}(c_2)}{G_{ss}(c_2)} > \frac{G_{H,ss}(c_1)}{G_{ss}(c_1)}$.

To prove the remainder of the proposition, note that the steady-state market size can be written as

$$M_{lr} = G_{H,ss}m_{H,lr} + G_{L,ss}m_{L,lr} = \frac{(\alpha_L - \alpha_H)N}{\alpha_H \alpha_L}\frac{G_{H,ss}}{G_{ss}} + \frac{N}{\alpha_L},$$

where we have used $m_{i,lr} = \frac{N}{G_{ss}\alpha_i}, i \in \{H, L\}$. The total market size (customers served) is strictly increasing in $\frac{G_{H,ss}}{G_{ss}}$. The average firm size equals $\frac{M_{lr}}{G_{ss}}$. Applying the formula of G_{ss} in equation (5), we obtain

$$\frac{M_{lr}}{G_{ss}} = \frac{(\alpha_L - \alpha_H)(\rho + \alpha_L)c}{\alpha_H \alpha_L}\frac{G_{H,ss}}{G_{ss}} + \frac{(\rho + \alpha_L)c}{\alpha_L},$$

which also increases in $\frac{G_{H,ss}}{G_{ss}}$.

Part 2: we first prove that $\frac{\partial g_H(t;\mu)}{\partial \mu} < 0$ for $t > t_{u0}(\mu)$, where $t_{u0}(\mu)$ denotes the initial entrants' time of upgrading under entry rate μ. Denote the last upgrader's time of upgrading under entry rate μ as $t_u(\mu)$. We know from Corollary OA.1.1 that $t_u(\mu) < T(\mu)$. Due to the continuity of $t_u(\mu)$ shown in Lemma OA.1.3, we can find sufficiently small $\Delta\mu$, such that $t_u(\mu - \Delta\mu) < T(\mu)$. We consider two cases $t_u(\mu - \Delta\mu) < t_u(\mu)$ and $t_u(\mu - \Delta\mu) > t_u(\mu)$. The knife edge case $t_u(\mu - \Delta\mu) = t_u(\mu)$ can be proved similarly.

First, suppose $t_u(\mu - \Delta\mu) < t_u(\mu)$. Note that $t_{u0}(\mu) > t_{u0}(\mu - \Delta\mu)$ because it takes more time to reach m^* when entry is faster and market competition is stronger. We can divide the entire range of time into four segments and discuss the relative size of $g_H(t;\mu)$ and $g_H(t;\mu - \Delta\mu)$, where the segments are divided using the cutoffs $t_{u0}(\mu - \Delta\mu), t_{u0}(\mu), t(\mu - \Delta\mu)$. On each segment, we compare the incentives of upgrading, $G_H(t)$ and $G(t)$ and conclude that $g_H(t;\mu - \Delta\mu) \geq g_H(t;\mu)$. Second, suppose $t_u(\mu - \Delta\mu) > t_u(\mu)$. We divide the entire range of time into five segments using the cutoffs $t_{u0}(\mu - \Delta\mu), t_{u0}(\mu), t(\mu), t(\mu - \Delta\mu)$ and again show that $g_H(t;\mu - \Delta\mu) \geq g_H(t;\mu)$. The inequality is strict on some of the segments. The detailed derivation is relegated to Online Appendix OA.1.1.

We have proved that in a small neighborhood of μ, $g_H(t;\mu)$ is decreasing. To show that the property holds "globally", we apply the differentiability of $g_H(t;\mu)$ with respect to μ from Lemma OA.1.4. Therefore, we must have $\partial g_H(t;\mu)/\partial\mu < 0$. For two arbitrary values of μ_1 and μ_2, we know that $g_H(t;\mu_1) = g_H(t;\mu_2)$ for $t < t_{u0}(\mu_2)$, and $g_H(t;\mu_1) > g_H(t;\mu_2)$ for $t \in [t_{u0}(\mu_1), t_{u0}(\mu_2)]$. For $t > t_{u0}(\mu_2)$, we can apply the negative derivative and obtain

$$g_H(t;\mu_2) = g_H(t;\mu_1) + \int_{\mu_1}^{\mu_2} \frac{\partial g_H(t;\mu)}{\partial \mu} d\mu < g_H(t;\mu_1).$$

We now compare the fraction of H-type firms under different flow fixed costs, $c_1 < c_2$. We denote the entry and upgrading time of the last upgrader in market 1 as $T_u(c_1)$ and $t_u(c_1)$, respectively. We know that competition in market 1 is stronger than in market 2 (strictly stronger after market 2 reaches its steady state at $T(c_2)$). This has two implications. First, the value of upgrading, $\Delta V(t_u; t_e)$ is higher in market 2 for any (t_u, t_e). Therefore, if an entrant in market 1 upgrades, it must upgrade in market 2. Second, it takes less time for any firm in market 2 to reach a customer base of m^*, the necessary condition for upgrading. Therefore, if a firm in market 1 entered at time t_e and has upgraded by time t_u, the entrant at t_e in market 2 must have upgraded at or before t_u. For $t \leq t_u(c_1)$, we must have

$$G_H(t; c_2) \geq G_H(t; c_1).$$

For $t > t_u(c_1)$, $G_H(t; c_1) = G_{H,ss}(c_1)$ is constant and $G_H(t; c_2)$ may further increase, so $G_H(t; c_2) \geq G_H(t; c_1)$ also holds in this range.

On the other hand, $G(t; c_1) \geq G(t; c_2)$ and the inequality becomes strict when $t > T(c_2)$. Therefore, we must have $g_H(t; c_1) \leq g_H(t; c_2)$ and the inequality is strict for some t. ■

Derivative-based Proof of Proposition 5. In the paper, we have provided a proof of the proposition based on a decomposition of the continuation value. Here, we provide an alternative proof based on the derivative of the value of a permanent L-type firm, V_L^e, with respect to the flow fixed cost, c. The value of entry of a permanent L-type firm is

$$V_L^e(t_e) = \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left(\int_{t_e}^{t} \frac{N}{G(s)} e^{-(1-\alpha_L)(t-s)} \mathrm{d}s \right) \mathrm{d}t - \frac{c}{\rho}. \tag{A-2}$$

Its derivative with respect to c is

$$\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} = \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \frac{e^{-\rho(T-t_e)}}{\rho(\rho+1-\alpha_L)} - \frac{1}{\rho} = \frac{e^{-\rho(T-t_e)} - 1}{\rho} \leq 0,$$

where we have applied $N/G(T) = (\rho + 1 - \alpha_L)c$ in the second equality. We provide a detailed derivation of the first equality in Online Appendix OA.1.3. Therefore, the value of entry decreases with c, and the decrease is strict when $t_e < T$. When comparing the value of entry in two markets with $c_1 < c_2$, we just need to integrate this derivative over c in this range if both markets are not saturated yet. It is straightforward to prove the case in which market 2 is saturated, but market 1 is not. We relegate the detailed proof to Online Appendix OA.1.3. ■

Proof of Proposition 6. We derive the value of a permanently L-type firm and its derivative with respect to c in the alternative, derivative-based proof of Proposition 5 in

Online Appendix Section OA.1.3. The additional value of upgrading (relative to being a permanently L-type) at t_u discounted to time t_e can be written as

$$\begin{aligned}\Delta V(t_u;t_e) &= \frac{\alpha_L-\alpha_H}{(\rho+\alpha_H)(\rho+\alpha_L)}e^{-\rho(t_u-t_e)}m(t_u;t_e)\\ &\quad+\int_{t_u}^{\infty}e^{-\rho(t-t_e)}\left[\int_{t_u}^{t}\frac{N\left(e^{-\alpha_H(t-s)}-e^{-\alpha_L(t-s)}\right)}{G(s)}\mathrm{d}s\right]\mathrm{d}t-e^{-\rho(t_u-t_e)}I\\ &=\frac{\alpha_L-\alpha_H}{(\rho+\alpha_H)(\rho+\alpha_L)}e^{-\rho t'_u}m(t'_u+t_e;t_e)\\ &\quad+\int_{t'_u}^{\infty}e^{-\rho t}\left[\int_{t'_u}^{t}\frac{N\left(e^{-\alpha_H(t-s)}-e^{-\alpha_L(t-s)}\right)}{G(s+t_e)}\mathrm{d}s\right]\mathrm{d}t-e^{-\rho t'_u}I,\end{aligned} \tag{A-3}$$

where we obtain the second equality by rewriting the integral defining $t'_u = t_u - t_e$ as the age of the firm when it upgrades. We denote $t'_u(t_e, c)$ as the age of a firm born at t_e in a market with fixed cost c when it optimally upgrades.

Note that the cutoff $\underline{I}$ is linear in c following the formula (13), the assumption that the last entrant does not upgrade implies that $I \geq \underline{I}(c_2) \geq \underline{I}(c) \geq \underline{I}(c_1), \forall c \in [c_1, c_2]$. Therefore, we have the partial upgrading case for all $c \in [c_1, c_2]$. From Corollary OA.1.1, we know that upgrading must happen at or before $T(c_2)$.

We now consider a firm that upgrades at $t_u \leq T(c_2)$. We can take the derivative of $\Delta V(t'_u(t_e, c), t_e; c)$ with respect to the parameter c:

$$\frac{\mathrm{d}\Delta V(t'_u(t_e,c),t_e;c)}{\mathrm{d}c} = \frac{\partial \Delta V(t'_u,t_e;c)}{\partial c}\bigg|_{t'_u=t'_u(t_e,c)} = \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c}\times\frac{(\alpha_L-\alpha_H)e^{-\rho(T-t_e)}}{\rho(\rho+\alpha_H)(\rho+\alpha_L)}.$$

where the first equality is the envelope theorem.[14] We provide detailed derivations of the second equality in Online Appendix OA.1.3. Combining this with the earlier expression of $\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c}$, we have

$$\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c} = \frac{\mathrm{d}\Delta V(t'_u(t_e),t_e)}{\mathrm{d}c} + \frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} = \frac{\frac{\rho+\alpha_L}{\rho+\alpha_H}e^{-\rho(T-t_e)}-1}{\rho}. \tag{A-4}$$

With $G(0)$ sufficiently small as well as $I > \underline{I}$, we can ensure that early entrants upgrade and late ones do not. In Online Appendix OA.1.3, we show that as part of a sufficient

[14] When we evaluate the derivative at $c = c_2$, we calculate the left-hand derivative so that the breakdown of the integral at $T - t_e$ is well-defined: a smaller c ensures that $t_u(t_e, c) - t_e < T(c) - t_e$ for all upgraders.

condition for the derivative to be positive, we need sufficiently large μ:

$$\mu > \frac{\rho N}{(\rho+\alpha_L)c \cdot \log\left(\frac{\rho+\alpha_L}{\rho+\alpha_H}\right)}.$$

Intuitively, a higher entry rate reduces T and shortens the period between t_e and T. The future benefit of less competition is discounted less, and $\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c}$ is more likely to be positive. When the derivative is positive for all $c \in [c_1, c_2]$, we integrate it over this range and can show that the value of entry is higher in market 2 than in market 1. However, for a late entrant that does not upgrade in either market, we apply Proposition 5 and show that market 1 is more attractive. ■

Proof of Proposition 7.

Part 1 We first show that the time of entry of the last entrant, T_j, must be the same for $j = 1, 2$. To see this, consider the case when $T_1 > T_2$. Then the last entrant in market 2 has a strictly lower value of entry at T_2 compared to that in market 1. Therefore, it would have entered market 1 instead of market 2. This leads to a contradiction. We can rule out $T_1 < T_2$ similarly. Therefore, we must have $T_1 = T_2$.

In the no-upgrading case, the last entrant does not upgrade, and the steady-state mass of firms in market j does not depend on the market-specific entry rates and must have the following expression:

$$G_{ss,j} = \frac{N}{(\rho+\alpha_L)c_j},\ j = 1, 2.$$

We can also use the steady-state mass of firms and total entry rate to calculate T_j:

$$T_j = \frac{G(T;c_1) + G(T;c_2) - G(0)}{\mu} \equiv T,\ j = 1, 2.$$

In Online Appendix OA.1.3, we derive an expression for $\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}t_e}$, which further implies

$$\frac{\mathrm{d}V^e(t_e;c_2)}{\mathrm{d}t_e} - \frac{\mathrm{d}V^e(t_e;c_1)}{\mathrm{d}t_e} = c_2 - c_1 - \left(\frac{N}{(\rho+\alpha_L)G(t_e;c_2)} - \frac{N}{(\rho+\alpha_L)G(t_e;c_1)}\right). \tag{A-5}$$

We know that $V^e(t_e;c_2) = V^e(t_e;c_1) = 0$ at $t_e = T$. Substituting in the expressions of $G_{ss,j}$, we have $\mathrm{d}V_L^e(t_e;c_2)/\mathrm{d}t_e = \mathrm{d}V_L^e(t_e;c_1)/\mathrm{d}t_e$, so $V^e(t_e;c_2) = V^e(t_e;c_1)$ in a small neighborhood of T_1. In fact, if we can maintain $\mathrm{d}V_L^e(t_e;c_2)/\mathrm{d}t_e = \mathrm{d}V_L^e(t_e;c_1)/\mathrm{d}t_e$ thus firms are indifferent between entering the two markets at any time point, together with $G(t_e;c_2) + G(t_e;c_1) = G(0) + \mu t_e$, we obtain solutions to $G(t_e;c_j)$ for $t_e \in [0, T]$. The analytical expression for this solution can be found in Online Appendix equation (OA.1.8).

To show the uniqueness of the equilibrium, we now rule out the possibility of an equilibrium in which one market has a strictly higher value of entry at some point. Since $T_1 = T_2$, this has to occur before the steady state. Without loss of generality, we assume that $V^e(t_e; c_1) > V^e(t_e; c_2)$ for on $t_e \in (s,t), s < t < T_1$, while $V^e(t_e; c_1) = V^e(t_e; c_2)$ for all $t_e \geq t$. This implies that $G'(t_e; c_1) = \mu, G'(t_e; c_2) = 0$ on the segment (s,t), and the value of equation (A-5) becomes positive.

Suppose at some point at or before s, there is entry in market 2 again, which requires $V^e(t_e; c_1) \leq V^e(t_e; c_2)$. Without loss of generality, we assume this happens exactly at s. However, this leads to an immediate contradiction because $V^e(t; c_1) = V^e(t; c_2)$ and $\frac{\mathrm{d}V^e(t_e;c_2)}{\mathrm{d}t_e} - \frac{\mathrm{d}V^e(t_e;c_1)}{\mathrm{d}t_e} > 0$ for all $t_e \in (s,t)$, which implies that $V^e(t_e; c_1) > V^e(t_e; c_2)$ for all $t_e \in (s,t)$. Therefore, there is no entry at or before t in market 2. This leads to another contradiction: according to equation (OA.1.8), we must have $G(t; c_2) > 0$. We can obtain similar contradictions if we assume $V^e(t_e; c_2) > V^e(t_e; c_1)$ on a segment $t_e \in (s,t)$. Therefore, we conclude that there is a unique equilibrium in which entrepreneurs are indifferent between entering either market at any time, characterized by equation (OA.1.8).

Setting the value of equation (A-5) to zero and taking derivative with respect to t_e, we have

$$\frac{G'(t_e; c_1)}{G^2(t_e; c_1)} = \frac{G'(t_e; c_2)}{G^2(t_e; c_2)}.$$

Since $G(t_e; c_2) < G(t_e; c_1)$, we establish the comparison of entry rates: $G'(t_e; c_2) < G'(t_e; c_1)$.

Part 2 Since $I > \underline{I}(\rho, \alpha_L, \alpha_H, c_2) > \underline{I}(\rho, \alpha_L, \alpha_H, c_1)$, the last entrant does not upgrade in either market. The steady-state mass of firms, therefore, is the same as in Part 1. Since $G(t; c_j)$ is continuous in t and $G(T; c_1) > G(T; c_2)$, we must have $G(t; c_1) > G(t; c_2)$ in a neighborhood of $(T - \epsilon, T]$. Market 1 has more firms when t is sufficiently large.

Now suppose $G(t; c_1) \geq G(t; c_2)$ for all t. We seek contradictions as well as parameter combinations that lead to such contradictions. We also choose $G(0)$ to be small enough such that $G(0) < G(T; c_2)$. We now construct a collection of paths to facilitate our proof. These paths depend on a hypothetical fixed cost parameter $c \in [c_2, c_1]$ and are defined as follows

$$\hat{G}(t; c) \equiv \min\{G(t; c_1), G(T; c)\},$$

where $G(T; c) = \frac{N}{(\rho+\alpha_L)c}$ is the steady-state mass of firms under fixed cost c. These paths coincide with $G(t; c_1)$ until it reaches $G(T; c)$. Since $G(T; c_2) \leq G(T; c) \leq G(T; c_1)$, all these paths are higher than $G(t; c_2)$ but lower than $G(t; c_1)$.

We now attempt to derive a contradiction by showing that the value of entry in market

2 is strictly higher than that in market 1 at $t = 0$ under certain parameter restrictions. Consider a potential entrant at $t = 0$ facing a future path of firm mass $\hat{G}(t;c)$, and denote its value of entry as $\hat{V}^e(0;c)$. We restrict $G(0)$ to be sufficiently small so that the initial entrants in all markets with $c \in [c_1, c_2]$ upgrade. In Online Appendix OA.1.3, we show that as long as

$$\mu > \frac{\rho(G(T;c_1) + G(T;c_2))}{\log\left(\frac{\rho+\alpha_L}{\rho+\alpha_H}\right)},$$

we can ensure that $\frac{\partial \hat{V}^e(0;c)}{\partial c} > 0$ for all $c \in [c_1, c_2]$. Therefore, we have

$$V^e(0;c_1) = \hat{V}^e(0;c_1) < \hat{V}^e(0;c_2) \leq V^e(0;c_2),$$

which implies that $G(0;c_1) = 0 < G(0;c_2)$ and contradicts the hypothesis that $G(t;c_2) \leq G(t;c_1)$ for all $t \geq 0$. ■

Proof of Proposition 8. We first prove a stronger result that when ρ' is sufficiently large or when ρ is sufficiently small, we must have the value of a permanent L-type firm $V_L^e(t_e)$ increases with c. In the proof of Proposition 5, we have obtained

$$\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} = \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c}\frac{e^{-\rho(T-t_e)}}{\rho(\rho+\alpha_L)} - \frac{1}{\rho}.$$

The steady-state customer arrival rate $N/G(T)$ is determined by the prevailing discount factor in the market, ρ', rather than the focal firm's discount factor, ρ. Since we assume that the last entrant does not upgrade, $N/G(T)$ equals $(\rho' + \alpha_L)c$. We can simplify the above expression as

$$\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} = \frac{\frac{\rho'+\alpha_L}{\rho+\alpha_L}e^{-\rho(T-t_e)} - 1}{\rho}.$$

When $\rho' \to \infty$, we have $T - t_e < \frac{\frac{N}{(\rho'+\alpha_L)c} - G(0)}{\mu} \to 0$. Therefore, $\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} > 0$ when ρ' is sufficiently large. When $\rho \to 0$, the numerator of the above expression converges to ρ'/α_L, which also implies $\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} > 0$.

In the proof or Proposition 6, we have also shown that $\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c} \geq \frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c}$. Therefore, $\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c}$ is positive when ρ' is sufficiently large. Therefore, the value of entry increases with c whether the entrepreneur upgrades or not.[15] ■

[15] When $c = c^*$ is such that the patient entrant at t_e is indifferent between upgrading and not upgrading, there is a kink in $V^e(t_e;c)$ as a function of c at c^*. Since a large c leads to upgrading, we have

$$V^e(t_e;c^{*+}) = \Delta V(t_u(t_e), t_e; c^*) + V_L^e(t_e;c^*) \geq V_L^e(t_e;c^*) = V^e(t_e;c^{*-}).$$

B Continuous-time Dynamic Choices by Consumers

In this section, we provide a model of consumers' dynamic choices in continuous time, which micro-founds the constant depreciation of customer capital at any instant. It also provides a setup where we can discuss the firm's optimal pricing.

