Selection and Sorting of Heterogeneous Firms through Competitive Pressures*

Kiminori Matsuyama
Northwestern University, USA
k-matsuyama@northwestern.edu

Philip Ushchev
HSE University, St.Petersburg, Russia
fuschev@hse.ru

2022-03-06: 11:43 AM

Abstract: To understand theoretically how competitive pressures affect selection and sorting of firms with different productivity, we study the Melitz (2003) model under the H.S.A. (*Homothetic with a Single Aggregator)* class of demand systems. H.S.A. is tractable due to its homotheticity and to its single aggregator that serves as a sufficient statistic for competitive pressures, which acts as a magnifier of firm heterogeneity. It is also flexible enough to allow for the choke price, the 2nd law of demand--"a higher price leads to a higher price elasticity"--, and the 3rd law of demand--"a higher price leads to a smaller rate of change in the price elasticity." We show, among others:

- More productive firms have higher profits and revenues; they have higher markup rates under the 2nd law and lower pass-through rates under the 3rd law. Employments are not monotone in firm productivity; they are *hump-shaped* under the 2nd and 3rd laws. The 2nd law also implies the procompetitive effect and strategic complementarity in pricing.
- *A lower entry cost* leads to more competitive pressures, which reduces the markup rates of all firms under the 2nd law and raises the pass-through rates of all firms under the 3rd law. The profits of all firms decline (at faster rates among less productive firms under the 2nd law), which leads to a tougher selection. The revenues of all firms also decline (at faster rates among less productive firms under the 3rd law). *A lower overhead cost* has similar effects when the employment is decreasing in firm productivity, which occurs under the 2nd and the 3rd laws for a sufficiently high overhead cost.
- *Larger market size* also leads to more competitive pressures, reducing the markup rates of all firms under the 2nd law and raises the pass-through rates of all firms under the 3rd law. The profits among more productive firms increase, while those among less productive decline under the 2nd law, which leads to a tougher selection. The revenues among more productive firms also increase, while those among less productive decline under the 3rd law at least when the overhead cost is not too large.
- The impacts on the masses of entrants and of active firms depend, often crucially, on whether the elasticity of the distribution of the marginal cost is increasing or decreasing with Pareto-distributed productivity being the knife-edge case.
- Both a lower entry cost and larger market size may cause an *increase* in the average markup rate under the 2nd law and a *decline* in the average pass-through under the 3rd law due to *the composition effect,* since they also lead to a tougher selection, forcing less productive firms with lower markup rates and higher pass-through rates to shrink and to exit. This suggests that a rise of the markup may occur due to *increased* competitive pressures, causing a shift from the less productive/smaller to the more productive/larger.
- In a multi-market setting, competitive pressures are stronger in larger markets. And more productive firms sort themselves into larger markets under the 2nd law. Due to this *composition effect*, the average markup (pass-through) rates can be *higher* (*lower* under the 3rd Law) in larger (thus more competitive) markets. This result suggests a caution when interpreting the evidence that compares the average markup and pass-through rates across markets with different sizes.

Keywords: Heterogeneous firms, The Melitz model, H.S.A., market size, competitive pressures, the 2nd law, markup rates, the 3rd law, pass-through rates, selection, sorting, the composition effect, log-supermodularity.

*We thank the conference and seminar participants at (chronologically) Berkeley, VEAMS (Virtual East Asia Macroeconomics Seminar Series), Tokyo, Michigan, MIT, and Hitotsubashi Conference on International Trade and FDI for their valuable comments.

1. Introduction

How do firms with different productivity respond differently to increased competitive pressures caused by a lower entry cost or an increase in market size? How do these changes affect selection of heterogeneous firms? Or sorting of heterogeneous firms across different markets? And what are the impacts on the firm distribution? In the Melitz (2003) model of monopolistic competition with firm heterogeneity, its assumption of the CES demand system implies that all firms sell their products at an exogenous and common markup rate. The pricing behaviors of all firms are thus unresponsive to competitive pressures. Furthermore, a change in market size has no effect on the distribution of the firm types and their behaviors, with all adjustments taking place at the extensive margin.

In this paper, we extend the (closed economy version) of the Melitz (2003) model by relaxing the CES assumption, thereby allowing for heterogeneous firms to set different markup rates, which are responsive to a change in competitive pressures. We do so by using the H.S.A. (*Homotheticity with a Single Aggregator*) class of demand systems, originally introduced by Matsuyama and Ushchev (2017) and first applied to monopolistic competition by Matsuyama and Ushchev (2022).[1] The H.S.A. class of demand systems has many attractive features that make it suitable for the Melitz model

First, H.S.A. is homothetic, unlike most non-CES demand systems that have been applied to monopolistic competition.[2] Homotheticity allows us to define unambiguously a single measure of market size, even though market size can change for a variety of reasons, such as labor productivity growth, globalization, a sectoral shift in demand, a change in the population size, etc., because the composition of market demand does not matter.[3] It also helps to isolate the

[1] More recent applications of H.S.A. to monopolistic competition include Matsuyama and Ushchev (2020a, 2020b), Baqaee, Fahri, and Sangani (2021), Fujiwara and Matsuyama (2022), and Grossman, Helpman, and Lhuiller (2021).

[2] For example, Dixit and Stiglitz (1977, Section II) extended their monopolistic competition model to the directly explicitly additive (DEA) demand systems, which have been further explored by Krugman (1979), Behrens and Murata (2007), Zhelobodko, et.al. (2012), Melitz (2018), Dhingra and Morrow (2019), Latzer, Matsuyama, and Parenti (2019), Behrens et.al. (2020), among many others. This class can be also used to rationalize the reduced-form profit functions assumed in Mrázová-Neary (2017; 2019) and Nocke (2006). Though Dixit and Stiglitz called this class, "Variable Elasticity Case," the well-known Bergson's Law states that, within this class of demand systems, they are homothetic if and only if they are CES. In other words, any departure from CES within this class introduces nonhomotheticity. The linear-quadratic demand system introduced by Ottaviano, Tabuchi, and Thisse, (2002) and applied to the Melitz model by Melitz and Ottaviano (2008) is also nonhomothetic. See Thisse and Ushchev (2018) for a survey of monopolistic competition with non-CES demand systems. Parenti, Ushchev and Thisse (2017) provides a unified treatment of this literature.

[3] Using the linear-quadratic demand system with the outside good, Melitz and Ottaviano (2008) studied the market size effect by changing the population size. Many of the comparative statics go in the opposite directions, if the

effects of endogenous markup rates from those of nonhomotheticity. Furthermore, homotheticity makes it straightforward to use the Melitz model under H.S.A. as a building block in multi-sector general equilibrium models.

Second, H.S.A. is flexible. It can accommodate (but does not necessitate) the choke price, as well as the so-called Marshall's 2nd law of demand, "a higher price leads to a higher price elasticity," which implies incomplete pass-through--less productive firms have lower markup rates--, and what we call the 3rd law of demand, "a higher price leads to a smaller rate of change in the price elasticity," which implies that less productive firms have higher pass-through rate,[4] for which there is some supporting empirical evidence.[5] Furthermore, since this class contains CES (as well as translog) as a special case, H.S.A. can be used to perform the robustness check; it helps us understand which properties of the original Melitz model carry over to a broader class of the demand system.[6]

Third, the Melitz model under H.S.A. retains much of the tractability of the original Melitz model under CES. This is partly due to its homotheticity but also due to its single aggregator, which serves as a sufficient statistic for capturing any change in competitive pressures, whether caused by a change in the mass of active firms or by a change in the prices of competing products. The single aggregator enters all firm-specific variables (the markup and pass-through rates, the profit, the revenue and the employment) proportionately with the firm's marginal cost, hence acting as a magnifier of firm heterogeneity. This allows us to take advantage of log-supermodularity[7] to study the differential impacts of competitive pressures on

---

market size effect is studied by changing the per capita expenditure with a shock to the weight attached to the outside good in the preferences. Also under DEA, how firms respond to a market size change depends on whether it is caused by a change in the population size or by a change in the per capita expenditure.