B.1 Model

We consider a typical customer of an incumbent firm. A Poisson arrival process determines when the customer can change his/her consumption decision. At rate λ, the customer decides whether to stay with the current firm ($j = 1$) or to take the outside option permanently ($j = 0$). Denoting the consumer discount factor as ρ^c and the flow utility of being attached to the current firm as u_1, we can write the Bellman equation from time t to time $t + h$ as

$$V_1^c = u_1 h + \frac{1}{1+\rho^c h}\left[(1-\lambda h)V_1^c + \lambda h E_{max}\left\{V_1^c + \varepsilon_1, V_0^c + \varepsilon_0\right\}\right], \tag{B-1}$$

where ε_j is an instantaneous choice-specific payoff shock. The value of taking the outside option is $V_0^c = u_0/\rho$. Following Arcidiacono et al. (2016), we assume that the shocks are i.i.d. across consumers and over time. We also assume $E_{max}\left\{\varepsilon_1, \varepsilon_0\right\}$ is finite to ensure the value function is bounded. Taking $h \to 0$, we have

$$V_1^c = \frac{u_1 + \lambda E_{max}\left\{V_1^c + \varepsilon_1, V_0^c + \varepsilon_0\right\}}{\rho + \lambda} \equiv \Gamma V_1^c, \tag{B-2}$$

where we define the right-hand side of the above equation as a mapping Γ from V_1^c to ΓV_1^c. Online Appendix Lemma OA.1.5 shows that Γ is a contraction mapping. Therefore, there is a unique solution V_1^c to equation (B-2).

To obtain a constant customer attrition rate α, we need

$$\alpha = \lambda \Pr(V_0^c + \varepsilon_0 \geq V_1^c + \varepsilon_1) \tag{B-3}$$

As a normalization, we can set $u_0 = V_0^c = 0$. Therefore, we can back out the value of V_1^c that is consistent with α, which depends on α, λ and the joint density of $(\varepsilon_0, \varepsilon_1)$.

Proposition B-1 *For two firms, if $\alpha_H < \alpha_L$, we must have $V_{1,H}^c > V_{1,L}^c$ and $u_{1,H}^c > u_{1,L}^c$.*

Proof. When $\alpha_H < \alpha_L$, it is immediate from equation (B-3) that $V^c_{1,H} > V^c_{1,L}$. To indicate the dependence of the contraction mapping (B-1) on flow utility u_1, we now write it as $\Gamma(V^c_1; u_1)$. $\Gamma(V^c_1; u_1)$ strictly increases in u_1. Therefore, $V^c_{1,H} > V^c_{1,L}$ must imply that $u^c_{1,H} > u^c_{1,L}$. ■

B.2 Optimal Pricing

We now consider firms' optimal prices under the above microfoundation. We assume that firms cannot commit to future prices and that consumers do not expect firms to upgrade. We solve for a price p that maximizes the value of an existing customer. We assume that consumers have quasi-linear utility, so we can rewrite the value function as

$$V^c_1 = \frac{s + \lambda E_{max}\left\{V^c_1 + \varepsilon_1, V^c_0 + \varepsilon_0\right\}}{\rho + \lambda}$$

where $s \equiv u_1 - p$ is the consumer surplus. The value of matching with a firm is now a function of $u_1 - p$ instead of u_1. Therefore, the customer attrition rate is $\alpha(u_1 - p; \cdots)$, where the dots represent other parameters such as λ and the joint distribution of $(\varepsilon_0, \varepsilon_1)$.

A firm chooses p by trading off a higher flow profit against a higher rate of the consumer leaving $\alpha(u_1 - p)$. Consider the value of an old customer when the firm charges p,

$$v(p) = ph + \frac{1}{1 + \rho h}\left[1 - \alpha(u_1 - p)h\right] v(p).$$

Taking $h \to 0$, we have

$$v(p) = \frac{p}{\rho + \alpha(u_1 - p)},$$

where $\alpha(u_1 - p) \in [0, \lambda]$. We impose a limit price $\bar{p}$ above which the customer prefers not to buy anything from the firm and $v(p) = 0$. This rules out the scenario where the firm chooses to charge an infinitely high price to extract value from existing customers who do not immediately have a chance to take the outside option.[16]

With the restriction that $p \in (0, \bar{p}]$, we search for an interior solution. The model predicts that consumers would derive higher utility from a higher-quality firm, despite the possibility of higher prices. In addition, we characterize the conditions under which a higher-quality firm charges a higher price:

[16] We can also eliminate this scenario if we allow λ to be an increasing function of p and the increase is fast enough.

Proposition B-2 *Suppose there exists* $p^* \in (0, \bar{p})$ *that maximizes* $v(p)$*, we must have*

$$\frac{\mathrm{d}(u_1 - p)}{\mathrm{d}u_1}\bigg|_{p=p^*} \geq 0.$$

If we further assume $p^*\alpha''/\alpha' < -1$*, where the derivatives are taken with respect to* $u_1 - p$ *at* $p = p^*$*, we have*

$$\frac{\mathrm{d}p^*}{\mathrm{d}u_1} > 0.$$

Online Appendix

OA.1 Additional Theoretical Results

In this section, we provide additional theoretical analysis of our main model. We provide detailed proofs of the propositions along with lemmas and other theoretical results that are needed for the proofs.

OA.1.1 Proofs of Propositions in Sections 3 and 4 and Additional Results

The following lemma characterizes the shape of the path $m(t;t_e)$ and is useful when proving Proposition 1 and other results in the paper.

Lemma OA.1.1 *If $G(t)$ is weakly increasing on $t \in [t_e, t_1]$, $m(t;t_e), t \in [t_e, t_1]$ is either increasing or hump-shaped.*

Proof of Lemma OA.1.1. We first show that on any segment $[t_e, t_1]$ where $G(t)$ is weakly increasing, $m(t;t_e)$ cannot first decrease and then increase. If this happens, there exists a local minimum of $m(t;t_e)$ at t_0. We will show that the existence of t_0 leads to a contradiction.

Note that the solution to the law of motion for the customer base, (1), is

$$m(t;t_e) = m(t_e;t_e)e^{-\alpha(t-t_e)} + \int_{t_e}^{t} \frac{N}{G(s)} e^{-\alpha(t-s)} \mathrm{d}s,$$

which is continuous in t. Since we have assumed $G(t)$ to be weakly increasing, we rule out mass exits, i.e., discontinuity in $G(t)$ such that $G(t^-) > G(t^+)$. Therefore, by construction, $G(t) = \int_0^t (\mu(s) - x(s))\, \mathrm{d}s$, where $x(s)$ is a finite exit rate. Since $G(t)$ is differentiable, it must be continuous. Given the law of motion, $\frac{\mathrm{d}m(t;t_e)}{\mathrm{d}t} = -\alpha m(t;t_e) + \frac{N}{G(t)}$, $\frac{\mathrm{d}m(t;t_e)}{\mathrm{d}t}$ is also continuous in t. Since t_0 is a local minimum, and $\frac{\mathrm{d}m(t;t_e)}{\mathrm{d}t}$ is continuous, we must have $\frac{\mathrm{d}m(t;t_e)}{\mathrm{d}t} < 0$ for a neighborhood on the left of t_0 and $\frac{\mathrm{d}m(t;t_e)}{\mathrm{d}t} > 0$ for a neighborhood on the right of t_0. We denote these two neighborhoods as $(t_0 - h, t_0)$ and $(t_0, t_0 + h)$ with $h > 0$. By the Mean

Value Theorem, we can find $h_1, h_2 \in (0, h)$ such that

$$\frac{m(t_0 - h; t_e) - m(t_0; t_e)}{-h} = \frac{\mathrm{d}m(t; t_e)}{\mathrm{d}t}|_{t=t_0-h_1} < 0,$$
$$\frac{m_{t_e}(t_0 + h) - m(t_0; t_e)}{h} = \frac{\mathrm{d}m(t; t_e)}{\mathrm{d}t}|_{t=t_0+h_2} > 0.$$

This implies that

$$m(t_0 - h; t_e) > m(t_0; t_e), \quad m(t_0 + h; t_e) > m(t_0; t_e).$$

By continuity of $m(t; t_e)$, there exists $h_1', h_2' \in (0, h)$ such that

$$m(t_0 - h_1'; t_e) = m(t_0 + h_2'; t_e) > m(t_0; t_e).$$

Since $G(t)$ is weakly decreasing, unless $m(t; t_e)$ is a constant, we must have

$$\frac{N}{G(t_0 - h_1')} < \alpha m(t_0 - h_1'; t_e) = \alpha m(t_0 + h_2'; t_e) < \frac{N}{G(t_0 + h_2')}.$$

We obtained a contradiction.

Now we have ruled out the possibility of $m(t; t_e)$ first decreasing and then increasing, $m(t; t_e)$ can be either increasing, decreasing, or hump-shaped (first increasing and then decreasing). When $m(t_e; t_e) = 0$, $m(t; t_e)$ can only be either increasing or hump-shaped. In general, when $m(t_e; t_e)$ can be positive, $m(t; t_e)$ can be decreasing on $t \in [t_e, t_1]$. ■

Proof of Proposition 1.

We first show that incumbents never exit in equilibrium. To see this, assume the first exit happens at time t_x for a firm that entered at $t_e < t_x$ (the firm would not have entered if $t_e = t_x$). We argue that there must be new entries between t_e and t_x, i.e., $G(t_x) > G(t_e)$. If not, $G(t) = G(t_e), \forall t \in [t_e, t_x]$. We can derive a closed-form expression of $m(t; t_e)$ on this segment

$$m(t; t_e) = \int_{t_e}^{t} \frac{N}{G(t_e)} e^{-\alpha(t-s)} \mathrm{d}s = \frac{N}{\alpha G(t_e)} \left(1 - e^{-\alpha(t-t_e)}\right).$$

Therefore, $m(t; t_e)$ strictly increases in $[t_e, t_x]$. At the time of exit, the firm must make a negative flow profit, so

$$m(t_x; t_e) - c \leq 0.$$

Together with the monotonicity of $m(t; t_e)$, the firm makes negative profits at every instant between $[t_e, t_x]$, so it would not have entered. Therefore, there must be new entries between

t_e and t_x.

Consider a new entrant that entered at $t'_e \in (t_e, t_x)$ and would optimally exit at t'_x, where t'_x can potentially equal infinity. By construction, $t'_x \geq t_x$ because we have assumed the first exit happens at t_x. Denoting the continuation value of a firm with a customer base m at time t by $V_t(m)$, we examine the continuation value of the new entrant at t_x. In particular,

$$V_{t_x}(m(t_x; t'_e)) = \int_{t_x}^{t'_x} e^{-\rho(t-t_x)} \left(m(t; t'_e) - c\right) \mathrm{d}t.$$

The law of motion of customer base (2) implies that

$$m(t; t'_e) < m(t; t_e), \forall t \geq t'_e.$$

Therefore, we must have

$$V_{t_x}(m(t_x; t'_e)) \leq \int_{t_x}^{t'_x} e^{-\rho(t-t_x)} \left(m(t; t_e) - c\right) \mathrm{d}t,$$

where equality is achieved only when $t'_x = t_x$. The right-hand side of the above inequality is the value of the earlier entrant (having entered at t_e) operating from t_x to t'_x. We know that this value cannot be positive, otherwise this firm would not exit at t_x. This implies that $V_{t_x}(m(t_x; t'_e)) \leq 0$ and the inequality is strict if $t'_x > t_x$. It is contradictory since $V_{t_x}(m(t_x; t'_e)) < 0$ and this later entrant will exit no later than t_x. This leaves us with the only possibility that $t'_x = t_x$.

We now show that $t'_x = t_x$ is also contradictory. First, since $G(t)$ is bounded by $G(0)$ and $G(0) + \mu G(t_x)$ on $t \in [0, t_x]$, $m(t; t_e)$ and $m(t; t'_e)$ must be continuous. We know that at the time of exit, $m(t_x; t_e) - c \leq 0$. Therefore,

$$m(t_x; t'_e) - c < m(t_x; t_e) - c \leq 0.$$

From Lemma OA.1.1, we know that $m(t; t'_e)$ can either be increasing or hump-shaped on $t \in [t'_e, t_x]$. Therefore, it is either always below c or it reaches a peak and then hits c from above at time $t''_x < t_x$. In the former case, the firm would not have entered because it always makes negative profits before t_x and the continuation value at t_x is non-positive. In the latter case, the firm would have exited before t_x because it makes negative profits on $t \in [t''_x, t_x]$.

Since we have shown that no firms exit in equilibrium, the mass of firms $G(t)$ is weakly increasing over time. The value of entry can be written as

$$V^e(t_e) = \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left[\int_{t_e}^{t} \frac{Ne^{-\alpha(t-s)}}{G(s)} \mathrm{d}s - c \right] \mathrm{d}t.$$

$G(t)$ is weakly increasing, so later entrants face more competition over their life cycles and $V^e(t_e)$ is weakly decreasing in t_e.

We now show that there exists a sufficiently large value of $\bar{G}$ such that the value of entry $V^e(t_e)$ must be negative if $G(t) \geq \bar{G}, \forall t \geq t_e$. We can calculate an upper bound

$$V^e(t_e) \leq \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left[\int_{t_e}^{t} \frac{Ne^{-\alpha(t-s)}}{\bar{G}} \mathrm{d}s - c \right] \mathrm{d}t = \frac{N}{\rho(\rho+\alpha)\bar{G}} - \frac{c}{\rho}$$

Therefore, $V^e(t_e)$ must be negative if $\bar{G} > \frac{N}{(\rho+\alpha)c}$. One implication is that entry must stop within a finite time. Otherwise, firms enter at a constant rate μ, and the mass of firms will exceed $\frac{N}{(\rho+\alpha)c}$ at some point. Since we know that $V^e(T) = 0$ and $G(t) = G(T), \forall t \geq T$, we can solve $G(T)$ as in equation (5). ■

OA.1.2 Proof of Proposition 3

Before proving the proposition that characterizes the equilibrium when firms can upgrade from α_L to α_H, we derive some useful expressions and a lemma. We first derive general expressions for the customer base and the net gain of upgrading for firms that enter at t_e and upgrade at t_u. We use $\Delta m(t; t_e), t \geq t_u$ to denote the difference in customer base after upgrading, $\Delta\tilde{V}(t_e, t_u)$ to denote the net gain of upgrading evaluated at the time of upgrading $t = t_u$. We have the following:

$$\begin{aligned}
\Delta m(t; t_e) &= m(t_u; t_e) \left(e^{-\alpha_H(t-t_u)} - e^{-\alpha_L(t-t_u)} \right) + \int_{t_u}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)} \right)}{G(s)} \mathrm{d}s \\
\Delta\tilde{V}(t_u; t_e) &= \frac{\alpha_L - \alpha_H}{(\rho+\alpha_H)(\rho+\alpha_L)} m(t_u; t_e) \\
&\quad + \int_{t_u}^{\infty} e^{-\rho(t-t_u)} \left[\int_{t_u}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)} \right)}{G(s)} \mathrm{d}s \right] \mathrm{d}t - I. \qquad \text{(OA.1.1)}
\end{aligned}$$

Since we define $\Delta V(t_e, t_u)$ as the net gain of upgrading evaluated at the time entry $t = t_e$, we have

$$\Delta V(t_u; t_e) = e^{-\rho(t_u - t_e)} \Delta\tilde{V}(t_u; t_e).$$

Taking the derivative of $\Delta V(t_u; t_e)$ with respect to t_u, we obtain the necessary condition for upgrading (9). For the second component on the right-hand side of equation (OA.1.1), we can further simplify it if $t_u \geq T$:

$$\int_{t_u}^{\infty} e^{-\rho(t-t_u)} \left[\int_{t_u}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(s)} \mathrm{d}s \right] \mathrm{d}t = \frac{(\alpha_L - \alpha_H)N}{\rho(\rho+\alpha_H)(\rho+\alpha_L)G(T)}$$

Before we characterize the equilibrium regarding whether some or all entrants upgrade, we first prove a useful property of the equilibrium:

Lemma OA.1.2 *If the firms that entered at time t_e upgrade, then all firms that entered before t_e also upgrade. Earlier entrants wait less time before upgrading.*

Proof of Lemma OA.1.2. Suppose that a firm that entered at time t_e upgrades at t_u, and suppose $t_e' < t_e$. We know that

$$m(t_u; t_e') > m(t_u; t_e)$$

because the earlier entrants accumulated customers in $t \in [t_e', t_e)$. From the expression of $\Delta\tilde{V}(t_u; t_e)$, equation (OA.1.1), we obtain immediately

$$\Delta\tilde{V}(t_u; t_e') > \Delta\tilde{V}(t_u; t_e) \geq 0,$$

where the last inequality holds because the firm that entered at time t_e finds it profitable to upgrade at t_u. Therefore, it is also profitable for the earlier entrants to upgrade at t_u, though upgrading at the optimal $t_u(t_e')$ will bring them even larger discounted net gains $\Delta V(t_u(t_e'); t_e')$.

To show that earlier entrants wait less time before upgrading, we use the necessary condition for upgrading in Lemma 1. Suppose two entrants enter at t_e and t_e', and upgrade at t_u and t_u', respectively, and $t_e < t_e'$. We know that

$$m(t_u; t_e) = m(t_u'; t_e') = m^*.$$

However, since $t_e < t_e'$, the competition faced by the later entrant over its life cycle is strictly stronger than that of the earlier entrant. Therefore, the earlier entrant accumulates customers at a faster rate compared to the later entrant when they are of the same age, i.e., $t - t_e = t' - t_e'$. Therefore, $m(t; t_e)$ must have crossed m^* before $t_e + (t_u' - t_e')$ if $m(t_u'; t_e') = m^*$. The earlier entrant waits less before upgrading. ■

Proof of Proposition 3. To show that firms never exit, we apply similar arguments as in the proof of Proposition 1. The only complication now is that firms may upgrade, and it changes the path $m(t;t_e)$. Consider the first exit that happens at t_x for a firm entered at t_e. We can show that, even if this firm upgrades between t_e and t_x, without new entrants in $[t_e, t_x]$, $m(t;t_e)$ monotonically increases. The firm exiting at t_x implies that it has never made positive profits in $[t_e, t_x]$, and it would not have entered. Therefore, there must be new entries between t_e and t_x.

Consider a firm that enters at $t'_e \in (t_e, t_x)$. Lemma OA.1.2 implies that the later entrant upgrades later. Therefore, at any time $t > t'_e$, we still have $m(t;t'_e) < m(t;t_e)$. This ensures that the continuation value of the later entrant at time t_x is non-positive and leaves us with the only possibility that $t'_x = t_x$. To rule out $t'_x = t_x$, we recognize that Lemma OA.1.1 still holds with upgrading because firms always upgrade when $m(t;t'_e)$ crosses m^* from below. Therefore, we have $m(t;t'_e) < c$ on $[t'_e, t_x]$ or $m(t;t'_e) < c$ on $[t''_x, t_x]$. In the former case, it would not have entered, and in the latter case, it would have exited before t_x. Therefore, firms never exit.

We first consider the case where all firms upgrade. The last entrant faces a constant number of competing firms, $G(T)$, after entry. Therefore, without upgrading,

$$m(t;T) = \frac{N}{G(T)\alpha_L}\left(1 - e^{-\alpha_L(t-T)}\right), \forall t \geq T,$$

we can solve for the optimal time for upgrading, $t_u(T)$, by setting $m(t;T) = m^*$

$$t_u(T) = T + \frac{-\log\left(1 - \frac{G(T)\alpha_L}{N}m^*\right)}{\alpha_L}.$$

Note that this also implies a necessary condition of this case, involving an endogenous variable $G(T)$: the steady-state customer capital of low-type firms, $m_{L,lr} \equiv \frac{N}{G(T)\alpha_L}$, is above the cutoff customer capital, m^*. This ensures that the last entrant's customer base will reach m^* at some point.

In this case, the zero-profit condition for the last entrant becomes

$$\left(\max_{t_u} \Delta V(t_u;T)\right) + V_L^e(T) = \left(\max_{t_u} \Delta V(t_u;T)\right) + \frac{N}{G(T)\rho(\rho+\alpha_L)} - \frac{c}{\rho} = 0,$$

where we have explicitly calculated the value of being a low-type firm forever since T:

$$V_L^e(T) = \int_T^\infty e^{-\rho(t-T)}(m(t;T) - c)\mathrm{d}t = \frac{N}{G(T)\rho(\rho+\alpha_L)} - \frac{c}{\rho}$$

The gains from optimal upgrading are

$$\max_{t_u} \Delta V(t_u;T) = \frac{\left(1 - m^* \frac{G(T)\alpha_L}{N}\right)^{\frac{\rho}{\alpha_L}}}{\rho + \alpha_L} \left[\frac{N(\alpha_L - \alpha_H)}{G(T)\rho(\rho+\alpha_H)} - \alpha_L I\right], \tag{OA.1.2}$$

where we have substituted in the expression of $t_u(T)$. We now define the entry value of the last entrant as

$$\begin{aligned} F(x) &\equiv \left(\max_{t_u} \Delta V(t_u;T)\right) + V_L^e(T) \\ &= \frac{\left(1 - \frac{\alpha_L m^*}{x}\right)^{\frac{\rho}{\alpha_L}}}{\rho + \alpha_L} \left[\frac{N(\alpha_L - \alpha_H)}{G(T)\rho(\rho+\alpha_H)} - \alpha_L I\right] + \frac{x}{\rho(\rho+\alpha_L)} - \frac{c}{\rho}, \end{aligned} \tag{OA.1.3}$$

where $x = N/G(T) \in [\alpha_L m^*, \infty)$ is the steady-state customer arrival rate. It is clear that $F(x)$ is strictly increasing in $N/G(T)$ and

$$\lim_{x\to\infty} F(x) = \infty, \quad \lim_{x\to\alpha_L m^*} F(x) = \frac{\alpha_L m^*}{\rho(\rho+\alpha_L)} - \frac{c}{\rho}.$$

Therefore, as long as $\frac{\alpha_L m^*}{\rho(\rho+\alpha_L)} - \frac{c}{\rho} < 0$, there exists a unique solution x (thus $G(T)$) in the corresponding range. This is equivalent to the inequality

$$I < \underline{I} = \frac{(\alpha_L - \alpha_H)(\rho+\alpha_L)c}{\rho\alpha_L(\rho+\alpha_H)}.$$

When $I \geq \underline{I}$, we know that we cannot find a solution $G(T)$ such that $F(G(T)) = 0$. Therefore, at least some entrants (later entrants) will not upgrade. To ensure that at least some firms upgrade, we must also ensure that the first entrant upgrades. Formally, we need

$$\max_{t_u} \Delta V(t_u; 0, I) \geq 0.$$

We can show that there is a unique $\bar{I}$ such that

$$\max_{t_u} \Delta V(t_u; 0, \bar{I}) = 0,$$

and for $I > \bar{I}$, firms that enter at $t = 0$ do not upgrade. We first show that $\max_{t_u} \Delta V(t_u; 0, I)$ strictly decreases in I. To see this, consider $I_1 < I_2$ and denote the optimal upgrading time under different investment costs as $t_u(0, I)$. We have the following relationship

$$\Delta V(t_u(0, I_2); 0, I_2) < \Delta V(t_u(0, I_2); 0, I_1) \leq \Delta V(t_u(0, I_1); 0, I_1),$$

where the first inequality can be obtained by observing equation (OA.1.1) and the relationship between ΔV and $\Delta\tilde{V}$, and the second inequality comes from the fact that $\Delta V(t_u; 0, I_1)$ is maximized under $t_u(0, I_1)$. Next, we show that $\lim_{I\to\infty} \max_{t_u} \Delta V(t_u; 0, I) \leq 0$ and $\max_{t_u} \Delta V(t_u; 0, I)|_{I=\underline{I}} > 0$ under sufficiently small $G(0)$. The first inequality results from the fact that $\Delta\tilde{V}(t_u; 0) + I$ does not depend on I, and it is bounded from above by the value of being a high type from $t = 0$. However, the value of being a high type is also bounded from above because $G(t) \geq G(0), \forall t$. Therefore, we can find sufficiently large I such that $\Delta V(t_u; 0, I) < 0$ for any t_u. We obtain $\lim_{I\to\infty} \max_{t_u} \Delta V(t_u; 0, I) \leq 0$. To prove the second inequality, note that when $I = 0$, upgrading always results in positive value $\Delta V(t_u; 0) > 0$ as long as $\alpha_H > \alpha_L$. In sum, we have shown the existence and uniqueness of $\bar{I}$.

We now show that $\bar{I} > \underline{I}$ for sufficiently small $G(0)$. In particular, we show that the first entrant's value of upgrading is arbitrarily large when $G(0)$ is sufficiently small. Consider a firm that enters at $t = 0$ and immediately upgrades ($t_u = 0$). The net value of upgrading can be written as

$$\Delta V(0; 0, I) = \int_0^\infty e^{-\rho t} \left[\int_0^t \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(0) + \mu s} \mathrm{d}s \right] \mathrm{d}t - I.$$

For any t, the integral inside the first component is unbounded when $G(0) \to 0$. To see this,

$$\begin{aligned} \int_0^t \frac{e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}}{G(0) + \mu s} \mathrm{d}s &\geq \int_0^{t_1} \frac{e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}}{G(0) + \mu s} \mathrm{d}s \\ \geq \left(\min_{s\in[0,t]} e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)} \right) &\times \int_0^t \frac{1}{G(0) + \mu s} \mathrm{d}s, \end{aligned}$$

where the first minimum exists and is strictly positive because $e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}$ is a positive continuous function on $[0, t]$ bounded by (0,1]. In contrast, the second integral is unbounded when $G(0) \to 0$.

Finally, the characterization of equilibrium when $I \in [\underline{I}, \bar{I}]$ follows from the construction of $\underline{I}$ and $\bar{I}$ as well as Lemma OA.1.2. ■

Corollary OA.1.1 *If $I \in [\underline{I}, \bar{I}]$, firms that ever upgrade must do so no later than T. When $I = \underline{I}$, the last upgrader must upgrade at T. When $I \in (\underline{I}, \bar{I}]$, the last upgrader upgrades strictly before T.*

Proof. We now show that, if $t_u(t_e)$ is finite (i.e., the firm that entered at t_e upgrades), it must be below T. Suppose at T, the firm has not upgraded yet, i.e., $m(T;t_e) < m^*$. Since $I \geq \underline{I}$, we must have

$$m_{L,lr} = \frac{N}{G(T)\alpha_L} \leq \frac{(\rho+\alpha_L)c}{\alpha_L} < m^* = \frac{\rho(\rho+\alpha_H)}{\alpha_L-\alpha_H}I.$$

After T, $m(t;t_e)$ converges monotonically to $m_{L,lr}$ either from above or below, depending on whether $m(T;t_e)$ is larger or smaller than $m_{L,lr}$. Therefore, $m(t;t_e)$ can never reach m^*.

We denote the time of entry of the last upgrader by T_u and the time of upgrading of a firm that entered at any time t_e as $t_u(t_e)$. Therefore, the last upgrader's time of upgrading is $t_u(T_u)$. When $I = \underline{I}$, we have $m_{L,lr} = m^*$. Therefore, $m(t;T_u)$ crosses m^* at $t_u(T_u)$ from below. Because m^* is also the steady-state value of $m(t;T_u)$, $m(t;T_u)$ continues to rise to its peak and then converges monotonically to m^*. Since $t_u(T_u) < T$, we have $m(T;T_u) > m^*$. It is straightforward from equation (2) that $m(t_u;t_e)$ is continuous in t_e (customer arrival rates are bounded). Therefore, we can find $\delta > 0$ such that

$$m(T;T_u+\delta) > m^*.$$

Therefore, the entrant who enters at $T_u+\delta$ will also upgrade. This contradicts the definition of T_u. We must have $t_u(T_u) = T$.

When $I > \underline{I}$, $G(T)$ and $m_{L,lr}$ are unchanged but m^* is strictly higher, so $m_{L,lr} < m^*$. Suppose that the last upgrader upgrades at T, which implies that $m(T;T_u) = m^*$. We know that after T, $m(t;T_u)$ converges monotonically to $m_{L,lr}$. Therefore, $m(t;T_u) \leq m^*$ for all $t \geq T_u$, and T is a local minimum instead of a local maximum. We obtain a contradiction. Therefore, $t_u(T_u) < T$. ■

Before we prove (the second part of) Proposition 4, we prove the continuity of the upgrading time of the last upgrader, $t_u(T_u)$, with respect to μ.

Lemma OA.1.3 *Suppose $I < \bar{I}$. The time of upgrading of the last upgrader, $t_u(T_u)$, is continuous in μ.*

Proof. We first show that the entry time of the last upgrader, T_u, is continuous in μ. We discuss two cases, $I < \underline{I}$ and $I \in [\underline{I}, \bar{I})$. In the first case, $T_u = T = \frac{G(T;\mu)-G(0)}{\mu}$. Setting $F(G(T)) \equiv (\max_{t_u} \Delta V(t_u;T)) + V_L^e(T) = 0$, we see that μ does not affect the solution to $G(T)$. Therefore, T_u is continuous in μ.

When $I \in [\underline{I}, \bar{I})$, we solve T_u by

$$\Delta V(t_u(T_u, \mu); T_u, \mu) = 0,$$

where $t_u(t_e, \mu)$ solves

$$\begin{aligned} f(t_u; t_e, \mu) \equiv & m(t_u; t_e, \mu) - m^* \left(1 - \mathbf{1}\left(\frac{dm(t_u; t_e, \mu)}{dt_u} < 0\right)\right) \\ = & \int_{t_e}^{t_u} \frac{Ne^{-\alpha_L(t_u - s)}}{G(s)} \mathrm{d}s - \frac{\rho(\rho + \alpha_H)I}{\alpha_L - \alpha_H}\left(1 - \mathbf{1}\left(\frac{dm(t_u; t_e, \mu)}{dt_u} < 0\right)\right) \end{aligned}$$

The multiplier to m^* ensures that we solve for t_u when $m(t_u; t_e, \mu)$ crosses m^* from below. Note that we can write $G(s) = G(0) + \mu s$ because we know $t_u \le T$. Focusing on the rising segment of $m(t; t_e)$, the function $f(t_u; t_e, \mu)$ is continuously differentiable in t_u, t_e and μ. Consider ranges of t_e and μ so that firms ever upgrade. There must be a solution $t_u(t_e, \mu)$ such that $f(t_u; t_e, \mu) = 0$. In addition, $t_u(t_e, \mu)$ cannot be the single peak of the function so $f_{t_u}(t_u; t_e, \mu) > 0$. We can then apply the implicit function theorem and conclude that $t_u(t_e, \mu)$ is continuously differentiable.

We can then express the "indirect" value of upgrading by $\Delta V(t_u(t_e, \mu); t_e, \mu)$ as a function of t_e and μ. Since $t_u(t_e, \mu)$ is continuously differentiable in t_e and μ, the indirect value must be continuously differentiable in the two variables due to the chain rule. In addition, we can apply the chain rule and calculate

$$\begin{aligned} \frac{\partial \Delta V(t_u(t_e, \mu); t_e, \mu)}{\partial t_e} &= \frac{\partial \Delta V(t_u; t_e, \mu)}{\partial t_u}\Big|_{t_u = t_u(t_e, \mu)} \times \frac{\partial t_u(t_e, \mu)}{\partial t_e} + \frac{\partial \Delta V(t_u; t_e, \mu)}{\partial t_e}\Big|_{t_u = t_u(t_e, \mu)} \\ &= \frac{\partial \Delta V(t_u; t_e, \mu)}{\partial t_e}\Big|_{t_u = t_u(t_e, \mu)} \\ &= e^{-\rho(t_u - t_e)}\left(\rho \Delta \tilde{V}(t_u(t_e, \mu); t_e, \mu) + \frac{\Delta \tilde{V}(t_u; t_e, \mu)}{\partial t_e}\Big|_{t_u = t_u(t_e, \mu)}\right) \\ &= -e^{-\rho(t_u - t_e)} \times \frac{\alpha_L - \alpha_H}{(\rho + \alpha_H)(\rho + \alpha_L)} \frac{Ne^{-\alpha_L(t_u(t_e, \mu) - t_e)}}{G(t_e)} < 0 \end{aligned}$$

where we have applied an envelope theorem in the second equality, the relationship $\Delta V = e^{-\rho(t_u - t_e)} \Delta \tilde{V}$ in the third inequality and the formula (OA.1.1) and the fact that $\Delta \tilde{V}(t_u(t_e, \mu); t_e, \mu) = 0$ at $t_e = T_u$ in the fourth equality. Therefore, we can apply the implicit function theorem at an open neighborhood of μ and obtain a continuously differentiable function $T_u(\mu)$. Therefore, $t_u(T_u(\mu), \mu)$ must be a continuously differentiable (thus continuous) function of μ. ■

We also need to show the differentiability of the path $g_H(t; \mu) \equiv G_H(t; \mu)/G(t; \mu)$ with respect to μ. Note that the $G(0)$ initial entrants must simultaneously reach the threshold customer base m^*. Therefore, if they upgrade, they will do it at the same time. We denote this time point as t_{u0}, which depends on model parameters such as μ. We prove the following lemma:

Lemma OA.1.4 *The share of H-type firms at any time point t, $g_H(t;\mu)$, is differentiable in μ for $t > t_{u0}(\mu)$.*

Proof. We first characterize the $G_H(t)$ as the solution to an ordinary differential equation. We omit μ for now but it is understood that equilibrium objects depend on μ. We consider a firm that enters at t_e and upgrades at t_u, and another firm that enters at $t'_e \geq t_e$. Because the later entrant faces more competition, it takes longer for it to reach m^* and we must have $t'_u > t_u$. Denoting $\Delta t_e \equiv t'_e - t_e$ and $\Delta t_u \equiv t'_u - t_u$, we have

$$\begin{aligned}
&m(t_u;t_e) = m(t'_e;t_e)e^{-\alpha_L(t_u-t'_e)} + m(t_u;t'_e) \Rightarrow \\
&m(t_u;t'_e) \approx m(t_u;t_e) - \frac{Ne^{-\alpha_L(t_u-t_e)}}{G(t_e)}\Delta t_e \\
&m(t'_u;t'_e) = m(t_u;t'_e)(1-\alpha_L\Delta t_u) + \frac{N}{G(t_u)}\Delta t_u \Rightarrow \\
&m(t'_u;t'_e) - m(t_u;t_e) = -\alpha_L m(t_u;t_e)\Delta t_u - \frac{Ne^{-\alpha_L(t_u-t_e)}}{G(t_e)}\Delta t_e + \frac{N}{G(t_u)}\Delta t_u
\end{aligned} \tag{OA.1.4}$$

To obtain an ODE for $G_H(t)$, we use the following relationships

$$G_H(t_u) = G_0 + \mu t_e, \quad G_H(t'_u) = G_0 + \mu t'_e.$$

From these equations, we can easily back out $t_e = \frac{G_H(t_u)-G_0}{\mu}$ given $G_H(t_u)$. When $t'_e > t_e$, we can also express Δt_e as

$$\Delta t_e = \frac{G_H(t'_u) - G_H(t_u)}{\mu} \approx \frac{G'_H(t_u)\Delta t_u}{\mu}. \tag{OA.1.5}$$

Since $m(t'_u;t'_e) = m(t_u;t_e) = m^*$, we set the value of equation (OA.1.4) to zero and combine it with equation (OA.1.5) and $G_H(t_u) = G(t_e)$ to solve

$$G'_H(t)\left[\frac{Ne^{-\alpha_L(t-t_e)}}{\mu G_H(t)}\right] = \frac{N}{G(t)} - \frac{\alpha_L}{\alpha_L-\alpha_H}\left(\rho(\rho+\alpha_H)I\right). \tag{OA.1.6}$$

where

$$t_e = \frac{G(t_e)-G(0)}{\mu} = \frac{G_H(t)-G(0)}{\mu},$$

and the initial condition is $G_H(t_{u0}(\mu)) = G(0)$.

We can rewrite the ODE as $G'_H = f(t, G_H, t_{u0}(\mu), \mu)$ with the initial condition $G_H(t_{u0}(\mu)) = G(0)$. We can first ignore the dependence of t_{u0} on μ and consider the less restrictive version:

$$G'_H = f(t, G_H, t_{u0}, \mu), \quad G_H(t_{u0}) = G(0).$$

Since f is continuously differentiable in all four elements, $G_H(t, t_{u0}, \mu)$ must be continuously differentiable. In addition, $t_{u0}(\mu)$ must be differentiable with respect to μ because t_{u0} is the solution to

$$m(t_{u0}; 0, \mu) = \int_0^{t_{u0}} \frac{Ne^{-\alpha_L(t_{u0}-s)}}{G(s)} \mathrm{d}s = m^*.$$

Therefore, $G_H(t, G_H, t_{u0}(\mu), \mu)$ must be differentiable with respect to μ due to the chain rule. ■

Proof of Proposition 4. In this proof, we denote the steady-state mass of firms $G(T)$ as G_{ss}. Correspondingly, we denote the steady-state mass of H-type firms as $G_{H,ss}$.

Part 1: Consider two different market entry rates $\mu_1 < \mu_2$. We discuss the impact of raising μ_1 to μ_2 in two scenarios: (1) $G_{H,ss}(\mu_2) = 0$ and (2) $G_{H,ss}(\mu_2) > 0$. In scenario (1), since $G_{H,ss}(\mu_1) \geq 0$, it is trivial that $G_{H,ss}(\mu_1) \geq G_{H,ss}(\mu_2)$. In scenario (2), denote the entry time of the last upgrader as $T_u(\mu_2) \geq 0$. We prove that a firm that enters at the same time under μ_1 will upgrade.

First, note that there will be entry at $t = T_u(\mu_2)$ under μ_1. This is because mu does not affect the firm's profit after the last firm enters; therefore, it does not affect G_{ss}, whether the last firm upgrades or not. With lower market entry rates, it takes longer for the economy to reach the steady state, $T(\mu_1) > T(\mu_2)$. Since $T_u(\mu_2) \leq T(\mu_2)$, we must have $T_u(\mu_2) < T(\mu_1)$. Next, we show that the value of upgrading under μ_1 is non-negative for entrants at $T_u(\mu_2)$. To see this, note that the equilibrium under μ_2 has $G_{H,ss}(\mu_2)$ entrants at $t = T_u(\mu_2)$. Consider an entrant who enters at

$$t_e = \frac{G_{H,ss}(\mu_2) - G(0)}{\mu_1},$$

under the new entry rate μ_1. At this point, the economy has $G_{H,ss}(\mu_2)$ firms, the same as the number of firms up to $T_u(\mu_2)$ in the equilibrium under μ_2. In the equilibrium under μ_1, $G(t)$ rises from $G_{H,ss}(\mu_2)$ to G_{ss} at the rate μ_1 from t_e onward, while in the equilibrium under μ_2, $G(t)$ rises from $G_{H,ss}(\mu_2)$ to G_{ss} at the rate μ_2 from $T_u(\mu_2)$ onward. Therefore, from the perspective of the focal entrants at t_e, the future competition is stronger under μ_2 and net gains from upgrading are higher under μ_1. Therefore, the entrant at time t_e will upgrade in the equilibrium with μ_1. By Lemma OA.1.2, we must have

$$G_{H,ss}(\mu_1) \geq G_{H,ss}(\mu_2).$$

The inequality is strict if $G_{H,ss}(\mu_2) < G_{ss}$ because in this case, the future path under μ_2 has a strictly increasing segment, and the focal entrant under μ_1 has a strictly higher value of upgrading.[17]

Consider two different flow fixed costs, $c_1 < c_2$. We compare the equilibrium in three cases: (1) $G_{H,ss}(c_1)/G_{ss}(c_1) = 1$ (2) $G_{H,ss}(c_1)/G_{ss}(c_1) \in (0,1)$, and (3) $G_{H,ss}(c_1) = 0$. When $G_{H,ss}(c_1)/G_{ss}(c_1) = 1$, we must have $I < \underline{I}(c_1)$, where $\underline{I}(c_1)$ is the same cutoff defined in equation (13) with c_1 as an explicit parameter. Since $\underline{I}(c)$ strictly increases with c, we must have $I < \underline{I}(c_1) < \underline{I}(c_2)$. Therefore, $G_{H,ss}(c_2)/G_{ss}(c_2) = 1 = G_{H,ss}(c_1)/G_{ss}(c_1)$.