[4] Regarding the terminology, Marshall's 1st law of demand states that a higher price reduces demand; it imposes the restriction on the 1st derivative of the demand curve. The 2nd law states that a higher price increases the price elasticity; it imposes the restriction on the 2nd derivative. We call the law stating that a higher price reduces the rate of change in the price elasticity as the 3rd law because it imposes the restriction on the 3rd derivative.

[5] For the empirical evidence on the 2nd law and incomplete pass-through, as well as the closely related concepts of the procompetitive effect and strategic complementarity in pricing, see Campbell and Hopenhayn (2005); Burstein-Gopinath (2014), DeLoecker and Goldberg (2014), Feenstra and Weinstein (2017), and Amiti, Itskhoki, and Konings (2019); For the empirical evidence on the 3rd law, see Berman, Martin, and Mayer (2012) and Amiti, Itskhoki, and Konings (2014). Recently, Baqaee, Fahri, and Sangani (2021, Appendix G) nonparametrically calibrated H.S.A. using the firm-level data from Belgium in support of the 2nd and the 3rd laws.

[6] In contrast, translog, applied to monopolistic competition by Feenstra (2003) and others, imposes the 2nd law, while violating the 3rd law. It is also an isolated example and hence cannot be used as a tool for the robustness check for CES. This motivated Matsuyama and Ushchev (2020a, 2022) to develop Generalized Translog, a family within H.S.A. that nests both CES and translog. See Appendix D.1.

[7] See, for example, Costinot (2009) and Costinot and Vogel (2010; 2015).

heterogeneous firms. It also enables us to use simple diagrams to prove the existence and uniqueness of free-entry equilibrium with firm heterogeneity and to conduct most comparative statics without imposing any parametric restrictions on the demand system and productivity distribution.[8] Moreover, unlike Melitz and Ottaviano (2008) and Arkolakis et.al. (2019) and many others that introduce the procompetitive effect in the Melitz model, there is no need to assume zero overhead cost for tractability. This is important not only because it makes the Melitz model under H.S.A. applicable also to the sectors characterized by high overhead costs, but also because it allows us to study the effects of the recent rise in overhead costs. Indeed, a combination of firm heterogeneity and the 2nd and 3rd laws of demand generates some new insights when the overhead cost is sufficiently high.[9]

Here are the main findings on the Melitz model under H.S.A.

- More productive firms have higher profits and revenues. They have higher markup rates under the 2nd law and lower pass-through rates under the 3rd law. Employments are not monotone in firm productivity; they are *hump-shaped* under the 2nd and 3rd laws. The 2nd law also implies the procompetitive effect and strategic complementarity in pricing.
- *A lower entry cost* leads to more competitive pressures, which reduces the markup rates of all firms under the 2nd law and raises the pass-through rates of all firms under the 3rd law. The profits of all firms decline (at faster rates among less productive firms under the 2nd law), which leads to a tougher selection. The revenues of all firms also decline (at faster rates among less productive firms under the 3rd law). *A lower overhead cost* has similar effects when the employment is decreasing in firm productivity, which occurs under the 2nd and the 3rd laws for a sufficiently high overhead cost.
- *Larger market size* also leads to more competitive pressures, reducing the markup rates of all firms under the 2nd law and raises the pass-through rates of all firms under the 3rd law. The

[8] In contrast, under the two other classes of demand systems studied in Matsuyama and Ushchev (2020a), HDIA, which contains the Kimball (1995) demand system as a special case, and HIIA, we need the two aggregators, one for competitive pressures due to a change in the pricing of competing firms, and another for competitive pressures due to a change in the mass of firms. This poses a challenge for ensuring the existence and the uniqueness of the free-entry equilibrium even in a single-market setting, since it would require further restrictions on the firm productivity distribution and the demand system. (Matsuyama and Ushchev (2020a) found the condition of the existence and the uniqueness under HDIA and HIIA only for the case of homogeneous firms.) The problem of ensuring the existence and the uniqueness under HDIA and HIIA would be even more challenging in a multi-market setting, which we develop in section 6 to study sorting of firms across markets.

[9] Another advantage of H.S.A., pointed out by Kasahara and Sugita (2020), is that the market share (in revenue) functions are the primitive of H.S.A., hence it can be readily identified with the typical firm-level data, which contain revenue, but not the output.

profits among more productive firms increase, while those among less productive decline under the 2nd law, which leads to a tougher selection. The revenues among more productive firms also increase, while those among less productive decline under the 3rd law at least when the overhead cost is not too large.

- The impacts on the masses of entrants and of active firms depend, often critically, on whether the elasticity of the distribution of the marginal cost is increasing or decreasing with Pareto-distributed productivity being the knife-edge case.
- Both a lower entry cost and larger market size may cause an *increase* in the average markup rate under the 2nd law and a *decline* in the average pass-through rate under the 3rd law due to the *composition effect*, since they also lead to a tougher selection, forcing less productive firms with lower markup rates and higher pass-through rates to shrink and to exit. This suggests that a rise of the markup may occur due to *increased* competitive pressures, causing a shift from the less productive, smaller firms to the more productive, larger firms, hence it should not be interpreted as the prima-facie evidence for reduced competitive pressures.
- In a multi-market setting, competitive pressures are stronger in larger markets. And more productive firms sort themselves into larger markets under the 2nd Law. Due to this *composition effect*, the average markup (pass-through) rates can be *higher* (*lower* under the 3rd Law) in larger (thus more competitive) markets. This result suggests a caution when interpreting the evidence that compares the average markup and pass-through rates across markets with different sizes.

Here's the roadmap. In section 2, we formally introduce the H.S.A. class of demand systems and apply it to the (closed economy version of) Melitz model. We show, under some mild regularity conditions, that the markup and pass-through rates of firms with the marginal cost ψ can be expressed as $\mu(\psi/A)$ and $\rho(\psi/A)$, both differentiable functions of a single variable, $\psi/A,$ the firm's "normalized cost", where A is the inverse measure of competitive pressures; it is the equilibrium value of the single aggregator, which serves a sufficient statistic that captures all the equilibrium interactions across firms, and acts as a magnifier of firm heterogeneity. We also show that the profit, the revenue, and the employment of a ψ-firm can be expressed as $\pi(\psi/A)L$, $r(\psi/A)L$ and $\ell(\psi/A)L$, all differentiable functions of $\psi/A,$ multiplied by market size L. Then, we derive the equilibrium conditions in terms of A and the cutoff marginal cost, ψ_c and show that the equilibrium is uniquely determined (Figure 1) as a differentiable function of F_e/L and F/L, where F_e is the entry cost and F the overhead cost.

In section 3, we revisit the Melitz model under CES, which implies constant markup rate $\mu(\psi/A) = \mu > 1$ and complete pass-through, $\rho(\psi/A) = 1$. We offer a simpler proof of the existence of the unique equilibrium (Figure 2) and a reproduction of the well-known results; We also show that the sign of the elasticity of the marginal cost distribution determines comparative statics on the masses of the entrance and active firms, with Pareto-distributed firm productivity being the knife-edge case (Proposition 1).

Then, we depart from CES. In section 4, we consider the cross-sectional implications of more competitive pressures (a lower A) under the 2nd law, i.e., when $\mu(\psi/A)$ is strictly decreasing (Proposition 2), and under the weak or strong 3rd law, i.e., when $\rho(\psi/A)$ is weakly or strictly increasing (Propositions 3, 4, and 5). These results are summarized in Figure 3. In section 5, we conduct general equilibrium analysis to study the impacts of changes in F_e, L and F on competitive pressures, A, and selection, ψ_c (Proposition 6; Figure 4). We look at the market size effect on the profit and the revenue (Proposition 7). Figure 5 puts together these results. Then, we study how the average markup and pass-through rates change through the composition effect (Proposition 8) and the effects on the masses of the entrants and active firms (Proposition 9). At the end of section 5, we look at the limit case of no overhead cost, where the cutoff firms operate at the choke price (Figure 6).