In the second scenario, early entrants upgrade while late ones do not under c_1. We can follow the same strategy when analyzing the comparative statics with respect to μ. Denote the entry time of the last upgrader with $t_e \equiv T_u(c_1) < T(c_1)$. There are three sub-cases: (2-a) there is no entry at t_e under c_2, i.e., $t_e \geq T(c_2)$, (2-b) there is still entry at t_e under c_2, i.e., $t_e < T(c_2)$, and the last entrant does not upgrade, and (2-c) there is still entry at t_e and the last entrant upgrades. In case (2-a), consider the last entrant under c_2. It faces constant $G(t) = G_{ss}(c_2)$ from $T(c_2)$ onward. Under both c_1 and c_2, the market entry rate is the same thus $G(t_e) \geq G_{ss}(c_2)$. Therefore, the firm that enters at t_e under c_1 faces tougher competition over its life cycle than the last entrant under c_2. Since the former upgrades, the latter must upgrade and $G_{H,ss}(c_2)/G_{ss}(c_2) = 1 > G_{H,ss}(c_1)/G_{ss}(c_1)$. In case (2-b), since $c_1 < c_2$, and the last entrant does not upgrade, by equation (5), we must have $G_{ss}(c_1) > G_{ss}(c_2)$. Therefore, the firm that enters at t_e under c_2 faces less competition over its life cycle (for some part strictly less) than the firm that enters at t_e under c_1. This gives stronger incentive for it to upgrade, $\max_{t_u} \Delta V(t_u; t_e, c_2) > 0$. Therefore, $G_{H,ss}(c_2) \geq G_{H,ss}(c_1)$, and

$$\frac{G_{H,ss}(c_2)}{G_{ss}(c_2)} > \frac{G_{H,ss}(c_1)}{G_{ss}(c_1)}.$$

Case (2-c) is trivial because by construction, $G_{H,ss}(c_2)/G_{ss}(c_2) = 1 > G_{H,ss}(c_1)/G_{ss}(c_1)$.

In the third scenario, it is trivial that $G_{H,ss}(c_2)/G_{ss}(c_2) \geq 0 = G_{H,ss}(c_1)/G_{ss}(c_1)$. In addition, since $G_{ss}(c_2) < G_{ss}(c_1)$, it is possible that the reduction in $G_{H,ss}$ is large and gives

[17] Formally, the value of upgrading is positive if the entrant at time t_e under μ_1 upgrades at the same age as the entrant at time $T_u(\mu_2)$, i.e., $\Delta V(t_e + t_u(T_u(\mu_2); \mu_2) - T_u(\mu_2); t_e, \mu_1) > 0$. Due to the continuity of $\Delta V(t_u; t_e)$ in both t_u and t_e, in a small neighborhood $t'_e \in (t_e, t_e + \Delta t)$, we have $\Delta V(t_e + t_u(T_u(\mu_2); \mu_2) - T_u(\mu_2); t'_e, \mu_1) > 0$.

firms enough incentives to upgrade so $G_{H,ss}(c_2) > 0$ and the inequality becomes strict.

To prove the remainder of the proposition, note that the steady-state market size can be written as

$$M_{lr} = G_{H,ss}m_{H,lr} + G_{L,ss}m_{L,lr} = \frac{(\alpha_L - \alpha_H)N}{\alpha_H \alpha_L}\frac{G_{H,ss}}{G_{ss}} + \frac{N}{\alpha_L},$$

where we have used $m_{i,lr} = \frac{N}{G_{ss}\alpha_i}, i \in \{H, L\}$. The total market size (customers served) is strictly increasing in $\frac{G_{H,ss}}{G_{ss}}$. The average firm size equals $\frac{M_{lr}}{G_{ss}}$. Applying the formula of G_{ss} in equation (5), we obtain

$$\frac{M_{lr}}{G_{ss}} = \frac{(\alpha_L - \alpha_H)(\rho + \alpha_L)c}{\alpha_H \alpha_L}\frac{G_{H,ss}}{G_{ss}} + \frac{(\rho + \alpha_L)c}{\alpha_L},$$

which also increases in $\frac{G_{H,ss}}{G_{ss}}$.

Part 2: we first prove that $\frac{\partial g_H(t;\mu)}{\partial \mu} < 0$ for $t > t_{u0}(\mu)$. Denote the last upgrader's time of upgrading under entry rate μ as $t_u(\mu)$. We know from Corollary OA.1.1 that $t_u(\mu) < T(\mu)$. Due to the continuity of $t_u(\mu)$ shown in Lemma OA.1.3, we can find sufficiently small $\Delta\mu$, such that $t_u(\mu - \Delta\mu) < T(\mu)$. We consider two cases $t_u(\mu - \Delta\mu) < t_u(\mu)$ and $t_u(\mu - \Delta\mu) > t_u(\mu)$. The knife edge case $t_u(\mu - \Delta\mu) = t_u(\mu)$ can be proved similarly.

First, suppose $t_u(\mu - \Delta\mu) < t_u(\mu)$. We can divide the entire range of time into four segments and discuss the relative size of $g_H(t;\mu)$ and $g_H(t;\mu - \Delta\mu)$. Note that $t_{u0}(\mu) > t_{u0}(\mu - \Delta\mu)$ because it takes more time to reach m^* when entry is faster and market competition is stronger.

1. When $t \in [0, t_{u0}(\mu - \Delta\mu))$, we have $g_H(t;\mu - \Delta\mu) = g_H(t;\mu) = 0$.

2. When $t \in [t_{u0}(\mu - \Delta\mu), t_{u0}(\mu))$, we have $g_H(t;\mu - \Delta\mu) > 0$ and $g_H(t;\mu) = 0$ so $g_H(t;\mu - \Delta\mu) > g_H(t;\mu)$.

3. When $t \in [t_{u0}(\mu), t_u(\mu - \Delta\mu)$, both $G_H(t;\mu - \Delta\mu)$ and $G_H(t;\mu)$ are rising. However, since the accumulation of customer capital is faster under $\mu - \Delta\mu$, it takes less time for an entrant to reach m^*. Suppose the time of entry corresponding to an upgrader at time t is $t_u^{-1}(t;\mu)$ and $t_u^{-1}(t;\mu - \Delta\mu)$, where $t_u^{-1}(\cdot)$ is the inverse function of $t_u(t_e)$. We must have

$$t_u^{-1}(t;\mu) < t_u^{-1}(t;\mu - \Delta\mu).$$

Therefore, at time t, we must have

$$g_H(t;\mu-\Delta\mu) = \frac{G(0)+(\mu-\Delta\mu)t_u^{-1}(t;\mu-\Delta\mu)}{G(0)+(\mu-\Delta\mu)t}$$
$$> \frac{G(0)+(\mu-\Delta\mu)t_u^{-1}(t;\mu)}{G(0)+(\mu-\Delta\mu)t} > \frac{G(0)+\mu t_u^{-1}(t;\mu)}{G(0)+\mu t} = g_H(t;\mu)$$

4. When $t \in [t_u(\mu-\Delta\mu), \infty)$, we have $G_H(t;\mu-\Delta\mu) = G_{H,ss}(\mu-\Delta\mu) > G_{H,ss}(\mu) \geq G_H(t;\mu)$. Meanwhile, $G(t;\mu-\Delta\mu) \leq G(t;\mu)$. Therefore,

$$g_H(t;\mu-\Delta\mu) > g_H(t;\mu).$$

Second, suppose $t_u(\mu-\Delta\mu) > t_u(\mu)$. Figure OA.1.1 illustrates such a case. We divide the entire range of time into five segments as

1. When $t \in [0, t_{u0}(\mu-\Delta\mu))$, we have $g_H(t;\mu-\Delta\mu) = g_H(t;\mu) = 0$.
2. When $t \in [t_{u0}(\mu-\Delta\mu), t_{u0}(\mu))$, we have $g_H(t;\mu-\Delta\mu) > 0$ and $g_H(t;\mu) = 0$ so $g_H(t;\mu-\Delta\mu) > g_H(t;\mu)$.
3. When $t \in [t_{u0}(\mu), t_u(\mu))$, both $G_H(t;\mu-\Delta\mu)$ and $G_H(t;\mu)$ are rising. We can apply the same argument as in the previous case #3 for $t \in [t_{u0}(\mu), t_u(T_u(\mu-\Delta\mu)))$ and obtain $g_H(t;\mu-\Delta\mu) > g_H(t;\mu)$.
4. When $t \in [t_u(\mu), t_u(\mu-\Delta\mu)$, we have $G_H(t;\mu) = G_{H,ss}(\mu)$ and $G_H(t;\mu-\Delta\mu) > G_H(t_u(\mu);\mu-\Delta\mu)$. At $t = t_u(\mu)$, we have

$$g_H(t_u(\mu);\mu-\Delta\mu) = \frac{G_H(t_u(\mu);\mu-\Delta\mu)}{G(t_u(\mu);\mu-\Delta\mu)} > \frac{G_H(t_u(\mu);\mu)}{G(t_u(\mu);\mu)} = g_H(t_u(\mu);\mu).$$

For $t > t_u(\mu)$, we can write

$$\begin{aligned} g_H(t; \mu - \Delta\mu) &> g_H(t_u(\mu); \mu - \Delta\mu) \times \frac{G(t_u(\mu); \mu - \Delta\mu)}{G(t; \mu - \Delta\mu)} \\ &= g_H(t_u(\mu); \mu - \Delta\mu) \times \frac{G(0) + (\mu - \Delta\mu)t_u(\mu)}{G(0) + (\mu - \Delta\mu)t} \\ &= g_H(t_u(\mu); \mu - \Delta\mu) \times \frac{G(0) + \mu t_u(\mu)}{G(0) + \mu t} \\ &= g_H(t_u(\mu); \mu - \Delta\mu) \times \frac{G(t_u(\mu); \mu)}{G(t; \mu)} \\ &> g_H(t_u(\mu); \mu) \times \frac{G(t_u(\mu); \mu)}{G(t; \mu)} \\ &> g_H(t; \mu) \times \frac{G(t_u(\mu); \mu)}{G(t; \mu)} \end{aligned}$$

The last inequality holds because $G_H(t; \mu)$ has reached its steady state while $G(t; \mu)$ is still rising.

5. When $t \in [t_u(\mu - \Delta\mu), \infty)$, both $G_H(t; \mu - \Delta\mu)$ and $G_H(t; \mu)$ have reached their steady states, with $G_{H,ss}(\mu) < G_{H,ss}(\mu - \Delta\mu)$ according to Proposition 4. Meanwhile, we know that $G(t; \mu) \geq G(t; \mu - \Delta\mu)$ for all t. Therefore, $g_H(t; \mu - \Delta\mu) > g_H(t; \mu)$.

We have proved that in a small neighborhood of μ, $g_H(t; \mu)$ is decreasing. To show that the property holds "globally", we apply the differentiability of $g_H(t; \mu)$ with respect to μ from Lemma OA.1.4. Therefore, we must have

$$\frac{\partial g_H(t; \mu)}{\partial \mu} < 0.$$

For two arbitrary values of μ_1 and μ_2, we know that $g_H(t; \mu_1) = g_H(t; \mu_2)$ for $t < t_{u0}(\mu_2)$, and $g_H(t; \mu_1) > g_H(t; \mu_2)$ for $t \in [t_{u0}(\mu_1), t_{u0}(\mu_2)]$. For $t > t_{u0}(\mu_2)$, we can apply the negative derivative and obtain

$$g_H(t; \mu_2) = g_H(t; \mu_1) + \int_{\mu_1}^{\mu_2} \frac{\partial g_H(t; \mu)}{\partial \mu} d\mu < g_H(t; \mu_1).$$

We now compare the fraction of H-type firms under different flow fixed costs, $c_1 < c_2$. We denote the entry and upgrading time of the last upgrader in market 1 as $T_u(c_1)$ and $t_u(c_1)$, respectively. We know that competition in market 1 is stronger than in market 2 (strictly stronger after market 2 reaches its steady state at $T(c_2)$). This has two implications. First, the value of upgrading, $\Delta V(t_u; t_e)$ is higher in market 2 for any (t_u, t_e). Therefore, if an entrant in market 1 upgrades, it must upgrade in market 2. Second, it takes less time for any firm in market 2 to reach a customer base of m^*, the necessary condition for upgrading. Therefore, if a firm in market 1 entered at time t_e and has upgraded by time t_u, the entrant at t_e in market 2 must have upgraded at or before t_u. For $t \le t_u(c_1)$, we must have

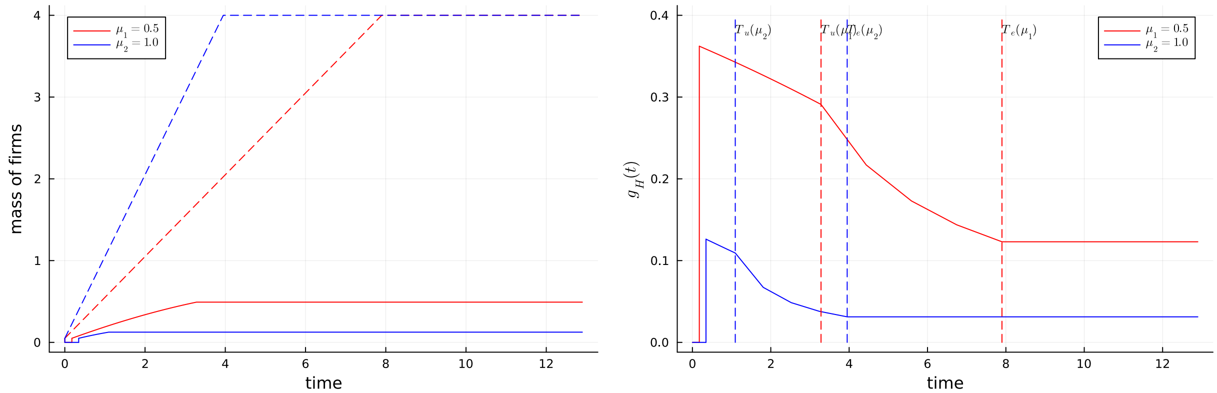
$$G_H(t; c_2) \ge G_H(t; c_1).$$

For $t > t_u(c_1)$, $G_H(t; c_1) = G_{H,ss}(c_1)$ is constant and $G_H(t; c_2)$ may further increase, so $G_H(t; c_2) \ge G_H(t; c_1)$ also holds in this range.

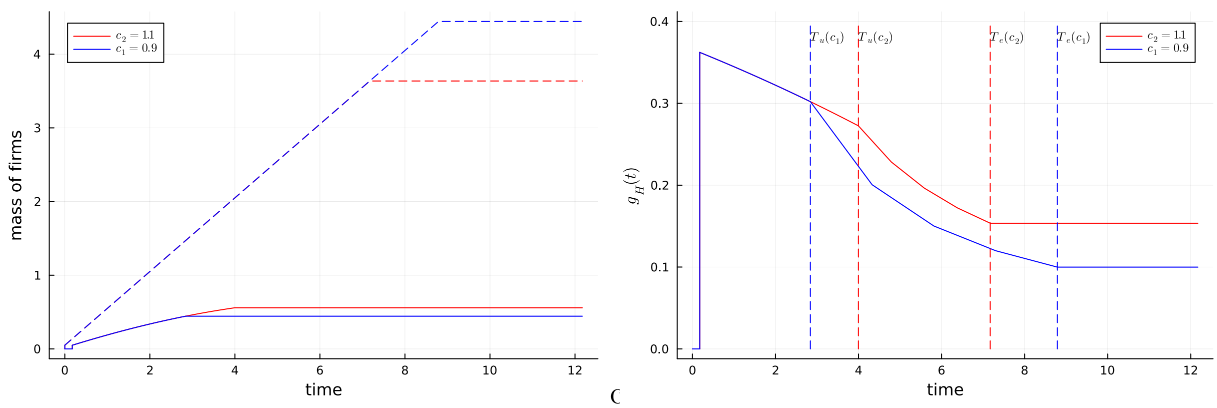
On the other hand, $G(t; c_1) \ge G(t; c_2)$ and the inequality becomes strict when $t > T(c_2)$. Therefore, we must have $g_H(t; c_1) \le g_H(t; c_2)$ and the inequality is strict for some t. ■

In Figure OA.1.1, We compare the paths of $G_H(t)$, $G(t)$ and $g_H(t)$ under two different values of μ ($\mu_1 = 0.3 < \mu_2 = 0.5$). As is shown in Panel (b), $g_H(t)$ jumps from zero to a positive value slightly earlier under the lower market entry rate. After that, $g_H(t; \mu_1)$ is always higher than $g_H(t; \mu_2)$. After the last entrant enters under μ_1, both paths converge to their steady states, and $g_{H,ss}(\mu_1) > g_{H,ss}(\mu_2)$, consistent with Proposition 4.

Figure OA.1.1: Comparative dynamics with respect to entry rate, μ



(a) Mass of Firms and μ (b) Share of H-type Firms and μ



(c) Mass of Firms and c (d) Share of H-type Firms and c

Notes. Other parameter values $\rho = 0.1, N = 1, \alpha_H = 0.1, \alpha_L = 0.15, I = 5, G_0 = 0.05$. We set $c = 1$ when varying μ and $\mu = 0.5$ when varying c.

OA.1.3 Proofs of Results Related to Entry

Derivative-based Proof of Proposition 5. We take the derivative of the value of a permanent L-type firm, V_L^e, with respect to the flow fixed cost, c. The value of entry of a permanent L-type firm is

$$V_L^e(t_e) = \int_{t_e}^{\infty} e^{-\rho(t-t_e)} \left(\int_{t_e}^{t} \frac{N}{G(s)} e^{-\alpha_L(t-s)} \mathrm{d}s \right) \mathrm{d}t - \frac{c}{\rho}. \tag{OA.1.1}$$

Alternatively, we can express the integration using firm age (time relative to entry) instead of absolute time

$$V_L^e(t_e) = \int_{0}^{\infty} e^{-\rho t} \left(\int_{0}^{t} \frac{N e^{-\alpha_L(t-s)}}{G(t_e + s)} \mathrm{d}s \right) \mathrm{d}t - \frac{c}{\rho}.$$

One immediate result from this expression is

$$\frac{\partial V_L^e(t_e)}{\partial t_e} \leq 0,$$

because $G(s+t_e)$ decreases with t_e. The derivative with respect to t_e is strict before the steady state, i.e., $s < T - t_e$.

We can further decompose the above expression into three terms, V_{L1}^e, V_{L2}^e and V_{L3}^e:

$$\begin{aligned}
V_{L1}^e &= \int_0^{T-t_e} e^{-\rho t} \left(\int_0^t \frac{N}{G(s)} e^{-\alpha_L (t-s)} \mathrm{d}s \right) \mathrm{d}t, \\
V_{L2}^e &= \int_{T-t_e}^{\infty} e^{-\rho t} \left(\int_0^{T-t_e} \frac{N}{G(s)} e^{-\alpha_L (t-s)} \mathrm{d}s \right) \mathrm{d}t, \\
V_{L3}^e &= \int_{T-t_e}^{\infty} e^{-\rho t} \left(\int_{T-t_e}^{t} \frac{N}{G(T)} e^{-\alpha_L (t-s)} \mathrm{d}s \right) \mathrm{d}t.
\end{aligned}$$

Taking the derivative of each term with respect to c, we have

$$\begin{aligned}
\frac{\mathrm{d}V_{L1}^e}{\mathrm{d}c} &= \frac{\mathrm{d}T}{\mathrm{d}c} \cdot e^{-\rho(T-t_e)} \int_0^{T-t_e} \frac{N e^{-\alpha_L (T-t_e-s)}}{G(0)+\mu(s+t_e)} \mathrm{d}s \\
\frac{\mathrm{d}V_{L2}^e}{\mathrm{d}c} &= -\frac{\mathrm{d}T}{\mathrm{d}c} e^{-\rho(T-t_e)} \int_0^{T-t_e} \frac{N e^{-\alpha_L (T-t_e-s)}}{G(0)+\mu(s+t_e)} \mathrm{d}s + \int_{T-t_e}^{\infty} e^{-\rho t} \frac{\mathrm{d}T}{\mathrm{d}c} \frac{N e^{-\alpha_L (t-(T-t_e))}}{G(T)} \mathrm{d}t \\
\frac{\mathrm{d}V_{L3}^e}{\mathrm{d}c} &= -\frac{\mathrm{d}T}{\mathrm{d}c} e^{-\rho(T-t_e)} \cdot 0 + \int_{T-t_e}^{\infty} e^{-\rho t} \left(-\frac{\mathrm{d}T}{\mathrm{d}c} \right) \frac{N e^{-\alpha_L (t-(T-t_e))}}{G(T)} \mathrm{d}t \\
&\quad + \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \int_{T-t_e}^{\infty} e^{-\rho t} \left(\int_{T-t_e}^{t} e^{-\alpha_L (t-s)} \mathrm{d}s \right) \mathrm{d}t
\end{aligned}$$

Therefore, we obtain

$$\frac{\mathrm{d}(V_{L1}^e + V_{L2}^e + V_{L3}^e)}{\mathrm{d}c} = \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \int_{T-t_e}^{\infty} e^{-\rho t} \left(\int_{T-t_e}^{t} e^{-\alpha_L (t-s)} \mathrm{d}s \right) \mathrm{d}t = \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \frac{e^{-\rho(T-t_e)}}{\rho(\rho+1-\alpha_L)}.$$

The above derivation suggests that net benefits caused by a marginal change in T are zero. Intuitively, a small increase in c induces a small reduction in T, but it reallocates profits between V_{L1}^e, V_{L2}^e, and V_{L3}^e, and has no first-order effect on the sum of the three, due to the continuity of the firm's revenue with respect to time. Eventually, the benefit of a higher c solely comes from the higher profits after the market is saturated, $t = T$. This benefit is discounted more when firms enter earlier.

We know that the steady-state customer arrival rate is $N/G(T) = (\rho + \alpha_L)c$. Since $V_L^e(t_e) = V_{L1}^e + V_{L2}^e + V_{L3}^e - c/\rho$, we can show that the benefit from higher steady-state profits is dominated by the cost:

$$\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} = \frac{e^{-\rho(T-t_e)}-1}{\rho} \leq 0.$$

Therefore, the value of entry decreases with c, and the decrease is strict when $t_e < T$.

It is also straightforward to compare the value of entry in two markets with a large gap in c. Suppose there are two markets, $c_1 < c_2$. We know that it takes longer for market 1 to become saturated than market 2, i.e., $T_1 > T_2$. For $t_e < T_2$, we must have

$$V_L^e(t_e;c_2) - V_L^e(t_e;c_1) = \int_{c_1}^{c_2} \frac{\mathrm{d}V_L^e(t_e;c)}{\mathrm{d}c}\mathrm{d}c < 0,$$

because the integrand is negative on $[c_1, c_2]$. When $t_e \in [T_2; T_1)$, we know that $V_L^e(t_e;c_1) > V_L^e(T_1;c_1) = 0 = V_L^e(t_e;c_2)$, where the first inequality comes from the fact that the value of entry decreases strictly with time before market saturation. ■

Proof of Proposition 6. In the proof of Proposition 5, we have shown how the value of entry of a permanently L-type firm varies with c. We now examine the impact of c on the actual value of entry, $V^e(t_e)$, when upgrading is possible.

The additional value of upgrading at t_u discounted to time t_e can be written as

$$\begin{aligned}
\Delta V(t_u;t_e) &= \frac{\alpha_L - \alpha_H}{(\rho+\alpha_H)(\rho+\alpha_L)} e^{-\rho(t_u-t_e)} m(t_u;t_e) \\
&\quad + \int_{t_u}^{\infty} e^{-\rho(t-t_e)} \left[\int_{t_u}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(s)} \mathrm{d}s \right] \mathrm{d}t - e^{-\rho(t_u-t_e)} I \\
&= \frac{\alpha_L - \alpha_H}{(\rho+\alpha_H)(\rho+\alpha_L)} e^{-\rho t_u'} m(t_u'+t_e;t_e) \\
&\quad + \int_{t_u'}^{\infty} e^{-\rho t} \left[\int_{t_u'}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(s+t_e)} \mathrm{d}s \right] \mathrm{d}t - e^{-\rho t_u'} I,
\end{aligned} \tag{OA.1.2}$$

where we obtain the second equality by rewriting the integral in the first one using firm age, and $t_u' = t_u - t_e$ is the age of the firm when it upgrades. We denote $t_u'(t_e, c)$ as the age of a firm born at t_e in a market with fixed cost c when it optimally upgrades.

Note that the cutoff $\underline{I}$ is linear in c following the formula (13), the assumption that the last entrant does not upgrade implies that

$$I \geq \underline{I}(c_2) \geq \underline{I}(c) \geq \underline{I}(c_1), \forall c \in [c_1, c_2].$$

Therefore, we have the partial upgrading case for all $c \in [c_1, c_2]$. From Corollary OA.1.1, we know that upgrading must happen at or before $T(c_2)$.

We now consider a firm that upgrades at $t_u \leq T(c_2)$. We can take the derivative of $\Delta V(t'_u(t_e, c), t_e; c)$ with respect to the parameter c:

$$\frac{\mathrm{d}\Delta V(t'_u(t_e, c), t_e; c)}{\mathrm{d}c} = \left.\frac{\partial \Delta V(t'_u, t_e; c)}{\partial c}\right|_{t'_u = t'_u(t_e, c)}, \tag{OA.1.3}$$

which is immediate from the envelope theorem.[18] We can then calculate this derivative by taking the derivative of each term on the RHS of equation (OA.1.2). The first and third terms are not directly affected by c. For the second term, we use the same strategy when we take the derivative of V_L^e in the proof of Proposition 5. We decompose it into three terms, $\Delta V_1, \Delta V_2$, and ΔV_3, by breaking the integral at $t = T - t_e$:

$$\begin{aligned}
\Delta V_1 &\equiv \int_{t'_u}^{T-t_e} e^{-\rho t} \left[\int_{t'_u}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(s+t_e)} \mathrm{d}s \right] \mathrm{d}t \\
\Delta V_2 &\equiv \int_{T-t_e}^{\infty} e^{-\rho t} \left[\int_{t'_u}^{T-t_e} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(s+t_e)} \mathrm{d}s \right] \mathrm{d}t \\
\Delta V_3 &\equiv \int_{T-t_e}^{\infty} e^{-\rho t} \left[\int_{T-t_e}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(T)} \mathrm{d}s \right] \mathrm{d}t
\end{aligned} \tag{OA.1.4}$$

Similar to the earlier derivation of $\mathrm{d}V_{Li}^e/\mathrm{d}c\,, i = \{1, 2, 3\}$, c does not directly affect $\Delta V(t'_u(t_e), t_e)$ through $T - t_e$. It directly affects $\Delta V(t'_u(t_e), t_e)$ only through the speed of customer accumulation after T. Therefore, we have

$$\begin{aligned}
\frac{\mathrm{d}\Delta V(t'_u(t_e), t_e)}{\mathrm{d}c} &= \frac{\partial}{\partial c} \int_{T-t_e}^{\infty} e^{-\rho t} \left[\int_{T-t_e}^{t} \frac{N\left(e^{-\alpha_H(t-s)} - e^{-\alpha_L(t-s)}\right)}{G(T)} \mathrm{d}s \right] \mathrm{d}t \\
&= \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \times \frac{(\alpha_L - \alpha_H)e^{-\rho(T-t_e)}}{\rho(\rho+\alpha_H)(\rho+\alpha_L)}.
\end{aligned}$$

[18]When we evaluate the derivative at $c = c_2$, we calculate the left-hand derivative so that the breakdown of the integral at $T - t_e$ is well-defined: a smaller c ensures that $t_u(t_e, c) - t_e < T(c) - t_e$ for all upgraders.

Combining this with the earlier expression of $\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c}$, we have

$$\begin{aligned}\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c} &= \frac{\mathrm{d}\Delta V(t_u'(t_e), t_e)}{\mathrm{d}c} + \frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}c} \\ &= \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \times \frac{(\alpha_L - \alpha_H)e^{-\rho(T-t_e)}}{\rho(\rho+\alpha_H)(\rho+\alpha_L)} + \frac{\mathrm{d}(N/G(T))}{\mathrm{d}c} \times \frac{e^{-\rho(T-t_e)}}{\rho(\rho+\alpha_L)} - \frac{1}{\rho} \\ &= \frac{\frac{\rho+\alpha_L}{\rho+\alpha_H}e^{-\rho(T-t_e)} - 1}{\rho}. \end{aligned} \tag{OA.1.5}$$

Intuitively, less competition in the steady state increases the benefit of upgrading; however, this benefit is only realized after the firm has actually upgraded.

We now find sufficient conditions for $\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c}$ to be positive and ensure that there exists $\bar{I}$ such that $I < \bar{I}$. We set $t_e = 0$ so that the derivative must be positive for $t_e \in (0, T_u]$. Note that

$$T = \frac{G(T) - G(0)}{\mu} = \frac{\frac{N}{(\rho+\alpha_L)c} - G(0)}{\mu} < \frac{N}{\mu(\rho+\alpha_L)c}.$$

Substituting this upper bound into $\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c}$ and letting it be positive, we obtain a sufficient condition assuming $I < \bar{I}$

$$\mu > \frac{\rho N}{(\rho+\alpha_L)c \cdot \log\left(\frac{\rho+\alpha_L}{\rho+\alpha_H}\right)}.$$

We now need to find conditions such that $I < \bar{I}$ (partial upgrading). We know from Proposition 3 that $\underline{I}$ is a function of $(\rho, \alpha_L, \alpha_H, c)$. Given μ, we can find a sufficiently small $G(0)$ such that $\bar{I} > I$. When this holds, all $I \in [\underline{I}, \bar{I})$ satisfy the restriction of partial upgrading and we must have $\frac{\mathrm{d}V^e(t_e)}{\mathrm{d}c} > 0$.

Finally, the difference in the value of entry in markets 2 and 1 for an early entrant that upgrades in both markets is

$$V^e(t_e; c_2) - V^e(t_e; c_1) = \int_{c_1}^{c_2} \frac{\frac{\rho+\alpha_L}{\rho+\alpha_H}e^{-\rho(T(c)-t_e)} - 1}{\rho}\mathrm{d}t > 0$$

under our parameter restrictions. However, for a late entrant that does not upgrade in either market, we must have

$$V^e(t_e; c_2) - V^e(t_e; c_1) = \int_{c_1}^{c_2} \frac{e^{-\rho(T(c)-t_e)} - 1}{\rho}\mathrm{d}t < 0.$$

Therefore, the early entrant enters the high-barrier market, while the late entrant prefers the low-barrier one. ■

Proof of Proposition 7.

Part 1 We first show that the time of entry of the last entrant, T_j, must be the same for $j = 1, 2$. To see this, consider the case when $T_1 > T_2$. Then the last entrant in market 2 has a strictly lower value of entry at T_2 compared to that in market 1. Therefore, it would have entered market 1 instead of market 2. This leads to a contradiction. We can rule out $T_1 < T_2$ similarly. Therefore, we must have $T_1 = T_2$.

In the no-upgrading case, the last entrant does not upgrade, and the steady-state mass of firms in market j does not depend on the market-specific entry rates and must have the following expression:

$$G_{ss,j} = \frac{N}{(\rho + \alpha_L)c_j}.$$

We can also use the steady-state mass of firms and total entry rate to calculate T_j:

$$T_j = \frac{G(T; c_1) + G(T; c_2) - G(0)}{\mu} \equiv T. \tag{OA.1.6}$$

To understand the equilibrium entry rates or the paths $G(t; c_j)$, we first check how the value of entry changes with respect to the time of entry in each market by calculating $\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}t_e}$. Using the expression of $V_L^e(t_e)$ expressed in absolute time (see Online Appendix equation (OA.1.1)), we have

$$\begin{aligned}
\frac{\mathrm{d}V_L^e(t_e)}{\mathrm{d}t_e} &= \frac{\mathrm{d}}{\mathrm{d}t_e}\left[\int_{t_e}^{\infty} e^{-\rho(t-t_e)}\left(\int_{t_e}^{t} \frac{N}{G(s)} e^{-\alpha(t-s)}\mathrm{d}s\right)\mathrm{d}t - \frac{c}{\rho}\right] \\
&= \int_{t_e}^{\infty} \rho e^{-\rho(t-t_e)}\left(\int_{t_e}^{t} \frac{N}{G(s)} e^{-\alpha(t-s)}\mathrm{d}s\right)\mathrm{d}t - \int_{t_e}^{\infty} e^{-\rho(t-t_e)}\frac{N}{G(t_e)} e^{-\alpha(t-t_e)}\mathrm{d}t \\
&= \rho\left(V_L^e(t_e) + \frac{c}{\rho}\right) - \frac{N}{(\rho+\alpha)G(t_e)}
\end{aligned}$$

The difference in the value of the derivative in the two markets is

$$\frac{\mathrm{d}V^e(t_e; c_2)}{\mathrm{d}t_e} - \frac{\mathrm{d}V^e(t_e; c_1)}{\mathrm{d}t_e} = c_2 - c_1 - \left(\frac{N}{(\rho+\alpha)G(t_e; c_2)} - \frac{N}{(\rho+\alpha)G(t_e; c_1)}\right).$$

We know that $V^e(t_e; c_2) = V^e(t_e; c_1) = 0$ at $t_e = T$. Substituting in the expressions of $G_{ss,j}$, we have $\mathrm{d}V_L^e(t_e; c_2)/\mathrm{d}t_e = \mathrm{d}V_L^e(t_e; c_1)/\mathrm{d}t_e$, so $V^e(t_e; c_2) = V^e(t_e; c_1)$ in a small neighborhood of T_1. In fact, if we can maintain $\mathrm{d}V_L^e(t_e; c_2)/\mathrm{d}t_e = \mathrm{d}V_L^e(t_e; c_1)/\mathrm{d}t_e$ as long as

$$\frac{1}{G(t_e; c_2)} - \frac{1}{G(t_e; c_1)} = \frac{(\rho+\alpha)(c_2 - c_1)}{N}. \tag{OA.1.7}$$

In an equilibrium in which entrepreneurs are indifferent between entering either market at any time, we can solve the mass of firms in each market at time t_e combining equation (OA.1.7) and the total mass of firms at time t_e:

$$G(t_e; c_2) + G(t_e; c_1) = G(0) + \mu t_e.$$

We can solve

$$\begin{aligned} G(t_e; c_1) &= \frac{G(0) + \mu t_e - 2\xi + \sqrt{(G(0) + \mu t_e)^2 + 4\xi^2}}{2}, \\ G(t_e; c_2) &= \frac{G(0) + \mu t_e + 2\xi - \sqrt{(G(0) + \mu t_e)^2 + 4\xi^2}}{2}, \end{aligned} \tag{OA.1.8}$$

where $\xi \equiv \frac{N}{(\rho+\alpha)(c_2-c_1)}$.

Taking derivative of equation (OA.1.7) with respect to t_e, we have

$$\frac{G'(t_e; c_1)}{G^2(t_e; c_1)} = \frac{G'(t_e; c_2)}{G^2(t_e; c_2)}.$$

Since $G(t_e; c_2) < G(t_e; c_1)$, we establish the comparison of entry rates: $G'(t_e; c_2) < G'(t_e; c_1)$.

Given the calculation above, it is straightforward that this is the only equilibrium in which entrepreneurs are indifferent between entering either market at any time. We now rule out the possibility of an equilibrium in which one market has a strictly higher value of entry at some point. Since $T_1 = T_2$, this has to occur before the steady state. Without loss of generality, we assume that $V^e(t_e; c_1) > V^e(t_e; c_2)$ for on $t_e \in (s, t), s < t < T_1$, while $V^e(t_e; c_1) = V^e(t_e; c_2)$ for all $t_e \geq t$. This implies that $G'(t_e; c_1) = \mu, G'(t_e; c_2) = 0$ on the segment (s, t), and

$$\frac{\mathrm{d}V^e(t_e; c_2)}{\mathrm{d}t_e} - \frac{\mathrm{d}V^e(t_e; c_1)}{\mathrm{d}t_e} = c_2 - c_1 - \left(\frac{N}{(\rho+\alpha)G(t_e; c_2)} - \frac{N}{(\rho+\alpha)G(t_e; c_1)} \right) > 0.$$

Suppose at some point at or before s, there is positive entry in market 2 again, which requires $V^e(t_e; c_1) \leq V^e(t_e; c_2)$. Without loss of generality, we assume this happens exactly at s. However, this leads to an immediate contradiction because $V^e(t; c_1) = V^e(t; c_2)$ and $\frac{\mathrm{d}V^e(t_e;c_2)}{\mathrm{d}t_e} - \frac{\mathrm{d}V^e(t_e;c_1)}{\mathrm{d}t_e} > 0$ for all $t_e \in (s, t)$, which implies that $V^e(t_e; c_1) > V^e(t_e; c_2)$ for all $t_e \in (s, t)$. Therefore, there is no entry at or before t in market 2. This leads to another contradiction: according to equation (OA.1.8), we must have $G(t; c_2) > 0$. We can obtain similar contradictions if we assume $V^e(t_e; c_2) > V^e(t_e; c_1)$ on a segment $t_e \in (s, t)$. Therefore, we conclude that there is a unique equilibrium in which entrepreneurs are indifferent between entering either market at any time, characterized by equation (OA.1.8).

Part 2 Since $I > \underline{I}(\rho, \alpha_L, \alpha_H, c_2) > \underline{I}(\rho, \alpha_L, \alpha_H, c_1)$, the last entrant does not upgrade in either market. The steady-state mass of firms, therefore, is the same as in Part 1. Since $G(t; c_j)$ is continuous in t and $G(T; c_1) > G(T; c_2)$, we must have $G(t; c_1) > G(t; c_2)$ in a neighborhood of $(T - \epsilon, T]$. Market 1 has more firms when t is sufficiently large.

Now suppose $G(t; c_1) \geq G(t; c_2)$ for all t. We seek contradictions as well as parameter combinations that lead to such contradictions. We also choose $G(0)$ to be small enough such that $G(0) < G(T; c_2)$. We now construct a collection of paths to facilitate our proof. These paths depend on a hypothetical fixed cost parameter $c \in [c_2, c_1]$ and are defined as follows

$$\hat{G}(t; c) \equiv \min\{G(t; c_1), G(T; c)\},$$

where $G(T; c) = \frac{N}{(\rho+\alpha_L)c}$ is the steady-state mass of firms under fixed cost c. These paths coincide with $G(t; c_1)$ until it reaches $G(T; c)$. Since $G(T; c_2) \leq G(T; c) \leq G(T; c_1)$, all these paths are higher than $G(t; c_2)$ but lower than $G(t; c_1)$.