Then, in section 6, we consider a multi-market extension, in which each firm, after learning its productivity, decides whether to stay or exit and, if it stays, chooses among markets with different sizes. We show that larger markets are more competitive and that, under the 2nd law, there is a positive assortative matching between firm productivity and market size (Proposition 10; Figure 7). Then, we show the cross-sectional implications across markets (Figure 8). Due to the composition effect, the average markup rate may be higher and the average pass-through rate may be lower in larger markets, and a shock that increases competitive pressures in all markets may lead to higher average markup rates and lower average pass-through rates in all markets in spite of the 2nd law and the 3rd law (Proposition 11).

We conclude in Section 7. Appendices A through C contain some technical materials, including the proofs of some lemmas and propositions. Appendix D discuss three parametric families of H.S.A. and discuss their key properties.

2. Selection of Heterogeneous Firms

2.1. A Single-Market Setting

The representative household inelastically supplies L units of labor, the only primary factor of production, which we take as the *numeraire,* and consumes X units of the single final good subject to the budget constraint, $PX = L$, where P is the price of the final good.[10] The final good is produced competitively by assembling a set of differentiated intermediate inputs using CRS technology, which can be represented by the linear homogenous, monotone, and quasi-concave, production function, $X = X(\mathbf{x})$. Here, $\mathbf{x} = \{x_\omega; \omega \in \Omega\}$ is a quantity vector of intermediate inputs where Ω denotes a set of intermediate input varieties available, indexed by ω. Alternatively, the CRS technology can also be represented by the linear homogenous, monotone, and quasi-concave, unit cost function, $P = P(\mathbf{p})$, where $\mathbf{p} = \{p_\omega; \omega \in \Omega\}$ is a price vector of the intermediate inputs. The duality theory tells us that the production function, $X(\mathbf{x})$, and the unit cost function, $P(\mathbf{p})$, are related to each other as follows:

$$X(\mathbf{x}) \equiv \min_{\mathbf{p}} \left\{ \mathbf{px} = \int_\Omega p_\omega x_\omega d\omega \,\middle|\, P(\mathbf{p}) \geq 1 \right\}$$

$$P(\mathbf{p}) \equiv \min_{\mathbf{x}} \left\{ \mathbf{px} = \int_\Omega p_\omega x_\omega d\omega \,\middle|\, X(\mathbf{x}) \geq 1 \right\}$$

Hence, one could use either $P(\mathbf{p})$ or $X(\mathbf{x})$ as a primitive of the CRS technology.

The above minimization problems yield the demand curve and the inverse demand curve:

$$x_\omega = X(\mathbf{x}) \frac{\partial P(\mathbf{p})}{\partial p_\omega};\; p_\omega = P(\mathbf{p}) \frac{\partial X(\mathbf{x})}{\partial x_\omega}$$

for each intermediate input variety ω. From either of these, we can show, by using the Euler's theorem of linear homogenous functions,

$$\mathbf{px} = P(\mathbf{p})X(\mathbf{x}) = PX = L.$$

Market size for the intermediate inputs is thus equal to the aggregate income.[11] The market share of each variety can be expressed as

$$\frac{p_\omega x_\omega}{L} = \frac{p_\omega x_\omega}{P(\mathbf{p})X(\mathbf{x})} = \frac{\partial \ln P(\mathbf{p})}{\partial \ln p_\omega} = \frac{\partial \ln X(\mathbf{x})}{\partial \ln x_\omega} \tag{1}$$

---

[10] This budget constraint anticipates that monopolistic competitive firms collectively earn zero net profit in equilibrium due to the free-entry and hence the representative household receive no dividend income.

[11] This is due to the one-market setting. In a multi-market setting later, size of each market differs from L.

2.2. Symmetric H.S.A. Demand System with Gross Substitutes

Melitz (2003) assumed that the production function, $X(\mathbf{x})$, and its corresponding unit cost function, $P(\mathbf{p})$, is given by symmetric CES with gross substitutes. In Matsuyama and Ushchev (2017, section 3), we studied a class of homothetic functions that we called *Homothetic with a Single Aggregator* (H.S.A.), and in Matsuyama and Ushchev (2020a, 2022), we restrict this class further by defining over a continuum of varieties and imposing the symmetry and gross substitutability in order to make it applicable to monopolistic competitive settings.

More specifically, a symmetric CRS technology belongs to H.S.A. if it generates the demand system for inputs such that the market share of each input, eq.(1), can also be written as

$$\frac{p_\omega x_\omega}{L} = \frac{p_\omega}{P(\mathbf{p})}\frac{\partial P(\mathbf{p})}{\partial p_\omega} = s\left(\frac{p_\omega}{A(\mathbf{p})}\right). \tag{2}$$

Here, $s: \mathbb{R}_{++} \to \mathbb{R}_{+}$ is the *market share function,* which is *strictly decreasing* as long as $s(z) > 0$ with $\lim_{z\to\bar{z}} s(z) = 0$, where $\bar{z} \equiv \inf\{z > 0 | s(z) = 0\}$,[12] and $A(\mathbf{p})$ is linear homogenous in $\mathbf{p}$, defined implicitly by the adding-up constraint,

$$\int_\Omega s\left(\frac{p_\omega}{A(\mathbf{p})}\right) d\omega \equiv 1, \tag{3}$$

which ensures, by construction, that the market shares of all inputs are added up to one.[13] Symmetric CES with gross substitutes is a special case of H.S.A, with $s(z) = \gamma z^{1-\sigma}$ $(\sigma > 1)$. Symmetric translog is another special case, with $s(z) = \max\{-\gamma \ln(z/\bar{z}), 0\}$.[14] Appendix D offers more parametric examples of symmetric H.S.A.

---

[12] We need to ensure that the pass-through rate function defined later $\rho(\cdot)$ is continuous, for which it suffices to assume $s(\cdot) \in C^2(0, \bar{z})$. However, some of the proofs are much simpler if $\rho(\cdot)$ is continuously differentiable. Only for this expositional reason, we assume $s(\cdot) \in C^3(0, \bar{z})$ in this paper. All the parametric examples in this paper satisfy $s(\cdot) \in C^\infty(0, \bar{z})$. Matsuyama and Ushchev (2022; Appendix A) discusses how the analysis of monopolistic competition under H.S.A. might need to be modified if $s(\cdot)$ is *piecewise* $C^2(0, \bar{z})$, i.e., if it has some kinks.

[13] For $A(\mathbf{p})$ to be well-defined for all $\mathbf{p} = \{p_\omega; \omega \in \Omega\}$ for any Lebesgue measure of Ω, it is necessary to assume $\lim_{z\to 0} s(z) = \infty$. Though satisfied by CES and translog, this assumption would rule out some properties of the demand system we want to explore. Instead, we assume that L is not too small to ensure that there will be enough firms to enter in equilibrium so that $A(\mathbf{p})$ will be well-defined, as will be seen later.

[14] For $s: \mathbb{R}_{++} \to \mathbb{R}_{+}$, satisfying the above conditions, a class of the market share functions, $s_\gamma(z) \equiv \gamma s(z)$ for $\gamma > 0$, generate the same demand system with the same common price aggregator. We just need to renormalize the indices of varieties, as $\omega' = \gamma\omega$, so that $\int_\Omega s_\gamma(p_\omega/A(\mathbf{p}))d\omega = \int_\Omega s(p_{\omega'}/A(\mathbf{p}))d\omega' = 1$. In this sense, $s_\gamma(z) \equiv \gamma s(z)$ for $\gamma > 0$ are all equivalent. Note also that a class of the market share functions, $s_\lambda(z) \equiv s(\lambda z)$ for $\lambda > 0$, generate the same demand system, with $A_\lambda(\mathbf{p}) = \lambda A(\mathbf{p})$, because $s_\lambda(p_\omega/A_\lambda(\mathbf{p})) = s(\lambda p_\omega/A_\lambda(\mathbf{p})) = s(p_\omega/A(\mathbf{p}))$. In this sense, $s_\lambda(z) \equiv s(\lambda z)$ for $\lambda > 0$ are all equivalent. Using these equivalences, for example, one could obtain the CES case with $s(z) = z^{1-\sigma}$ $(\sigma > 1)$ by setting $\gamma = 1$ and the translog case, with $s(z) = \max\{-\ln(z/\bar{z}), 0\}$ by setting $\gamma = 1$ and $\lambda = 1/\bar{z} = 1$, without loss of generality.

Eqs.(2)-(3) state that the market share of an input is decreasing in its *relative price*, which is defined as its own price, p_ω, divided by the *common price aggregator,* $A(\mathbf{p})$. Notice that $A(\mathbf{p})$ is independent of ω; it is "the average input price" against which the relative prices of *all* inputs are measured. In other words, one could keep track of all the cross-price effects in the demand system by looking at a single aggregator, $A(\mathbf{p})$, which is the key feature of H.S.A. The assumption that the market share function, $s(\cdot)$, is independent of ω is not a defining feature of H.S.A.; it is due to the symmetry of the underlying production function that generates this demand system. The assumption that it is strictly decreasing in $z < \bar{z}$ means that inputs are gross substitutes. Furthermore, if $\bar{z} < \infty$, $\bar{z}A(\mathbf{p})$ is the choke price, at which demand for a variety goes to zero.

The unit cost function, $P(\mathbf{p})$, behind this H.S.A. demand system can be obtained by integrating eq.(2), which yields

$$\ln\left(\frac{P(\mathbf{p})}{A(\mathbf{p})}\right) = \text{const.} - \int_{\Omega} \left[\int_{p(\omega)/A(\mathbf{p})}^{\bar{z}} \frac{s(\xi)}{\xi} \mathrm{d}\xi\right] d\omega \tag{4}$$

The unit cost function, $P(\mathbf{p})$, satisfies the linear homogeneity, monotonicity, and strict quasi-concavity in the interior, and so does the corresponding production function, $X(\mathbf{x})$, which follows from Matsuyama and Ushchev (2017; Proposition 1-i)). This guarantees the existence of the underlying CRS technology, $X(\mathbf{x})$ or $P(\mathbf{p})$, that generates this H.S.A. demand system. In the case of CES, it is easy to verify that $P(\mathbf{p}) = cA(\mathbf{p})$, where $c > 0$ is a constant. However, it is important to note that, with the sole exception of CES, $P(\mathbf{p}) \neq cA(\mathbf{p})$ for any constant $c > 0$, as shown in Matsuyama and Ushchev (2020a; Corollary 2 of Lemma 2).[15] This should not come as a total surprise. After all, $A(\mathbf{p})$ is the "average input price", the inverse measure of competitive pressures for each input, which fully captures the *cross-price effects* in the demand system, while $P(\mathbf{p})$ is the inverse measure of TFP, which fully captures the *productivity (or welfare) effects* of price changes; there is no reason to think *a priori* that they should move together.[16]

[15]This holds also for asymmetric H.S.A., as well as H.S.A. with gross complements. See Matsuyama and Ushchev (2017; Proposition 1-iii).

[16] Note also that $A(\mathbf{p})$, the "average input price", depends on the unit of measurement of intermediate inputs, but not on the unit of measurement of the final good. In contrast, $P(\mathbf{p})$ is the cost of producing one unit of the final good, when the input prices are given by $\mathbf{p}$. Hence, it depends not only on the unit of measurement of intermediate inputs but also on that of the final good. Furthermore, a change in TFP affects $P(\mathbf{p})$ but not $A(\mathbf{p})$. This is the reason why eq.(4) has a constant term.

2.3. Monopolistically Competitive Differentiated Intermediate Inputs Producers

2.3.1. Timing

Differentiated intermediate inputs $\omega \in \Omega$ are produced in a monopolistically competitive industry a la Melitz, using labor (the numeraire) as the sole input, with the following timing.

- First, a continuum of ex-ante homogeneous monopolistically competitive firms, each identified by the input variety it produces and hence indexed by ω, decides whether to enter the industry. Every entrant pays a sunk cost $F_e > 0$, paid in labor.
- Second, each entrant draws its constant marginal cost $\psi \sim G(\psi)$, paid in labor, where $G(\psi)$ is a cdf, whose support is $\left(\underline{\psi}, \overline{\psi}\right) \subseteq (0, \infty)$.[17] Thus, firms become ex-post heterogeneous in their marginal costs of production. We assume that $G(\psi) \in C^3\left(\underline{\psi}, \overline{\psi}\right)$ and hence that its pdf, $g(\psi) \equiv G'(\psi) \in C^2\left(\underline{\psi}, \overline{\psi}\right)$, which ensures that $\mathcal{E}_G(\psi) \equiv \psi g(\psi)/G(\psi) \in C^2\left(\underline{\psi}, \overline{\psi}\right)$ and $\mathcal{E}_g(\psi) \equiv \psi g'(\psi)/g(\psi) \in C^1\left(\underline{\psi}, \overline{\psi}\right)$.[18],[19]
- After learning its constant marginal cost, ψ, each entrant chooses whether to exit without producing or stay and produce, in which case it pays an overhead cost $F > 0$. The set of firms that choose to stay and hence the set of intermediate input varieties produced is endogenously determined and denoted by Ω.
- Finally, each firm that chooses to stay sells its product at the profit-maximizing price.

2.3.2. Markup Rate and Pass-Through Rate Functions

After drawing its marginal cost, ψ_ω, firm ω would set its price p_ω to maximize its operating profit, if it would stay, as follows:

---

[17]Equivalently, each entrant draws its labor productivity, $\varphi = 1/\psi$, from its cdf, $1 - G(1/\varphi)$, whose support is $\varphi \in \left(\underline{\varphi}, \overline{\varphi}\right) \subseteq (0, \infty)$, with $\underline{\varphi} = 1/\overline{\psi}$ and $\overline{\varphi} = 1/\underline{\psi}$.

[18] For a differentiable positive-valued function $f(x) > 0$ of a single variable $x > 0$, we make frequent use of "the elasticity operator," $\mathcal{E}_f(x) \equiv d\ln f(x)/d\ln x = xf'(x)/f(x)$. Clearly, this operator satisfies the following properties: $\mathcal{E}_c(x) = 0$ and $\mathcal{E}_{cf}(x) = \mathcal{E}_f(x)$ for any constant $c > 0$; $\mathcal{E}_x(x) = 1$ for the identity function $x > 0$; $\mathcal{E}_{f_1 f_2}(x) = \mathcal{E}_{f_1}(x) + \mathcal{E}_{f_2}(x)$ for the product; $\mathcal{E}_{1/f}(x) = -\mathcal{E}_f(x)$ for the inverse; and the chain rule, $\mathcal{E}_{f_1 \circ f_2}(x) = \mathcal{E}_{f_1}(f_2(x))\mathcal{E}_{f_2}(x)$, for the composite $(f_1 \circ f_2)(x) \equiv f_1(f_2(x))$.