We now try to derive a contradiction by showing that the value of entry in market 2 is strictly higher than that in market 1 at $t = 0$ under certain parameter restrictions. Consider a potential entrant at $t = 0$ facing a future path of firm mass $\hat{G}(t; c)$, and denote its value of entry as $\hat{V}^e(0; c)$. When deriving $\partial V^e(t_e)/\partial c$ in Proposition 6, we have calculated the derivative as in equation (A-4). In addition, this derivation does not depend on the shape of $G(t)$ before the steady state, because the marginal benefit of a higher c comes solely from the intensified competition (thus higher profit) after the steady state, which is further amplified by upgrading. Therefore, equation (A-4) can be directly applied to $\hat{V}^e(0; c)$ and we have

$$\frac{\partial \hat{V}^e(0; c)}{\partial c} = \frac{\frac{\rho+\alpha_L}{\rho+\alpha_H} e^{-\rho T(c)} - 1}{\rho}.$$

We now find parameter values such that this derivative is positive for all $c \in [c_2, c_1]$. We first replace $T(c)$ with its upper bound, T, as expressed in equation (OA.1.6), and also set $G(0) = 0$ for simplicity. Setting $\frac{\partial \hat{V}^e(0;c)}{\partial c} > 0$ with this relaxation, we have

$$\mu > \frac{\rho(G(T; c_1) + G(T; c_2))}{\log\left(\frac{\rho+\alpha_L}{\rho+\alpha_H}\right)}.$$

Given μ, we can find a sufficiently small $G(0)$ such that the first entrant upgrades in both markets. If it upgrades in market 1, it also upgrades under the hypothetical equilibrium paths $\hat{G}(t;c)$ because $\hat{G}(t;c) \leq G(t;c_1), \forall t$. This, in turn, ensures that we can apply the expression of $\partial \hat{V}^e(0;c)/\partial c$ above. Therefore, we have

$$V^e(0;c_1) = \hat{V}^e(0;c_1) < \hat{V}^e(0;c_2) \leq V^e(0;c_2),$$

which implies that $G(0;c_1) = 0 < G(0;c_2)$ and contradicts the hypothesis that $G(t;c_2) \leq G(t;c_1)$ for all $t \geq 0$. ■

OA.1.4 Proofs for Paper Appendix B

Lemma OA.1.5 ΓV_1^c *is a contraction with modulus* $\frac{\lambda}{\lambda+\rho}$.

Proof. Note that it is straightforward to show that the mapping from V_1^c to ΓV_1^c satisfies the monotonicity and discounting in Blackwell's sufficient conditions for a contraction. To show that V_1^c is finite, we first consider

$$\hat{V}_0^c = \left(1 - \frac{\lambda}{\rho+\lambda}\right)^{-1} \frac{u_0}{\rho+\lambda} + E(\varepsilon_0) = \frac{u_0}{\rho} + E(\varepsilon_0),$$

which is the value of taking the outside option (permanently, as an absorbing state) in *discrete time* with discount rate $\frac{\lambda}{\rho+\lambda}$ and flow utility $\frac{u_0}{\rho+\lambda} + \varepsilon_0$. Therefore, we first construct a discrete-time dynamic choice problem with the solutions $\hat{V}_1^c, \hat{V}_0^c$ where

$$\hat{V}_1^c = \frac{u_1 + \lambda E_{max}\left\{\hat{V}_1^c + \varepsilon_1, \hat{V}_0^c + \varepsilon_0\right\}}{\rho+\lambda}, \quad \hat{V}_0^c = \frac{u_0}{\rho} + E(\varepsilon_0).$$

Under finite $E_{max}(\varepsilon_j)$, it is straightforward that $\hat{V}_0^c$ is finite. To show $\hat{V}_1^c$ is finite, we consider an alternative problem in which $j = 0$ is not an absorbing state. In this alternative problem, we can write the Bellman equation (with shocks) as

$$\tilde{v}^c(\varepsilon) = \max_{j\in\{0,1\}} \left\{\frac{u_j}{\rho+\lambda} + \varepsilon_j + \frac{\lambda}{\rho+\lambda} E_{max}\left\{\tilde{V}_1^c + \varepsilon_1, \tilde{V}_0^c + \varepsilon_0\right\}\right\}$$

Taking expectations with respect to ε, we have a mapping Γ from the integrated value function (Emax, $E\tilde{v}^c(\varepsilon)$) to $E\tilde{v}^c(\varepsilon)$. With the assumption that $E_{max}\left\{\varepsilon_1, \varepsilon_0\right\}$ is finite, we can write

$$E\tilde{v}^c(\varepsilon) \leq \frac{\frac{\max\{u_j\}}{\rho+\lambda} + E_{max}\{\varepsilon_j\}}{1 - \frac{\lambda}{\rho+\lambda}},$$

which is finite. This implies that $\tilde{V}_j^c$ is finite. Since this is a relaxed problem (no absorbing state), we conclude that $\hat{V}_1^c \leq \tilde{V}_1^c$ is finite. If $E(\varepsilon_0) \geq 0$, then finite $\hat{V}_1^c$ implies finite V_1^c. If $E(\varepsilon_0) < 0$, we have

$$V_1^c = \frac{u_1 + \lambda E_{max}\left\{V_1^c + \varepsilon_1, \hat{V}_0^c - E(\varepsilon_0) + \varepsilon_0\right\}}{\rho + \lambda} \leq \frac{u_1 + \lambda E_{max}\left\{V_1^c + \varepsilon_1, \hat{V}_0^c + \varepsilon_0\right\} - \lambda E(\varepsilon_0)}{\rho + \lambda}.$$

It is easy to establish a lower bound for V_1^c, as

$$V_1^c \geq \frac{u_1 + \lambda\left(\frac{u_0}{\rho} + E(\varepsilon_0)\right)}{1 - \frac{\lambda}{\rho+\lambda}}$$

Therefore, V_1^c must be bounded. We can then apply the Blackwell's sufficient conditions and conclude that Γ is a contraction mapping with modulus $\frac{\lambda}{\lambda+\rho}$. ■

Proof of Proposition B-2. The first-order condition for p is

$$v'(p) = \frac{\rho + \alpha + p\alpha'}{[\rho + \alpha]^2} = 0, \tag{OA.1.1}$$

where the derivative of α is taken with respect to the consumer flow surplus $u_1 - p$. For p to be a local maximum, it needs to satisfy the second-order condition

$$v''(p) = \frac{\frac{\partial}{\partial p}(\rho + \alpha + p\alpha')}{(\rho + \alpha)^2} = \frac{p\alpha''}{(\rho + \alpha)^2} \leq 0 \Rightarrow \alpha''(u_1 - p) \leq 0.$$

Taking derivative of equation (OA.1.1) with respect to u_1 at $p = p^*$, we have

$$\frac{dp^*}{du_1} = 1 + \frac{\alpha'}{p\alpha''}\bigg|_{p=p^*}.$$

Using $\alpha'' \leq 0$, we must have $\frac{dp^*}{du_1} \leq 1$ thus $\frac{d(u_1 - p^*)}{du_1} \geq 0$. Under the assumption that $p^*\alpha''/\alpha' < -1$, we must have

$$\frac{\mathrm{d}p^*}{\mathrm{d}u_1} > 0.$$

■

OA.2 An Extension: Endogenous Upgrading with Price Differences

We now consider an extension of the baseline model by allowing H- and L-type firms to charge different prices, $p_H \geq p_L$. The baseline model in the paper can be seen as the special case where $p_H = p_L = 1$. We can always normalize $p_L = 1$ without loss of generality. We choose not to so that we can keep track of p_L and p_H in the formulas.

The following lemma characterizes the timing of upgrading if a firm ever upgrades:

Lemma OA.2.1 *If a firm ever upgrades from L-type to H-type, it will happen*

1. *when $m(t; t_e)$ first reaches m^* if $m(t_e; t_e) < m^*$*
2. *at $t = t_e$ if $m(t_e; t_e) \geq m^*$*

The cutoff m^* has a closed-form expression:

$$m^* \equiv \frac{\rho(\rho + \alpha_H)}{(\rho + \alpha_H)(p_H - p_L) + (\alpha_L - \alpha_H)p_H} I. \tag{OA.2.1}$$

Proof. Consider a firm that enters at time t_e and contemplates whether to upgrade at t_u or $t_u + \Delta t$. The gains in profits from upgrading early have two components: (1) the gains in profits in $t \in [t_u, t_u + \Delta t]$ and (2) the gains in profits in $t \in [t_u + \Delta t, \infty)$. When Δt is small, we can use the following approximations for $s \leq \Delta t$

$$m(t_u + s; t_e) = m(t_u; t_e)[1 - \alpha_L s] + \frac{N}{G(t_u)} s$$
$$\tilde{m}(t_u + s; t_e) = m(t_u; t_e)[1 - \alpha_H s] + \frac{N}{G(t_u)} s$$

Since the firm charges customers p_H instead of p_L after upgrading, we write the difference in flow profit as

$$p_H \tilde{m}(t_u + s; t_e) - p_L m(t_u + s; t_e) \approx (p_H - p_L) m(t_u; t_e),$$

where we have dropped all terms involving s because their impact on the extra profit between t_u and $t_u + \Delta t$ is $O(\Delta t^2)$. The newly accrued customers after $t_u + \Delta t$ generate the same profits whether the firm upgrades at t_u or $t_u + \Delta t$. From $t_u + \Delta t$ onward, the gains in profits only come from the extra customer capital at $t_u + \Delta t$, $\Delta m(t_u + \Delta t; t_e)$, which depreciates at the rate α_H after time $t_u + \Delta t$. The gain in profits can be written as

$$\Delta \Pi = \int_{t_u+\Delta t}^{\infty} e^{-\rho(t-t_u-\Delta t)} p_H m(t_u; t_e)(\alpha_L - \alpha_H)\Delta t e^{-\alpha_H(t-t_u-\Delta t)} \mathrm{d}t = p_H \frac{m(t_u; t_e)(\alpha_L - \alpha_H)\Delta t}{\rho + \alpha_H},$$

which is $O(\Delta t)$. The additional costs of upgrading early is $I(1 - e^{-\rho \Delta t}) \approx \rho \Delta t I$. Equalizing the additional profits and the additional costs, we obtain the cutoff m^* as in equation (OA.2.1):

$$(p_H - p_L) m(t_u; t_e)\Delta t + p_H \frac{m(t_u; t_e)(\alpha_L - \alpha_H)\Delta t}{\rho + \alpha_H} = \rho I \Delta t.$$

The first term in the additional profits is due to higher prices, and the second term is due to higher customer retention.

Formally, denote the extra NPV of upgrading at time t_u evaluated at time t_u as $\Delta \tilde{V}(t_u; t_e)$, we can calculate it as

$$\begin{aligned}
\Delta \tilde{V}(t_u; t_e) &= \int_{t_u}^{\infty} e^{-\rho(t-t_u)} \left(p_H \tilde{m}(t; t_e) - p_L \tilde{m}(t; t_e)\right) \mathrm{d}t \\
&= \left(\frac{p_H}{\rho + \alpha_H} - \frac{p_L}{\rho + \alpha_L}\right) m(t_u; t_e) + \\
&\int_{t_u}^{\infty} e^{-\rho(t-t_u)} \left(\int_{t_u}^{t} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)}\right] \mathrm{d}s\right) \mathrm{d}t - I
\end{aligned} \tag{OA.2.2}$$

The NPV of upgrading evaluated at time t_e is $\Delta V(t_u; t_e) = e^{-\rho(t_u - t_e)} \Delta \tilde{V}(t_u; t_e)$. Taking the derivative with respect to t_u, we have

$$\frac{\partial \Delta V(t_u; t_e)}{\partial t_u} = e^{-\rho(t_u - t_e)} \left[\rho I - \left(\frac{\rho + \alpha_L}{\rho + \alpha_H} p_H - p_L\right) m(t_u; t_e)\right], \tag{OA.2.3}$$

which is positive when $m(t_u; t_e) < m^*$ and negative when $m(t_u; t_e) > m^*$.

To ensure that t_u is a global maximum of $\Delta V(t_u; t_e)$, $m(t; t_e)$ has to reach m^* from below (on its increasing segment), i.e., $\frac{\partial \Delta V(t; t_e)}{\partial t}|_{t==t_u^-} > 0$, $\frac{\partial \Delta V(t; t_e)}{\partial t}|_{t==t_u^+} < 0$. From Lemma OA.1.1 in the paper, we know $m(t; t_e)$ is either increasing, decreasing or first increasing and then decreasing. When $m(t; t_e)$ is increasing on $[t_e, \infty)$, the firm must upgrade when $m(t; t_e)$ (first) crosses m^* if it ever upgrades. When $m(t; t_e)$ is decreasing on $[t_e, \infty)$, the value $\Delta V(t_u; t_e)$ is maximized when $t_u = t_e$. Finally, when $m(t; t_e)$ is hump-shaped and crosses m^* twice at $t_1 < t_2$. We know from equation (OA.2.3), the firm either upgrades at t_1 or $t = \infty$ (t_2 is a local minimum). However, it is clear that $\lim_{t_u \to \infty} \Delta V(t_u; t_e) = 0$. Therefore, $\Delta V(t_u; t_e)$ monotonically converges to zero from below in the range of $t \in [t_2, \infty)$. Therefore, if the firm upgrades, it will only upgrade at t_1. ■

Before proving Proposition OA.2.1, we derive some useful expressions and a lemma. that characterizes the equilibrium when firms can upgrade from α_L to α_H, we derive some useful expressions and a lemma.

For the second component on the right-hand side of equation (OA.2.2), we can further simplify it if $t_u \geq T, G(s) = G(T), s \geq t_u$:

$$\int_{t_u}^{\infty} e^{-\rho(t-t_u)} \left(\int_{t_u}^{t} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)} \right] \mathrm{d}s \right) \mathrm{d}t = \frac{\kappa}{\rho} \frac{N}{G(T)}, \tag{OA.2.4}$$

where κ is defined as

$$\kappa \equiv \frac{p_H}{\rho + \alpha_H} - \frac{p_L}{\rho + \alpha_L} \tag{OA.2.5}$$

Note that κ can also be used to simplify the cutoff $m^* = \frac{\rho I}{(\rho+\alpha_L)\kappa}$.

When $t_u < T$, we can perform the integration for $t \in [t_u, T)$ and $t \in [T, \infty)$ separately

$$\begin{aligned}
&\int_{t_u}^{\infty} e^{-\rho(t-t_u)} \left(\int_{t_u}^{t} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)} \right] \mathrm{d}s \right) \mathrm{d}t \\
=& \int_{t_u}^{T} e^{-\rho(t-t_u)} \left(\int_{t_u}^{t} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)} \right] \mathrm{d}s \right) \mathrm{d}t \\
&+ \int_{T}^{\infty} e^{-\rho(t-t_u)} \left(\int_{t_u}^{T} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)} \right] \mathrm{d}s \right) \mathrm{d}t \\
&+ \int_{T}^{\infty} e^{-\rho(t-t_u)} \left(\int_{T}^{\infty} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)} \right] \mathrm{d}s \right) \mathrm{d}t
\end{aligned} \tag{OA.2.6}$$

Note that the third term is the same as equation (OA.2.4) multiplied by $e^{-\rho(T-t_u)}$. The second term can be simplified as

$$\begin{aligned}
&\int_{T}^{\infty} e^{-\rho(t-t_u)} \left(\int_{t_u}^{T} \frac{N}{G(s)} \left[p_H e^{-\alpha_H(t-s)} - p_L e^{-\alpha_L(t-s)} \right] \mathrm{d}s \right) \mathrm{d}t \\
=& p_H \int_{T}^{\infty} e^{-\rho(t-t_u)} \times e^{-\alpha_H(t-T)} \left[\int_{t_u}^{T} \frac{N e^{-\alpha_H(T-s)}}{G(s)} \mathrm{d}s \right] \mathrm{d}t - p_L \int_{T}^{\infty} e^{-\rho(t-t_u)} \times e^{-\alpha_L(t-T)} \left[\int_{t_u}^{T} \frac{N e^{-\alpha_L(T-s)}}{G(s)} \mathrm{d}s \right] \mathrm{d}t \\
=& \frac{e^{-\rho(T-t_u)} p_H}{\rho + \alpha_H} \left[\int_{t_u}^{T} \frac{N e^{-\alpha_H(T-s)}}{G(s)} \mathrm{d}s \right] - \frac{e^{-\rho(T-t_u)} p_L}{\rho + \alpha_L} \left[\int_{t_u}^{T} \frac{N e^{-\alpha_L(T-s)}}{G(s)} \mathrm{d}s \right]
\end{aligned}$$

This expression is useful for numerical implementation.

Lemma OA.2.2 *If the firms that entered at time t_e upgrade, then all firms that entered before t_e also upgrade.*

Proof. Suppose that a firm that entered at time t_e upgrades at t_u, and suppose $t_e' < t_e$. We know that

$$m(t_u; t_e') > m(t_u; t_e)$$

because the earlier entrants accumulated customers in $t \in [t_e', t_e)$. From the expression of $\Delta\tilde{V}(t_u; t_e)$, equation (OA.2.2), we obtain immediately

$$\Delta\tilde{V}(t_u; t_e') > \Delta\tilde{V}(t_u; t_e) \geq 0,$$

where the second inequality holds because the firm that entered at time t_e finds it profitable to upgrade at t_u. Therefore, it is also profitable for the earlier entrants to upgrade at t_u, though upgrading at the optimal $t_u(t_e')$ will bring them even larger discounted net gains $\Delta V(t_u(t_e'); t_e')$. ■

We characterize the equilibrium of entry and upgrading as follows

Proposition OA.2.1 *In an equilibrium, firms enter at constant rate μ until T and no firm exits. There exist two cutoff investment costs $\underline{I}$ and $\bar{I}$ such that*

1. *If $I \leq \underline{I}$, all entrants upgrade. This cutoff has a closed-form expression*

$$\underline{I} = \frac{(\rho + \alpha_L)^2 \kappa c}{\rho \alpha_L p_L}. \tag{OA.2.7}$$

2. *If $I > \bar{I}$, no entrants upgrade. This cutoff can be solved from*

$$\max_{t_u} \Delta V\left(t_u; 0, \bar{I}\right) = 0,$$

where we have written I as a parameter to indicate the value of upgrading depends on its cost.

3. *When $G(0)$ is sufficiently small, we have $\underline{I} < \bar{I}$. If $I \in (\underline{I}, \bar{I}]$, there exists a cutoff $T_u \in [0, T)$, such that firms that enter at $t \in [0, T_u]$ will upgrade and later entrants will not.*

In the second and third cases, the steady-state mass of firms, $G(T)$, equals

$$G(T) = \frac{N p_L}{(\rho + \alpha_L)c}. \tag{OA.2.8}$$

Proof of Proposition OA.2.1. We first consider the case where all firms upgrade. For the last entrant, it faces a constant number of competing firms, $G(T)$, after entry. Therefore, without upgrading,

$$m(t;T) = \frac{N}{G(T)\alpha_L}\left(1 - e^{-\alpha_L(t-T)}\right), \forall t \geq T,$$

we can solve for the optimal time for upgrading, $t_u(T)$, by setting $m(t;T) = m^*$

$$t_u(T) = T + \frac{-\log\left(1 - \frac{G(T)\alpha_L}{N}m^*\right)}{\alpha_L}.$$

Note that this also implies a necessary condition of this case, involving an endogenous variable $G(T)$: the steady-state customer capital of low-type firms, $m_{L,lr} \equiv \frac{N}{G(T)\alpha_L}$, is above the cutoff customer capital, m^*. This ensures that the last entrant's customer base will reach m^* at some point.