[19]We need to ensure that $\mathcal{E}_g(\cdot)$ is continuous, for which it suffices to assume $G(\cdot) \in C^2(\underline{\psi}, \overline{\psi})$. However, some of the proofs are much simpler if $\mathcal{E}_g(\cdot) \in C^1\left(\underline{\psi}, \overline{\psi}\right)$. Only for this expositional reason, we assume $G(\cdot) \in C^3(\underline{\psi}, \overline{\psi})$ in this paper. All the parametric distributions discussed in this paper satisfy $G(\cdot) \in C^\infty(\underline{\psi}, \overline{\psi})$.

$$\Pi_\omega = \max_{p_\omega}(p_\omega - \psi_\omega)x_\omega = \max_{\psi_\omega < p_\omega < \bar{z}A}\left(1 - \frac{\psi_\omega}{p_\omega}\right) s\left(\frac{p_\omega}{A}\right) L.$$

for $\psi_\omega/A < \bar{z}$, by taking L and A as given.[20] Or equivalently, it chooses its relative price, $z_\omega \equiv p_\omega/A < \bar{z}$, to solve

$$\max_{\psi_\omega/A < z_\omega < \bar{z}}\left(1 - \frac{\psi_\omega/A}{z_\omega}\right) s(z_\omega) \equiv \pi\left(\frac{\psi_\omega}{A}\right) > 0$$

for $\psi_\omega/A < \bar{z}$. The FOC is given by

$$z_\omega\left[1 - \frac{1}{\zeta(z_\omega)}\right] = \frac{\psi_\omega}{A},$$

with $\psi_\omega/A < z_\omega < \bar{z}$, where

$$\zeta(z) \equiv 1 - \frac{d\ln s(z)}{d\ln z} = 1 - \frac{zs'(z)}{s(z)} \equiv 1 - \mathcal{E}_s(z) > 1$$

is the price elasticity function, which is $C^2(0,\bar{z})$, satisfying $\lim_{z\to\bar{z}}\zeta(z) = -\lim_{z\to\bar{z}}\mathcal{E}_s(z) = \infty$, if $\bar{z} < \infty$. The markup rate is hence given by $\zeta(z_\omega)/(\zeta(z_\omega) - 1)$. Alternatively, starting from any price elasticity function satisfying $\zeta(z) > 1$ and $\lim_{z\to\bar{z}}\zeta(z) = \infty$, if $\bar{z} < \infty$, one could reverse-engineer the market share function as follows:

$$s(z) = \exp\left[\int_{z_0}^{z}\frac{1 - \zeta(\xi)}{\xi}d\xi\right]$$

where $z_0 \in (0,\bar{z})$ is a constant.

In what follows, we maintain the following regularity condition for the ease of exposition:

A1: For all $z \in (0,\bar{z})$,

$$\mathcal{E}_{z(\zeta-1)/\zeta}(z) = 1 + \mathcal{E}_{(\zeta-1)/\zeta}(z) = 1 + \frac{z\zeta'(z)}{[\zeta(z) - 1]\zeta(z)} = 1 - \frac{\mathcal{E}_\zeta(z)}{\mathcal{E}_s(z)} > 0$$

$$\Leftrightarrow \mathcal{E}_{\zeta/(\zeta-1)}(z) = \frac{\mathcal{E}_\zeta(z)}{\mathcal{E}_s(z)} < 1.$$

$$\Leftrightarrow \mathcal{E}_{s/\zeta}(z) = \mathcal{E}_s(z) - \mathcal{E}_\zeta(z) = \mathcal{E}_s(z)\left[1 + \mathcal{E}_{(\zeta-1)/\zeta}(z)\right] = 1 - \zeta(z) - \mathcal{E}_\zeta(z) < 0.$$

A1 states that the marginal revenue is strictly increasing in p_ω (hence strictly decreasing in x_ω) along the demand curve, so that LHS of FOC is strictly increasing in z (i.e., $\mathcal{E}_{z(\zeta-1)/\zeta}(z) > 0$). It is equivalent to the condition that the markup rate $\zeta(z)/(\zeta(z) - 1)$ cannot go up as fast as z

[20]For $\bar{z} < \infty$, no firm that draws $\psi_\omega > \bar{z}A$ would stay.

(i.e., $\mathcal{E}_{\zeta/(\zeta-1)}(z) < 1$), which is in turn equivalent to the condition that the price elasticity cannot go down as fast as the market share (i.e., $\mathcal{E}_{\zeta}(z) > \mathcal{E}_{s}(z)$). Since $\mathcal{E}_{s}(z) < 0$, **A1** holds if the price elasticity is increasing in z (i.e., $\mathcal{E}_{\zeta}(z) > 0$), hence the markup rate is decreasing in z (i.e., under **A2**, Marshall's 2nd Law, introduced later). **A1** is also equivalent to the condition that the profit when z satisfies FOC, $\Pi = [1 - \psi/(zA)]s(z)L = [s(z)/\zeta(z)]L$, is strictly decreasing in z.

Because LHS of FOC is C^2 and strictly increasing in z_ω under **A1**, the inverse function theorem implies that the profit maximizing relative price, z_ω, can be written as a strictly increasing C^2 function of the normalized cost, ψ_ω/A. Hence, the revenue, $R_\omega = s(z_\omega)L$, the profit, $\Pi_\omega = s(z_\omega)/\zeta(z_\omega)\, L$, can also be written as strictly decreasing C^2 functions of ψ_ω/A. The employment, $L_\omega = R_\omega - \Pi_\omega = [1 - 1/\zeta(z_\omega)]s(z_\omega)L$, can also be written as a C^2 function of ψ_ω/A.[21] Thus, all firms sharing the same ψ would set the same price and earn the same revenue and the same profit. Their outputs and employments are also the same. This allows us to index firms by their ψ. By denoting the profit-maximizing price of all ψ-firms by p_ψ and their relative price, $z_\psi \equiv p_\psi/A$, the FOC can now be written as:

Lerner Formula:

$$z_\psi\left[1 - \frac{1}{\zeta(z_\psi)}\right] = \frac{\psi}{A}$$

And the inverse function theorem allows us to solve for the profit-maximizing z_ψ as a strictly increasing C^2 function of $\psi/A \in (0, \bar{z})$:

Relative Price:

$$z_\psi \equiv \frac{p_\psi}{A} = Z\left(\frac{\psi}{A}\right)$$

satisfying $\psi/A < Z(\psi/A) < \bar{z}$ and $\lim\limits_{\psi/A \to \bar{z}} Z(\psi/A) = \bar{z}$. From this, the price elasticity at the point of the demand curve where ψ-firms choose to operate and their markup rate can both be written as C^2 function of $\psi/A \in (0, \bar{z})$:

Price Elasticity:

$$\zeta(z_\psi) = \zeta\left(Z\left(\frac{\psi}{A}\right)\right) \equiv \sigma\left(\frac{\psi}{A}\right) > 1,$$

Markup Rate:

$$\mu_\psi \equiv \frac{p_\psi}{\psi} = \frac{\zeta(Z(\psi/A))}{\zeta(Z(\psi/A)) - 1} = \frac{\sigma(\psi/A)}{\sigma(\psi/A) - 1} \equiv \mu\left(\frac{\psi}{A}\right) > 1.$$

---

[21]Even without A1, the profit maximizing z_ω would be strictly increasing and the maximized profit $\Pi_\omega = s(z_\omega)/\zeta(z_\omega)\, L$ would be strictly decreasing in the normalized cost ψ_ω/A. However, z_ω would be piecewise-continuous (i.e., it would jump up at some values of ψ_ω/A), and Π_ω would be piecewise-differentiable, which would complicate comparative static analysis.

Furthermore, by log-differentiating the Lerner formula, we obtain the pass-through rate as a C^1 function of $\psi/A \in (0, \bar{z})$:

Pass-Through Rate: $$\rho_\psi \equiv \frac{\partial \ln p_\psi}{\partial \ln \psi} = \mathcal{E}_Z\left(\frac{\psi}{A}\right) = \frac{1}{1 + \mathcal{E}_{1-1/\zeta}\big(Z(\psi/A)\big)} \equiv \rho\left(\frac{\psi}{A}\right) > 0,$$

where $\rho(\psi/A) > 0$ is ensured by **A1**. The elasticity of $\mu(\psi/A)$ is directly related to $\rho(\psi/A)$ as

$$\mathcal{E}_\mu\left(\frac{\psi}{A}\right) = -\mathcal{E}_{1/\mu} = -\mathcal{E}_{(1-1/\zeta)\circ Z}\left(\frac{\psi}{A}\right) = -\mathcal{E}_{1-1/\zeta}\left(Z\left(\frac{\psi}{A}\right)\right)\rho\left(\frac{\psi}{A}\right) = \rho\left(\frac{\psi}{A}\right) - 1.$$

It should be noted that, although $Z(\psi/A)$ is always strictly increasing in ψ/A, $\mu(\psi/A)$ and $\rho(\psi/A)$ can be increasing, decreasing, or nonmonotonic at this level of generality. Note also that market size, L, does not enter directly in $\mu(\psi/A)$ and $\rho(\psi/A)$, which means that market size may affect the markup and pass-through rates only indirectly through its effect on A.

2.3.3. Profit, Revenue and Employment Functions

The revenue, the (gross) profit and the (variable) employment are all expressed as functions of a single variable, ψ/A, multiplied by market size, L, as follows:

Revenue: $$R_\psi \equiv s(z_\psi)L = s\left(Z\left(\frac{\psi}{A}\right)\right)L \equiv r\left(\frac{\psi}{A}\right)L,$$

Profit: $$\Pi_\psi \equiv \left(1 - \frac{\psi/A}{z_\psi}\right)s(z_\psi)L = \frac{s\big(Z(\psi/A)\big)}{\zeta\big(Z(\psi/A)\big)}L = \frac{r(\psi/A)}{\sigma(\psi/A)}L \equiv \pi\left(\frac{\psi}{A}\right)L,$$

Employment: $$L_\omega \equiv R_\psi - \Pi_\psi = \left[1 - \frac{1}{\zeta(z_\psi)}\right]s(z_\psi)L = \frac{r(\psi/A)}{\mu(\psi/A)}L \equiv \ell\left(\frac{\psi}{A}\right)L,$$

Furthermore, the elasticities of $r(\psi/A)$, $\pi(\psi/A)$, and $\ell(\psi/A)$ can be written solely in terms of $\sigma(\psi/A)$ and $\rho(\psi/A)$ as follows:

$$\mathcal{E}_r\left(\frac{\psi}{A}\right) = \mathcal{E}_{s\circ Z}\left(\frac{\psi}{A}\right) = \mathcal{E}_s\left(Z\left(\frac{\psi}{A}\right)\right)\mathcal{E}_Z\left(\frac{\psi}{A}\right) = \left[1 - \sigma\left(\frac{\psi}{A}\right)\right]\rho\left(\frac{\psi}{A}\right) < 0.$$

$$\mathcal{E}_\pi\left(\frac{\psi}{A}\right) = \mathcal{E}_{s/\zeta}\left(Z\left(\frac{\psi}{A}\right)\right)\mathcal{E}_Z\left(\frac{\psi}{A}\right) = \mathcal{E}_s\left(Z\left(\frac{\psi}{A}\right)\right)\left[1 + \mathcal{E}_{1-1/\zeta}\left(Z\left(\frac{\psi}{A}\right)\right)\right]\rho\left(\frac{\psi}{A}\right) = 1 - \sigma\left(\frac{\psi}{A}\right) < 0$$

$$\mathcal{E}_\ell\left(\frac{\psi}{A}\right) = \mathcal{E}_r\left(\frac{\psi}{A}\right) - \mathcal{E}_\mu\left(\frac{\psi}{A}\right) = \mathcal{E}_r\left(\frac{\psi}{A}\right) + 1 - \rho\left(\frac{\psi}{A}\right) = 1 - \rho\left(\frac{\psi}{A}\right)\sigma\left(\frac{\psi}{A}\right) \gtreqless 0.$$

Because $\sigma(\cdot)$ is C^2 and $\rho(\cdot)$ is C^1, these elasticities are all C^1 functions of ψ/A. Since $\sigma(\cdot) > 1$, $\mathcal{E}_r(\cdot) < 0$ and $\mathcal{E}_\pi(\cdot) < 0$, and hence the revenue, $r(\psi/A)L$, and the profit, $\pi(\psi/A)L$, are always strictly decreasing in ψ/A. In contrast, $\mathcal{E}_\ell(\cdot)$ can change its sign, and hence the employment,

$\ell(\psi/A)L$, is generally nonmonotonic. However, its elasticity is related to those of the revenue and the markup rate. If the markup rate is decreasing in ψ/A (i.e., $-\mathcal{E}_\mu(\psi/A) > 0$), the employment cannot decline as fast as the revenue (i.e., $\mathcal{E}_\ell(\psi/A) = \mathcal{E}_r(\psi/A) - \mathcal{E}_\mu(\psi/A) > \mathcal{E}_r(\psi/A)$). Indeed, the employment is increasing in ψ/A, if the markup rate declines faster than the revenue (i.e., $-\mathcal{E}_\mu(\psi/A) > -\mathcal{E}_r(\psi/A) > 0$).

2.4. Equilibrium Conditions

Monopolistic competitive firms enter as long as their expected profit is equal to their entry cost. Assuming $F_e + F < \pi(0)L$, the free entry condition is given by

$$\int_{\underline{\psi}}^{\overline{\psi}} \max\{\Pi_\psi - F, 0\}\, dG(\psi) = \int_{\underline{\psi}}^{\overline{\psi}} \max\{\pi(\psi/A)L - F, 0\}\, dG(\psi) = F_e > 0.$$

where F_e is the sunk entry cost. Since $\pi(\psi/A)$ is strictly decreasing in ψ, there exists a unique **cutoff** level of the marginal cost, ψ_c, for each A given by

Cutoff Rule:

$$\pi\left(\frac{\psi_c}{A}\right)L = F \Leftrightarrow \frac{\psi_c}{A} = \pi^{-1}\left(\frac{F}{L}\right) < \bar{z} \tag{5}$$

such that firms stay and produce if $\psi \in \left(\underline{\psi}, \psi_c\right)$ and exit without producing if $\psi \in (\psi_c, \overline{\psi})$, assuming the interior solution, $0 < G(\psi_c) < 1$. Then, the free entry condition can be written as:

Free Entry Condition:

$$F_e = \int_{\underline{\psi}}^{\psi_c} \left[\pi\left(\frac{\psi}{A}\right)L - F\right] dG(\psi). \tag{6}$$

Figure 1 illustrates the cutoff, eq.(5), by the upward-sloping ray and the free entry condition, eq.(6), by the downward-sloping curve. The cutoff rule has a positive slope because more competitive pressures, a lower A, implies a tougher selection, a lower ψ_c**.** The free-entry condition has a negative slope because both more competitive pressures, a lower A and a tougher selection, a lower ψ_c, would make entry less attractive.[22] Clearly, these two conditions jointly determine the equilibrium values of $A = A(\mathbf{p})$ and ψ_c uniquely as C^2-functions of F_e/L and F/L. Furthermore, the interior solution, $0 < G(\psi_c) < 1$, is ensured under:

$$0 < \frac{F_e}{L} = \int_{\underline{\psi}}^{\psi_c} \left[\pi\left(\pi^{-1}\left(\frac{F}{L}\right)\frac{\psi}{\psi_c}\right) - \frac{F}{L}\right] dG(\psi) < \int_{\underline{\psi}}^{\overline{\psi}} \left[\pi\left(\pi^{-1}\left(\frac{F}{L}\right)\frac{\psi}{\overline{\psi}}\right) - \frac{F}{L}\right] dG(\psi),$$

[22]As $A \to \infty$, it is asymptotic to the horizontal line defined by $F_e = [\pi(0)L - F]G(\psi_c)$.

which is assumed to hold throughout the paper.[23] Note that this condition holds for a sufficiently small $F_e > 0$ with no further restrictions on $G(\cdot)$ or $s(\cdot)$.