In this case, the zero-profit condition for the last entrant becomes

$$\left(\max_{t_u} \Delta V(t_u;T)\right) + V_L^e(T) = 0,$$

where we the value of being a low-type firm forever since T can be calculated as:

$$V_L^e(T) = \int_T^\infty e^{-\rho(t-T)}(p_L m(t;T) - c)\mathrm{d}t = \frac{Np_L}{G(T)\rho(\rho+\alpha_L)} - \frac{c}{\rho}. \tag{OA.2.9}$$

From equations (OA.2.2) and (OA.2.4), the gains from optimal upgrading is

$$\begin{aligned}
\Delta V(t_u(T);T) &= e^{-\rho(t_u(T)-T)}\Delta\tilde{V}(t_u(T);T) \\
&= \left(1 - m^*\frac{G(T)\alpha_L}{N}\right)^{\frac{\rho}{\alpha_L}} \times \left(\kappa m^* + \frac{\kappa N}{\rho G(T)} - I\right) \\
&= \left(1 - m^*\frac{G(T)\alpha_L}{N}\right)^{\frac{\rho}{\alpha_L}} \times \left(\frac{\kappa N}{\rho G(T)} - \frac{\alpha_L}{\rho+\alpha_L}I\right)
\end{aligned}$$

Denote $F(G(T)) \equiv (\max_{t_u} \Delta V(t_u;T)) + V_L^e(T)$ where $G(T) \in \left(0, \frac{N}{\alpha_L m^*}\right)$. One can show that $F(G(T))$ is strictly decreasing in $G(T)$ and

$$\lim_{G(T)\to 0} F(G(T)) = \infty, \qquad \lim_{G(T)\to \frac{N}{\alpha_L m^*}} F(G(T)) = V_L^e(T) = \frac{\alpha_L p_L m^*}{\rho(\rho+\alpha_L)} - \frac{c}{\rho}.$$

Note that $\max_{t_u} \Delta V(t_u;T) \to 0$ when $G(T) = \frac{N}{\alpha_L m^*}$. This holds because when $G(T) =$

$\frac{N}{(1-\alpha_L)m^*}$, $m_{L,lr} = m^*$ and $m(t;T)$ converges to m^* from below. Therefore, $\Delta V(t_u;T)$ strictly increases in t_u and converges to its maximum when $t_u \to \infty$. In addition, $\Delta\tilde{V}(t_u;T) = \kappa m(t_u;T) + \frac{\kappa N}{\rho G(T)} - I$ is bounded because $m(t_u;T) \in [0, m^*)$. Therefore, $\Delta V(t_u;T) = e^{-\rho(t_u - T)}\Delta\tilde{V}(t_u;T) \to 0$ when $t_u \to \infty$.

Therefore, as long as $\frac{\alpha_L p_L m^*}{\rho(\rho+\alpha_L)} - \frac{c}{\rho} < 0$, there exists a unique solution $G(T)$. This is equivalent to $m_{L,lr} > m^*$ under $G(T) = \frac{Np_L}{(\rho+\alpha_L)c}$, as well as the inequality

$$I < \underline{I} \equiv \frac{(\rho+\alpha_L)^2\kappa c}{\rho\alpha_L p_L}.$$

When $I \geq \underline{I}$, we know that we cannot find a solution $G(T)$ such that $F(G(T)) = 0$. Therefore, at least some entrants (later entrants) will not upgrade. To ensure that at least some firms upgrade, we need to ensure that the first entrant will upgrade because of Lemma OA.2.2. Formally, we need

$$\max_{t_u} \Delta V(t_u; 0, I) \geq 0.$$

We can show that there is a unique $\bar{I}$ such that

$$\max_{t_u} \Delta V(t_u; 0, \bar{I}) = 0,$$

and for $I > \bar{I}$, firms that enter at $t = 0$ do not upgrade. We first show that $\max_{t_u} \Delta V(t_u; 0, I)$ strictly decreases in I. To see this, consider $I_1 < I_2$ and denote the optimal upgrading time under different investment costs as $t_u(0, I)$. We have the following relationship

$$\Delta V(t_u(0, I_2); 0, I_2) < \Delta V(t_u(0, I_2); 0, I_1) \leq \Delta V(t_u(0, I_1); 0, I_1),$$

where the first inequality can be obtained by observing equation (OA.2.2) and the relationship between ΔV and $\Delta\tilde{V}$, and the second inequality comes from the fact that $\Delta V(t_u; 0, I_1)$ is maximized under $t_u(0, I_1)$. Next, we show that $\lim_{I\to\infty} \max_{t_u} \Delta V(t_u; 0, I) \leq 0$ and $\max_{t_u} \Delta V(t_u; 0, I)|_{I=0} > 0$ under sufficiently small $G(0)$. The first inequality results from the fact that $\Delta\tilde{V}(t_u; 0) + I$ does not depend on I, and it is bounded from above by the value of being a high type from $t = 0$. However, the value of being a high type is also bounded from above because $G(t) \geq G(0), \forall t$. Therefore, we can find sufficiently large I such that $\Delta V(t_u; 0, I) < 0$ for any t_u. We obtain $\lim_{I\to\infty} \max_{t_u} \Delta V(t_u; 0, I) \leq 0$. To prove the second inequality, note that when $I = 0$, upgrading always results in positive value $\Delta V(t_u; 0) > 0$ as long as $\alpha_H > \alpha_L$ and/or $p_H > p_L$. In sum, we have shown the existence and uniqueness of $\bar{I}$.

We now show that $\bar{I} > \underline{I}$ for sufficiently small $G(0)$. In particular, we show that the first entrant's value of upgrading is arbitrarily large when $G(0)$ is sufficiently small. Consider a firm that enters at $t = 0$ and immediately upgrades ($t_u = 0$). The net value of upgrading can be written as

$$\Delta V(0; 0, I) = \int_0^\infty e^{-\rho t} \left[\int_0^t \frac{N}{G(0) + \mu s} \left(p_H e^{-\alpha_H (t-s)} - p_L e^{-\alpha_L (t-s)} \right) \mathrm{d}s \right] \mathrm{d}t - I.$$

For any t, the integral inside the first component is unbounded when $G(0) \to 0$. To see this,

$$\int_0^t \frac{p_H e^{-\alpha_H (t-s)} - p_L e^{-\alpha_L (t-s)}}{G(0) + \mu s} \mathrm{d}s \geq \left(\min_{s \in [0,t]} p_H e^{-\alpha_H (t-s)} - p_L e^{-\alpha_L (t-s)} \right) \times \int_0^t \frac{1}{G(0) + \mu s} \mathrm{d}s,$$

where the first minimum exists and is strictly positive because $p_H e^{-\alpha_H (t-s)} - p_L e^{-\alpha_L (t-s)}$ is a positive continuous function on $[0, t]$, while the second integral is unbounded when $G(0) \to 0$.

Finally, the characterization of equilibrium when $I \in [\underline{I}, \bar{I})$ follows from the construction of $\underline{I}$ and $\bar{I}$ as well as Lemma OA.2.2. ■

Corollary OA.2.1 *If $I \in [\underline{I}, \bar{I}]$, firms that ever upgrade must do so no later than T. When $I = \underline{I}$, the last upgrader must upgrade at T. When $I \in (\underline{I}, \bar{I}]$, the last upgrader upgrades strictly before T.*

Proof. The corollary and its proof are the same as Corollary OA.1.1, because the generalization to different prices does not alter the equivalence between $m_{L,lr} = m^*$ and $I = \underline{I}$. ■

Lemma OA.2.3 *Suppose $I < \bar{I}$. The time of upgrading of the last upgrader, $t_u(T_u)$, is continuous in μ.*

Proof. The lemma is the same as OA.1.3.

!!! to be reviewed !!!

As with the baseline model, we can show that the entry time of the last upgrader, T_u, is continuous in μ in the two cases, $I < \underline{I}$ and $I \in [\underline{I}, \bar{I})$. We use the same strategy to show that the indirect value of upgrading $\Delta V(t_u(t_e, \mu); t_e, \mu)$ is continuously differentiable in t_e

and μ. Given equation (OA.2.2), the derivative of $\Delta V(t_u(t_e,\mu);t_e,\mu)$ with respect to t_e is now

$$\frac{\partial \Delta V(t_u(t_e,\mu);t_e,\mu)}{\partial t_e} = -e^{-\rho(t_u-t_e)} \cdot \kappa \frac{Ne^{-\alpha_L(t_u(t_e,\mu)-t_e)}}{G(t_e)} < 0$$

Therefore, we can set $\Delta V(t_u(t_e,\mu);t_e,\mu) = 0$ and apply the implicit function theorem to an open neighborhood of μ and obtain a continuously differentiable function $T_u(\mu)$. Therefore, $t_u(T_u(\mu),\mu)$ must be a continuously differentiable (thus continuous) function of μ.

■

Lemma OA.2.4 *The share of H-type firms at any time point t, $g_H(t;\mu)$, is differentiable in μ for $t > t_{u0}(\mu)$.*

Proof. The lemma is the same as Lemma OA.1.4. The proof is also the same because allowing for different prices does not alter the expressions for $m(t;t_e)$ and $G(t)$, thus none of the expressions in the proof for Lemma OA.1.4 changes. The only difference is that now we have a different threshold m^*, which implies a different initial value $t_{u0}(\mu)$. ■

Proposition OA.2.2

1. *A higher market entry rate μ or a lower flow fixed cost c reduces the share of H-type firms in the steady state. The decrease is strict when $I \geq \underline{I}$.*
2. *Assume $I \geq \underline{I}$. The share of H-type firms at any time t, $g_H(t)$, decreases in μ and increases in c. The decrease/increase is strict for sufficiently large t.*

Proof. The statements and proofs are the same as Proposition 4. The key to the proof is that future competition reduces the incentives for upgrading, which remains unchanged when we allow for different prices. The auxiliary lemma used in the proof, Lemma OA.2.4, has also been proved earlier. ■

OA.3 Discerning Customers

In this section, we extend our model to allow the H-type firms to have not only a higher customer retention rate but also an advantage in attracting new customers over L-type firms. We assume that the customer arrival rate of the H-type firms is γ times that of the L-type firms ($\gamma \geq 1$). Denoting the mass of H- and L-type firms at time t as $G_H(t)$ and $G_L(t)$, the customer arrival rates must satisfy

$$A_H(t) = \frac{\gamma N}{\gamma G_H(t) + G_L(t)}, \quad A_L(t) = \frac{N}{\gamma G_H(t) + G_L(t)}. \tag{OA.3.1}$$

These rates ensures that $A_H(t)G_H(t) + A_L(t)G_L(t) = N$ and $A_H(t) = \gamma A_L(t)$. One can microfound these customer arrival rates assuming that each new customer chooses his/her favorite product from all the H- and L-type firms in a random utility model (Train (2003)). All firms of the same type deliver the same representative utility, and $\gamma > 1$ implies that H-type firms deliver higher utility on average.[19] We define the *quality-adjusted* mass of firms,

$$G_Q(t) \equiv \gamma G_H(t) + G_L(t) = (\gamma - 1)G_H(t) + G(t). \tag{OA.3.2}$$

This variable summarizes the competitiveness of the market at time t. Our baseline model in Section 4 of the paper is a special case of the current model with $\gamma = 1$.

We make two assumptions, one rules out exits on the equilibrium path and one rules out firm upgrading upon entry:

Assumption OA.3.1 *No firms ever exit in equilibrium.*

Assumption OA.3.2 *The mass of initial entrants, $G(0)$, is larger than a threshold $\underline{G}_0$:*

$$G(0) > \underline{G}_0 \equiv \frac{(\gamma - 1)N}{\rho(\rho + \alpha_H)I}. \tag{OA.3.3}$$

Note that the no-exit assumption also ensures that $G(t)$ and $G_Q(t)$ weakly increase on $[0, \infty)$.

[19] For example, we can assume that each new customer draws an idiosyncratic unobserved utility for each variety provided by the current $G_H(t)$ H-type and $G_L(t)$ L-type firms. The shock is independently, identically distributed extreme value. The common utility across consumers for a j-type variety is $V_{1,j}^c$, $j \in \{H, L\}$ as derived in Appendix 4.3 of the paper. The probability that a new customer chooses a j-type firm becomes $\frac{G_j(t)e^{V_{1,j}^c}}{G_H(t)e^{V_{1,H}^c} + G_L(t)e^{V_{1,L}^c}}$, which is consistent with (OA.3.1) if $\gamma = e^{V_{1,H}^c - V_{1,L}^c}$.

With these two assumptions, we can characterize a necessary condition for optimal upgrading as follows:

Lemma OA.3.1 *For each time t, there exists a cutoff customer capital level*

$$m^*(t) \equiv \frac{\rho(\rho+\alpha_H)I - \frac{(\gamma-1)N}{G_Q(t)}}{\alpha_L - \alpha_H} \tag{OA.3.4}$$

If a firm ever upgrades from L- to H-type, it will happen when $m(t;t_e)$ crosses $m^*(t)$ from below, i.e., $m(t;t_e) \le m^*(t)$ for a small neighborhood $t \in (t_u - \Delta t, t_u)$ and $m(t;t_e) \ge m^*(t)$ for a small neighborhood $t \in (t_u, t_u + \Delta t)$, if $m^*(t_e) > 0$.

Proof. Consider a firm that enters at time t_e and contemplates whether to upgrade at t_u or $t_u + \Delta t$. The gains in profits from upgrading early has two components: (1) the gains in profits in $t \in [t_u, t_u + \Delta t]$ and (2) the gains in profits in $t \in [t_u + \Delta t, \infty)$. When Δt is small, we can use the following approximations for $s \le \Delta t$

$$\begin{aligned} m(t_u + s; t_e) &= m(t_u; t_e)[1 - \alpha_L s] + A_L(t_u)s \\ \tilde{m}(t_u + s; t_e) &= m(t_u; t_e)[1 - \alpha_H s] + A_H(t_u)s \end{aligned}$$

Therefore, the difference in profits due to upgrading is

$$\Delta m(t_u + s; t_e) = m(t_u; t_e)(\alpha_L - \alpha_H)s + \frac{(\gamma-1)N}{G_Q(t_u)}s.$$

It is immediate that the gain in profits in $t \in [t_u, t_u + \Delta t)$ is $O(\Delta t^2)$. This is because the time length is Δt and the maximum difference in instantaneous profit is $m(t_u;t_e)(\alpha_L - \alpha_H)\Delta t + \frac{(\gamma-1)N}{G_Q(t_u)}\Delta t$ (discount rate can be ignored due to small Δt). The newly accrued customers after $t_u + \Delta t$ generate the same profits whether the firm upgrades at t_u or $t_u + \Delta t$. From $t_u + \Delta t$ onward, the gains in profits only come from the extra customer capital at $t_u + \Delta t$, $\Delta m(t_u + \Delta t; t_e)$, which depreciates at the rate α_H after time $t_u + \Delta t$. The gain in profits can be written as

$$\begin{aligned} \Delta\Pi &= \int_{t_u+\Delta t}^{\infty} e^{-\rho(t - t_u - \Delta t)} \Delta m(t_u + \Delta t; t_e) e^{-(1-\alpha_H)(t - t_u - \Delta t)} \mathrm{d}t \\ &= \left(\frac{(\alpha_L - \alpha_H) m(t_u; t_e)}{\rho + \alpha_H} + \frac{(\gamma-1)N}{(\rho+\alpha_H)G_Q(t_u)} \right) \Delta t, \end{aligned}$$

which is $O(\Delta t)$. Therefore, we can ignore the gain in profits between t_u and $t_u + \Delta t$. The additional costs of upgrading early is $I(1 - e^{-\rho\Delta t}) \approx \rho\Delta t I$. Equalizing the additional profits and the additional costs, we obtain the cutoff m^* as in equation (OA.3.4).

Formally, denote the extra NPV of upgrading at time t_u evaluated at time t_u as $\Delta\tilde{V}(t_u;t_e)$, we can write

$$\Delta\tilde{V}(t_u;t_e) = \frac{(\alpha_L-\alpha_H)m(t_u;t_e)}{(\rho+\alpha_L)(\rho+\alpha_H)} + \int_{t_u}^{\infty} e^{-\rho(t-t_u)} \int_{t_u}^{t} A_H(s)e^{-\alpha_H(t-s)}\mathrm{d}s\mathrm{d}t \\ - \int_{t_u}^{\infty} e^{-\rho(t-t_u)} \int_{t_u}^{t} A_L(s)e^{-\alpha_L(t-s)}\mathrm{d}s\mathrm{d}t - I. \quad \text{(OA.3.5)}$$

The derivative of the extra NPV of upgrading evaluated at time t_e, $\Delta V(t_u;t_e) = e^{-\rho(t_u-t_e)}\Delta\tilde{V}(t_u;t_e)$, is as follows:

$$\frac{\partial \Delta V(t_u;t_e)}{\partial t_u} = e^{-\rho(t_u-t_e)}\left(\rho I - \frac{(\alpha_L-\alpha_H)m(t_u;t_e) + \frac{(\gamma-1)N}{G_Q(t_u)}}{\rho+\alpha_H}\right), \quad \text{(OA.3.6)}$$

which is positive when $m(t_u;t_e) < m^*(t_u)$ and negative when $m(t_u;t_e) > m^*(t_u)$. To ensure that t_u is a global maximum of $\Delta V(t_u;t_e)$, $m(t;t_e)$ has to reach m^* from below, i.e., $\frac{\partial \Delta V(t;t_e)}{\partial t} \geq 0$ for a small neighborhood $t \in (t_u - \Delta t, t_u)$, and $\frac{\partial \Delta V(t;t_e)}{\partial t} \leq 0$ for a small neighborhood $t \in (t_u, t_u + \Delta t)$.

Similar to our baseline model there is another possibility of local maximum: $m(t_e;t_e) \geq m^*(t_e)$ so the value of upgrading has a local maximum at t_e, which may lead to upgrading upon entry. Since we assume $m(t_e;t_e) = 0$, this is equivalent to $m^*(t_e) \leq 0$. Due to Assumption OA.3.1, $m^*(t)$ is weakly increasing. The restriction $G(0) > \underline{G}_0$ in equation (OA.3.3) is equivalent to $m^*(0) > 0$, which implies $m^*(t_e) > 0 = m(t_e;t_e)$ for all t_e. Therefore, we can rule out this possibility. ■

When there is no quality difference between H- and L-type firms ($\gamma = 1$), the cutoff in equation (OA.3.4) becomes the same as the cutoff in equation (9) of the paper. When $\gamma > 1$, the cutoff becomes time-varying and weakly increasing as long as $G_Q(t)$ is weakly increasing. Since upgrading is a one-time decision, $G_H(t)$ must be weakly increasing. From equation (OA.3.2), a sufficient condition for $G_Q(t)$ to be weakly increasing is that no firm exits along the transition path. Taking everything else as given, a larger quality difference between H- and L-type firms (a higher γ) will lower the cutoff and give firms more incentives to upgrade.

Lemma OA.3.2 *$G_H(t)$ and $m^*(t)$ are continuous on $[0,\infty)$.*

Proof. This is a direct result from Lemma OA.3.1 and the fact that $m(t;t_e)$ is continuous. $m(t;t_e)$ is continuous because it is a solution to the ODE (1) in the paper, where the arrival rate of new customers is bounded – $G_Q(t)$ is bounded by $[G(0), \gamma G(T_e)]$.

Suppose $G_H(t)$ has a jump at $t > 0$, i.e., we have $G_H(t^+) > G_H(t^-)$, which implies $m^*(t^+) > m^*(t^-)$ from equation (OA.3.4).

The jump in $G_H(t)$ implies a positive mass of firms upgrade at time t. Local optimality requires that $\Delta V(t; t_e)$ increases before t and decreases after t. Therefore, for a sufficiently small value of Δt, we have $m(s; t_e) \leq m^*(s), s \in (t - \Delta t, t)$ and $m(s; t_e) \geq m^*(s), s \in (t, t + \Delta t)$. However, due to the gap $m^*(t^+) > m^*(t^-)$, we must have $m(s_1; t_e) > m(s_2; t_e), s_1 \in (t, t + \Delta t), s_2 \in (t - \Delta t, t)$. This is contradictory to the fact that $m(t; t_e)$ is continuous.

We have proved that $G_H(t)$ is continuous. It is immediate that $m^*(t)$ is continuous from equations (OA.3.2) and (OA.3.4). ■

Assumption OA.3.3 *For a firm that enters after $t = 0$, if it ever upgrades, it upgrades at the earliest time when the conditions in Lemma OA.3.1 are met.*

This assumption restricts the set of optimal upgrading time to be a singleton for firms with $t_e > 0$. In general, we need to know more about the functions $m(t; t_e)$ and $m^*(t)$ to prove the set is a singleton.