Having $A = A(\mathbf{p})$ and ψ_c pinned down uniquely by eq.(5) and eq.(6), let us now turn to the mass of the entrants, M, that pay the sunk cost F_e.[24] By rewriting the adding-up constraint, eq.(3) as:

$$1 \equiv \int_{\Omega} s\left(\frac{p_\omega}{A}\right) d\omega = M \int_{\underline{\psi}}^{\psi_c} s\left(Z\left(\frac{\psi}{A}\right)\right) dG(\psi) = M \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi),$$

the equilibrium values of M can be given by:

$$M = \left[\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi)\right]^{-1} = \left[\int_{\underline{\xi}}^{1} r\left(\pi^{-1}\left(\frac{F}{L}\right)\xi\right) dG(\psi_c \xi)\right]^{-1} \tag{7}$$

as a C^2-function of F_e/L and F/L.

Eq.(5)-eq.(7) fully determine the equilibrium.[25] For the equilibrium value of $MG(\psi_c)$, the mass of firms that stay, which is equal to the Lebesgue measure of Ω, we can use eq.(7) to obtain

$$MG(\psi_c) = \left[\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) \frac{dG(\psi)}{G(\psi_c)}\right]^{-1} = \left[\int_{\underline{\xi}}^{1} r\left(\pi^{-1}\left(\frac{F}{L}\right)\xi\right) d\tilde{G}(\xi; \psi_c)\right]^{-1}, \tag{8}$$

where the second equality is obtained by changing variables as $\xi \equiv \psi/\psi_c$ with $\underline{\xi} \equiv \underline{\psi}/\psi_c$, and

$$\tilde{G}(\xi; \psi_c) \equiv \frac{G(\psi_c \xi)}{G(\psi_c)}$$

is the cdf of the marginal cost relative to the cutoff marginal cost among the firms that stay. Lemma 2 of Appendix A shows that a lower ψ_c (tougher selection) shifts $\tilde{G}(\xi; \psi_c)$ to the right in the MLR ordering if $\mathcal{E}_g(\psi) \equiv \psi g'(\psi)/g(\psi)$ is strictly decreasing, and to the right in the FSD

---

[23]For $\bar{\psi} = \infty$, this condition is reduced to $\pi(0)L > F_e + F > F_e > 0$, which is already assumed. For $\bar{\psi} < \infty$, the upper bound on F_e is less than $\pi(0)L - F$, and simple algebra can show that this upper bound is independent of L under CES, while increasing in L under **A2** introduced later.

[24]What makes H.S.A. particularly tractable is this recursive structure. Under HDIA and HIIA, the two other classes of the demand system studied in Matsuyama and Ushchev (2020a), the market share of each firm depends on the two aggregators, one affecting the pricing decision of the firm and the other its entry decision. As a result, the free-entry equilibrium is determined jointly by the three conditions. This complicates not only comparative statics, but also ensuring the existence and the uniqueness of the equilibrium, which requires further assumptions on the firm distribution and the demand system.

[25]The Walras Law ensures the labor market equilibrium. This can be verified as: labor demand per entrant $= F_e + FG(\psi_c) + \int_{\underline{\psi}}^{\psi_c} \ell\left(\frac{\psi}{A}\right) L dG(\psi) = \int_{\underline{\psi}}^{\psi_c} \left[\pi\left(\frac{\psi}{A}\right) L + \ell\left(\frac{\psi}{A}\right) L\right] dG(\psi) = L \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi) = L/M$, where eq.(5) and eq.(3), are used in the second and the last equalities. Of course, for these equilibrium conditions to be well-defined, the integrals in eq.(6) and eq.(7) must be finite, which is clearly the case if $\underline{\psi} > 0$. For $\underline{\psi} = 0$, Lemma 4 of Appendix B shows that $1 \le \lim_{z \to 0} \zeta(z) < 2 + \lim_{\psi \to 0} \mathcal{E}_g(\psi) < \infty$ is a sufficient condition.

ordering if $\mathcal{E}_G(\psi) \equiv \psi g(\psi)/G(\psi)$ is strictly decreasing. For example, $\mathcal{E}_g(\cdot)$ and $\mathcal{E}_G(\cdot)$ are both strictly decreasing for Fréchet, Weibull, and Lognormal, as shown in Appendix A.[26] If $G(\cdot)$ is a power function (hence, firm productivity is Pareto-distributed), $\mathcal{E}_g(\cdot)$ and $\mathcal{E}_G(\cdot)$ are both constant, and $\tilde{G}(\xi;\psi_c)$ is independent of ψ_c.[27]

3. CES Benchmark: The Original Melitz Model

As a benchmark, consider first the case studied by Melitz, CES, a special case of H.S.A., $\zeta(z) = \sigma > 1$ for all $z \in (0,\infty)$ or equivalently, $s(z) = \gamma z^{1-\sigma}$ for all $z \in (0,\infty)$. Even though the original Melitz model is well-known, it is instructive to obtain its properties as a special case of Melitz under H.S.A., because his analysis and its countless reproduction by those who follow-- see a survey by Melitz and Redding (2014)--make heavy use of CES from the very beginning. This makes it hard to see which properties of the Melitz model are specific to CES or which ones can be generalized under H.S.A.

The markup rate is simply $\mu(\psi/A) = \sigma/(\sigma-1)$, hence uniform across all firms and unaffected by L, F_e, F, $G(\cdot)$, A, ψ_c, thus it never changes across equilibriums, and the pass-through rate is $\rho(\psi/A) = 1$. The profit is $\pi(\psi/A)L = c_0 L(\psi/A)^{1-\sigma}$, where $c_0 \equiv (\gamma/\sigma)(1-1/\sigma)^{\sigma-1}$. Thus, the cutoff rule, eq.(5), and free entry condition, eq.(6), become:

Cutoff Rule:
$$c_0 L\left(\frac{\psi_c}{A}\right)^{1-\sigma} = F.$$

Free Entry Condition:
$$\int_{\underline{\psi}}^{\psi_c}\left[c_0 L\left(\frac{\psi}{A}\right)^{1-\sigma} - F\right] dG(\psi) = F_e;$$

Figure 2 depicts the upward-sloping cutoff rule and the downward-sloping free-entry condition with the unique intersection.[28] An increase in L would shift both conditions to the left, so that the intersection would move to the left. To see how the intersection moves, eliminate L from these two conditions to obtain

---

[26]Lemma 1 of Appendix A shows that $\mathcal{E}_g'(\psi) < 0$ always implies $\mathcal{E}_G'(\psi) < 0$, while $\mathcal{E}_g'(\psi) \geq 0$ implies $\mathcal{E}_G'(\psi) \geq 0$ only with some additional conditions.

[27] The unit cost, $P = P(\mathbf{p})$, and TFP, $X/L = X(\mathbf{x})/L$, can be obtained from eq.(4), as $\ln(X/L) = \ln(1/P) = \ln\left(\frac{1}{cA}\right) + M\int_0^{\psi_c}\left[\int_{Z(\psi/A)}^{\bar{z}}\frac{s(\xi)}{\xi}d\xi\right]dG(\psi) = \ln\left(\frac{1}{cA}\right) + \frac{\int_{\underline{\psi}}^{\psi_c}\left[\int_{Z(\psi/A)}^{\bar{z}}\frac{s(\xi)}{\xi}d\xi\right]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c} s\left(Z\left(\frac{\psi}{A}\right)\right)dG(\psi)} = \ln\left(\frac{1}{cA}\right) + \frac{\int_{\underline{\psi}}^{\psi_c}\left[\int_{Z(\psi/A)}^{\bar{z}}\frac{s(\xi)}{\xi}d\xi\right]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c}\left[\int_{Z(\psi/A)}^{\bar{z}}(\zeta(\xi)-1)\frac{s(\xi)}{\xi}d\xi\right]dG(\psi)}$.