Lemma OA.3.3 *If late entrants upgrade, early entrants must upgrade. Among upgrading firms, earlier entrants upgrade strictly earlier.*

Proof. Consider the value of upgrading of a firm that enters at $t_e < t'_e$ at the time when the later entrant optimally upgrades, denoted as t'_u. Using equation (OA.3.5), we have

$$\begin{aligned}
\Delta\tilde{V}(t'_u; t_e) &= \frac{(\alpha_L - \alpha_H)m(t'_u; t_e)}{(\rho + \alpha_H)(\rho + \alpha_L)} - I + \left[\Delta\tilde{V}(t'_u; t'_e) - \frac{(\alpha_L - \alpha_H)m(t'_u; t'_e)}{(\rho + \alpha_H)(\rho + \alpha_L)} + I\right] \\
&= \frac{(\alpha_L - \alpha_H)[m(t'_u; t_e) - m(t'_u; t'_e)]}{(\rho + \alpha_H)(\rho + \alpha_L)} + \Delta\tilde{V}(t'_u; t'_e)
\end{aligned}$$

where the terms in the bracket in the first equation represent the additional value accumulated due to new customers after t'_u, which does not depend on the history before t'_u, thus independent of the time of entry. In addition, we can write the customer base at time t'_u

$$m(t'_u; t_e) = m(t'_e; t_e)e^{-\alpha_L(t'_u - t'_e)} + m(t'_u; t'_e) > m(t'_u; t'_e).$$

Therefore,

$$\Delta\tilde{V}(t'_u; t_e) > \Delta\tilde{V}(t'_u; t'_e) \geq 0.$$

This ensures that the earlier entrant will upgrade, though the optimal upgrading time can differ from t'_u.

We now prove that $t'_u > t_u$. We rely on Assumption OA.3.3. Because the earlier entrant enters at t_e and upgrades at t_u, we know that $m(t;t_e) \leq m^*(t)$ for all $t \in [t_e, t_u]$. Given the law of motion for the customer capital, we know that

$$m(t;t'_e) < m(t;t_e) \leq m^*(t), \; \forall t \in [t'_e, t_u]$$

Therefore, the later entrant does not upgrade between $[t'_e, t_u]$ and $t'_u > t_u$. ■

Lemma OA.3.4 *Under the assumptions OA.3.1, OA.3.2 and OA.3.3, the path of $G_H(t)$ can be characterized by three cutoffs $t_{u0,0}$, $t_{u0,1}$ and $\bar{t}_u$ and two ODEs*

1. $t \in [0, t_{u0,0}]$, *no firm upgrades so* $G_H(t) = 0$.
2. $t \in (t_{u0,0}, t_{u0,1}]$, *initial entrants are indifferent between upgrading earlier and later.* $G_H(t)$ *can be characterized by the following ODE:*

$$\begin{aligned}
&G'_H(t_u) \times \frac{(\gamma-1)^2 N}{(\alpha_L - \alpha_H) G_Q(t_u)^2} \\
=&\frac{N}{G_Q(t_u)} - \frac{\alpha_L}{\alpha_L - \alpha_H}\left(\rho(\rho+\alpha_H)I - \frac{(\gamma-1)N}{G_Q(t_u)}\right) - \frac{(\gamma-1)N\mu\mathbf{1}(t_u \leq T_e)}{(\alpha_L-\alpha_H)G_Q(t_u)^2}.
\end{aligned} \quad \text{(OA.3.7)}$$

3. $t \in (t_{u0,1}, \bar{t}_u)$, $G_H(t)$ *can be characterized by the following ODE:*

$$\begin{aligned}
&G'_H(t_u)\left[\frac{(\gamma-1)^2 N}{(\alpha_L - \alpha_H) G_Q(t_u)^2} + \frac{N e^{-\alpha_L(t_u - t_e)}}{\mu G_Q(t_e)}\right] \\
=&\frac{N}{G_Q(t_u)} - \frac{\alpha_L}{\alpha_L - \alpha_H}\left(\rho(\rho+\alpha_H)I - \frac{(\gamma-1)N}{G_Q(t_u)}\right) - \frac{(\gamma-1)N\mu\mathbf{1}(t_u \leq T_e)}{(\alpha_L-\alpha_H)G_Q(t_u)^2}.
\end{aligned} \quad \text{(OA.3.8)}$$

   where

$$t_e = \frac{G(t_e) - G(0)}{\mu} = \frac{G_H(t_u) - G(0)}{\mu}, \quad G_Q(t_e) = (\gamma-1)G_H(t_e) + G_H(t_u).$$

   The initial condition is $G_H(t_{u0,1}) = G(0)$.

4. $t \in [\bar{t}_u, \infty)$, *no firm upgrades and* $G_H(t) = G_{H,ss}$

Proof. To derive $G_H(t)$ after $t_{u0,0}$, we consider a firm that enters at t_e and upgrades at t_u, and another firm that enters at $t'_e \geq t_e$ and upgrades at $t'_u > t_u$ (see Lemma OA.3.3). Note that we have ruled out multiplicity in optimal upgrading time using Assumption OA.3.3 when $t'_e > t_e$. Denoting $\Delta t_e \equiv t'_e - t_e$ and $\Delta t_u \equiv t'_u - t_u$, we have

$$\begin{aligned}
&m(t_u; t_e) = m(t'_e; t_e)e^{-\alpha_L(t_u - t'_e)} + m(t_u; t'_e) \Rightarrow \\
&m(t_u; t'_e) \approx m(t_u; t_e) - \frac{Ne^{-\alpha_L(t_u - t_e)}}{G_Q(t_e)}\Delta t_e \\
&m(t'_u; t'_e) = m(t_u; t'_e)(1 - \alpha_L \Delta t_u) + \frac{N}{G_Q(t_u)}\Delta t_u \Rightarrow \\
&m(t'_u; t'_e) - m(t_u; t_e) = -\alpha_L m(t_u; t_e)\Delta t_u - \frac{Ne^{-\alpha_L(t_u - t_e)}}{G_Q(t_e)}\Delta t_e + \frac{N}{G_Q(t_u)}\Delta t_u \qquad \text{(OA.3.9)}
\end{aligned}$$

To obtain an ODE for $G_H(t)$, we use the following relationships

$$G_H(t_u) = G_0 + \mu t_e, \quad G_H(t'_u) = G_0 + \mu t'_e.$$

From these equations, we can easily back out $t_e = \frac{G_H(t_u) - G_0}{\mu}$ given $t_u, G_H(t_u)$. When $t'_e > t_e$, we can also express Δt_e as

$$\Delta t_e = \frac{G_H(t'_u) - G_H(t_u)}{\mu} \approx \frac{G'_H(t_u)\Delta t_u}{\mu}. \qquad \text{(OA.3.10)}$$

Applying the cutoff equation (OA.3.4), we have

$$m(t'_u; t'_e) - m(t_u; t_e) = \frac{(\gamma - 1)N}{\alpha_L - \alpha_H}\left(\frac{1}{G_Q(t_u)} - \frac{1}{G_Q(t'_u)}\right) = \frac{(\gamma - 1)NG'_Q(t_u)}{(\alpha_L - \alpha_H)G_Q(t_u)^2}\Delta t_u \quad \text{(OA.3.11)}$$

where

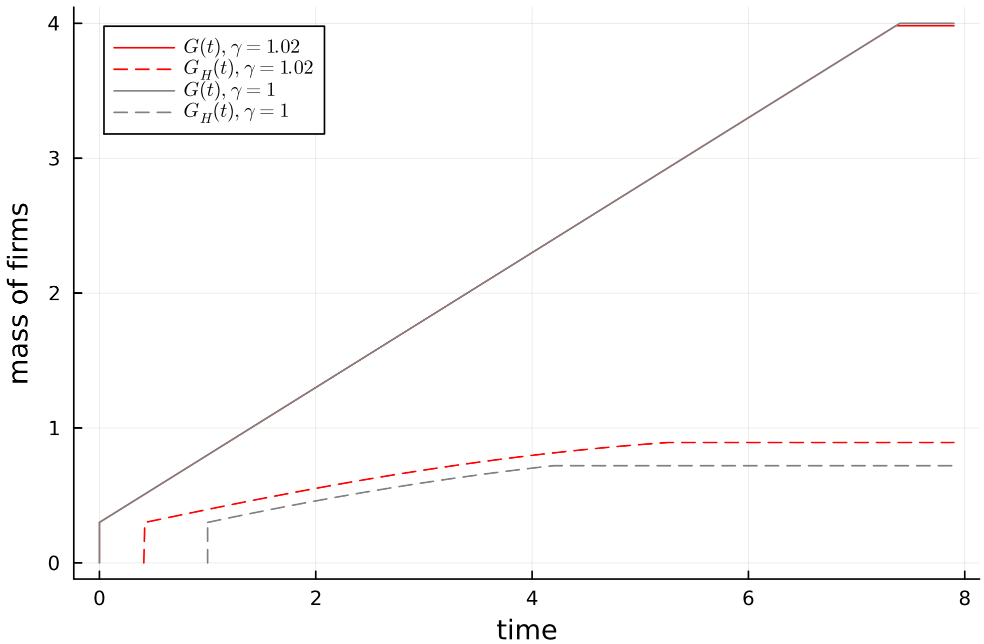
$$G'_Q(t_u) = \mu \mathbf{1}(t_u \leq T_e) + (\gamma - 1)G'_H(t_u)$$

Assuming $t_e = t'_e$, we combine equations (OA.3.9) and (OA.3.11) and obtain the ODE in (OA.3.7). Assuming $t'_e > t_e$, we combine equations (OA.3.9), (OA.3.10) and (OA.3.11) and obtain the ODE in (OA.3.8). ■

We can compare the equilibrium paths of $G_H(t)$ and $G(t)$ when $\gamma > 1$ and $\gamma = 1$ (baseline model). In the baseline model, all initial entrants upgrade at a particular time t_{u0} so $G_H(t)$ has a jump to $G(0)$ there. This is consistent with setting $\gamma = 1$ in equation (OA.3.7), which implies $G'_H(t_{u0}) \to \infty$. We can also set $\gamma = 1$ in equation (OA.3.8) and obtain the path of $G_H(t)$ in the baseline model. In Figure OA.3.1, we illustrate the impact of γ on the equilibrium paths. The grey solid and dashed lines represent $G(t)$ and $G_H(t)$ in our baseline model, respectively. We then raise γ from 1 to 1.02, and plot the corresponding $G(t)$ and

$G_H(t)$ with red solid and red dashed lines. When γ is higher, we see upgrading happens earlier and there are more steady-state H-type firms. More H-type firms also "crowd-out" L-type firms and the total mass of firms is lower when γ is higher.

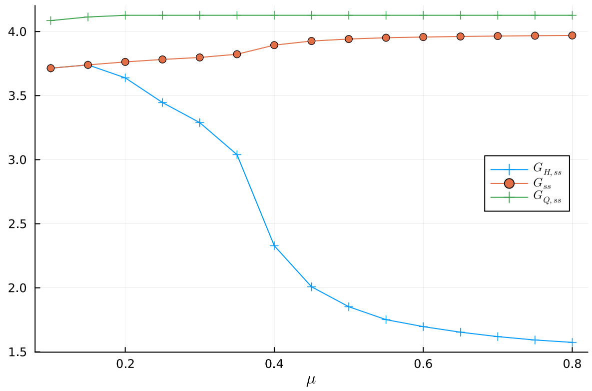
Figure OA.3.1: Compare $\gamma > 1$ and $\gamma = 1$



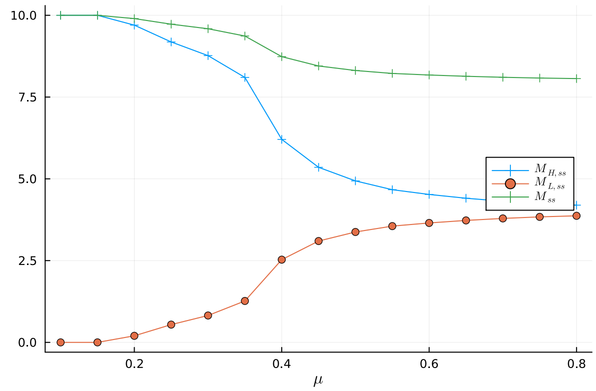
Notes: Base set of parameters $c = 1, N = 1, \rho = 0.1, \alpha_H = 0.1, \alpha_L = 0.15, \mu = 0.5, I = 4.5$.

Figure OA.3.2, we present comparative statics of steady-state mass of firms and total customers served with respect to market entry rate μ and flow fixed cost c. As we increase μ, firms face stronger competition from higher $G(t)$ at any time point, which reduces steady-state share and mass of H-type firms. The reduction in the share of H-type firms also reduces the total number of customers served, M_{lr}. A reduction in flow fixed cost c has a similar effect to an increase μ, since a lower c would support more firms in the market, thus raises G_{ss} and lowers $G_{H,ss}$. This eventually reduces the total number of customers served. In sum, we confirm that the comparative statics results are similar to the baseline model, at least within the range of parameters considered in Figure OA.3.1.

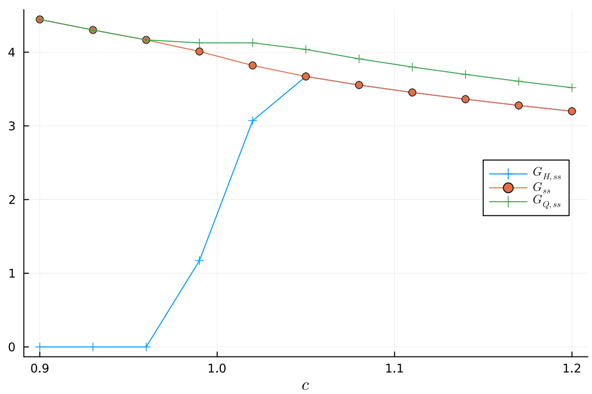
Figure OA.3.2: Comparative statics with respect to μ and c



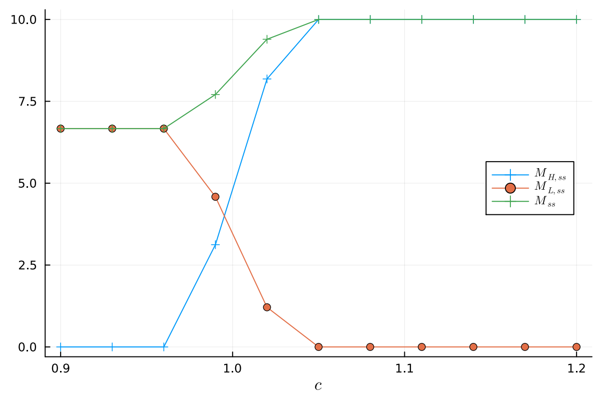
(a) mass of firms and μ



(b) customers served and μ



(c) mass of firms and c



(d) customers served and c

Notes: Base set of parameters $\rho = 0.1, N = 1, c = 1, \alpha_H = 0.1, \alpha_L = 0.15, \mu = 0.5, I = 5.25, \gamma = 1.1, G_0 = 1.5$.

In general, it is difficult to prove comparative statics as in our baseline model. However, we can compare the cases with sufficiently large and small market entry rates as follows

Proposition OA.3.1 *Suppose the upgrading cost is larger than a threshold value*

$$I > \frac{\rho + \alpha_L}{\rho(\rho + \alpha_H)} \left(\gamma - 1 + \frac{\alpha_L - \alpha_H}{\alpha_L} \right) c.$$

No firm upgrades when μ is sufficiently large and some firms upgrade when μ is sufficiently small.

Proof. We first consider the extreme case that all firms enter at $t = 0$, which corresponds to $\mu = \infty$. Since we assume that no firms exit, we have $G(t) = G_{ss}$ for all t. The value of $G(0)$ becomes irrelevant. We examine whether a typical firm' customer capital will reach the cutoff $m^*(t)$ assuming that no firms upgrade. In this case, we have $G_Q(t) = G_{ss}$ for all t and

$$m^*(t) = \frac{\rho(\rho + \alpha_H)I - \frac{(\gamma-1)N}{G_{ss}}}{\alpha_L - \alpha_H}.$$

The steady-state customer capital is

$$m_{L,lr} = \frac{N}{G_{ss}\alpha_L}.$$

The zero-profit condition at $t = 0$ implies that G_{ss} is the same as the expression (5) in the paper. Substitute G_{ss} into $m^*(t)$ and $m_{L,lr}$, we see that the restriction stated in the proposition ensures that

$$m^*(t) > m_{L,lr}.$$

We know that $m(t;0)$ increases and converges to $m_{L,lr}$. Therefore, $m(t;0)$ never crosses $m^*(t)$. According to Lemma OA.3.1, no firm upgrades.

This result applies to all μ that are sufficiently large. Under large but finite μ, $G(t)$ reaches G_{ss} at an arbitrarily small T_e. Therefore, the customer capital at T_e can be bounded as

$$m(T_e;0) = \int_0^{T_e} \frac{Ne^{-\alpha_L t}}{G(t)}\mathrm{d}t < \int_0^{T_e} \frac{N}{G(0)}\mathrm{d}t = \frac{NT_e}{G(0)} < m_{L,lr}.$$

Therefore, $m(t;0)$ does not cross $m^*(t)$ with sufficiently large μ, and no firm will upgrade.

When μ is sufficiently small, we search for a sufficient condition for the initial entrants to have positive value of upgrading. Assume that these firms upgrade at $t_u = 0$. We can re-write (OA.3.5) as

$$\Delta\tilde{V}(0;0) = \int_0^{\infty} e^{-\rho t} \int_0^t \left(A_H(s)e^{-\alpha_H(t-s)} - A_L(s)e^{-\alpha_L(t-s)}\right)\mathrm{d}s\mathrm{d}t - I.$$

We now perform the integration up to T instead of ∞. Since the integrand is always positive, we have

$$\begin{aligned}
\Delta\tilde{V}(0;0) &> \int_0^{T} e^{-\rho t} \int_0^t \left(A_H(s)e^{-\alpha_H(t-s)} - A_L(s)e^{-\alpha_L(t-s)}\right)\mathrm{d}s\mathrm{d}t - I \\
&> \int_0^{T} e^{-\rho t} \int_0^t \left(\frac{\gamma Ne^{-\alpha_H(t-s)}}{G(t)} - \frac{Ne^{-\alpha_L(t-s)}}{G(t)}\right)\mathrm{d}s\mathrm{d}t - I \\
&> \int_0^{T} e^{-\rho t} \int_0^t \left(\frac{Ne^{-\alpha_H(t-s)}}{G(T)} - \frac{Ne^{-\alpha_L(t-s)}}{G(T)}\right)\mathrm{d}s\mathrm{d}t - I
\end{aligned}$$

For any value of T and ε, we can find sufficiently small μ such that

$$G(T) - G(0) < \varepsilon.$$

Therefore,

$$\Delta\tilde{V}(0;0) > \frac{N}{G(0)+\varepsilon}\left[\frac{1-e^{\rho T}}{\rho\alpha_H} - \frac{1-e^{-(\rho+\alpha_H)T}}{\alpha_H(\rho+\alpha_H)} - \frac{1-e^{-\rho T}}{\rho\alpha_L} + \frac{1-e^{-(\rho+\alpha_L)T}}{\alpha_L(\rho+\alpha_L)}\right].$$

When T is sufficiently large, the term in the bracket will be sufficiently close to its limit, $\frac{\alpha_L-\alpha_H}{\rho(\rho+\alpha_L)(\rho+\alpha_H)}$. Suppose we set $\varepsilon = G(0)$, we can restrict $G(0) < \frac{(\alpha_L-\alpha_H)N}{2\rho(\rho+\alpha_L)(\rho+\alpha_H)I}$ so that for sufficiently small μ, initial entrants find positive value of upgrading. Note that this restriction can be fully consistent with $G(0) > \underline{G}_0$. This requires

$$\frac{(\gamma-1)N}{\rho(\rho+\alpha_H)I} < \frac{(\alpha_L-\alpha_H)N}{2\rho(\rho+\alpha_L)(\rho+\alpha_H)I} \Leftrightarrow \gamma - 1 < \frac{\rho(\alpha_L-\alpha_H)}{2(\rho+\alpha_L)}.$$

■

References

Train, Kenneth, *Discrete Choice Methods With Simulation*, Cambridge University Press, 2003.