[28] This proof of the existence and uniqueness of the equilibrium is simpler than Melitz (2003; Appendix B).

$$\int_{\underline{\psi}}^{\psi_c}\left(\left(\frac{\psi}{\psi_c}\right)^{1-\sigma}-1\right)dG(\psi)=\frac{F_e}{F}. \tag{9}$$

As L increases, the intersection moves to the left along the locus given by eq.(9), which is independent of A, as depicted by the horizontal dotted line in Figure 2. This means that the equilibrium cutoff, ψ_c, is independent of L. On the other hand, the equilibrium cutoff, ψ_c, declines in response to a lower F_e/F and to an improvement in productivity distribution, captured by a first-order stochastic dominant (FSD) shift of $\psi \sim G(\cdot)$ *to the left*. Furthermore, A can be expressed as

$$A=\psi_c\left(\frac{c_0L}{F}\right)^{\frac{1}{1-\sigma}}=\left(\frac{c_0L}{F_e}\int_{\underline{\psi}}^{\psi_c}[(\psi)^{1-\sigma}-(\psi_c)^{1-\sigma}]dG(\psi)\right)^{\frac{1}{1-\sigma}}.$$

Thus, a higher L, a lower F_e, a lower F, and a FSD shift of $\psi \sim G(\cdot)$ to the left all lead to more competitive pressures, a lower A. Since $P=cA$ for a constant $c>0$ under CES, the effect on P is the same, and the effect on TFP, $X/L=1/P$, goes the opposite direction.

The revenue, the (gross) profit and the (variable) employment of a ψ-firm are:

Revenue: $$r\left(\frac{\psi}{A}\right)L=\sigma c_0L\left(\frac{\psi}{A}\right)^{1-\sigma}=\sigma F\left(\frac{\psi}{\psi_c}\right)^{1-\sigma}\geq \sigma F$$

Profit: $$\pi\left(\frac{\psi}{A}\right)L=c_0L\left(\frac{\psi}{A}\right)^{1-\sigma}=F\left(\frac{\psi}{\psi_c}\right)^{1-\sigma}\geq F$$

Employment: $$\ell\left(\frac{\psi}{A}\right)L=(\sigma-1)c_0L\left(\frac{\psi}{A}\right)^{1-\sigma}=(\sigma-1)F\left(\frac{\psi}{\psi_c}\right)^{1-\sigma}\geq(\sigma-1)F$$

which are all decreasing power functions in ψ with the exponent, $1-\sigma<0$. Thus, their ratios across two firms with $\psi,\psi' \in (\underline{\psi},\psi_c)$, given by $(\psi/\psi')^{1-\sigma}>1$, are independent of L, F_e, F and $G(\cdot)$, as well as A and ψ_c. Hence, the relative size of two firms, whether measured in the revenue, profit, or variable employment, never changes across different equilibriums.

From the free entry condition and the adding-up constraint, $M[F_e+G(\psi_c)F]=L/\sigma$, which states that the aggregate entry cost plus the aggregate expected fixed cost is equal to the aggregate profit. Using eq.(9), this can be further rewritten to obtain:

$$M=\frac{L/\sigma}{F_e+G(\psi_c)F}=\frac{L}{\sigma F_e}\left[1-\frac{1}{H(\psi_c)}\right];MG(\psi_c)=\frac{L/\sigma}{F_e/G(\psi_c)+F}=\frac{L}{H(\psi_c)\sigma F},$$

where $H(\psi_c)\equiv\int_{\underline{\xi}}^{1}(\xi)^{1-\sigma}d\tilde{G}(\xi;\,\psi_c)$. Since $(\xi)^{1-\sigma}$ is decreasing, Lemma 2 implies

$$\mathcal{E}_G'(\cdot) \gtreqless 0 \Rightarrow H'(\psi_c) \lesseqgtr 0.$$

Hence, it is straightforward to verify the following:

Proposition 1: Under CES,

1a: A higher L keeps ψ_c unaffected and increases both M and $MG(\psi_c)$ *proportionately;*

1b: A lower F_e decreases ψ_c and increases M; It increases $MG(\psi_c)$ if $\mathcal{E}_G'(\psi) < 0$, decreases $MG(\psi_c)$ if $\mathcal{E}_G'(\psi) > 0$ and keeps $MG(\psi_c)$ unaffected if $\mathcal{E}_G'(\psi) = 0$;

1c: A lower F increases ψ_c and increases $MG(\psi_c)$; It increases M if $\mathcal{E}_G'(\psi) < 0$, decreases M if $\mathcal{E}_G'(\psi) > 0$ and keeps M unaffected if $\mathcal{E}_G'(\psi) = 0$.

Although most of these results are known, the result that the sign of $d[MG(\psi_c)]/dF_e$ and the sign of dM/dF are the same with the sign of $\mathcal{E}_G'(\psi)$ seems new.[29] A FSD shift of $G(\cdot)$ to the left reduces ψ_c. However, its effects on M and $MG(\psi_c)$ are ambiguous in general.[30]

To summarize the market size effects under CES, the markup rate is independent of market size and uniform across all active firms. Furthermore, market size has no effect on the cutoff, ψ_c, and hence on the productivity distribution as well as the revenue and employment across active firms, which are all monotonically increasing in the firm's productivity. Market size only increases the masses of entrants and of active firms proportionately. All adjustments are at the extensive margin.

4. Melitz under H.S.A.: Cross-Sectional Implications

We now depart from CES. Even though the 2nd and the 3rd laws may not be the universal laws, satisfied in every single sector in every single country, there seems to be ample evidence in their support, as cited in the introduction, so that we will primarily focus on the implications of the 2nd and the 3rd law. In this section, we explore how the impacts of more competitive pressures (a lower A) vary across heterogeneous firms, first under the 2nd law and then under the 3rd law.

[29] Appendix A shows that, $\mathcal{E}_g'(\cdot) < 0$ and $\mathcal{E}_G'(\cdot) < 0$ for Fréchet, Weibull, and Lognormal, which suggests, among others, that the results obtained by some recent studies on the Melitz model under the lognormal distribution, e.g., Head, Mayer, and Theonig (2014), are qualitatively robust to any distribution with $\mathcal{E}_g'(\cdot) < 0$.

[30] To see this, consider the case of power-distributed marginal cost (i.e., Pareto-distributed productivity), $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}, 0 < \psi < \overline{\psi}, \kappa > \sigma - 1$, so that $\mathcal{E}_G'(\cdot) = 0$ and $\tilde{G}(\xi;\psi_c) = \xi^{\kappa}$, and $H(\psi_c) = \int_0^1 \kappa(\xi)^{\kappa-\sigma} d\xi = \frac{\kappa}{\kappa-\sigma+1} > 1$ is independent of ψ_c. Under the condition that ensures the interior solution, $G(\psi_c) = \frac{\kappa-\sigma+1}{\sigma-1}\left(\frac{F_e}{F}\right) < 1$, we have $M = \frac{\sigma-1}{\kappa}\left(\frac{L}{\sigma F_e}\right) > MG(\psi_c) = \frac{\kappa-\sigma+1}{\kappa}\left(\frac{L}{\sigma F}\right)$. Thus, a FSD shift in G, due to a change in $\overline{\psi}$, has no effect on $G(\psi_c), M$ nor $MG(\psi_c)$, while a FSD shift in G, due to a change in κ, affects $G(\psi_c), M$ and $MG(\psi_c)$.

Of course, A is an endogenous variable, whose change must be triggered by a change in some exogenous variables in general equilibrium. Nevertheless, we postpone the general equilibrium comparative statics analysis to the next section.

4.1. Cross-Sectional Implications of the 2nd Law of Demand

A2: $\zeta'(z) > 0$ for all $z \in (0, \bar{z}) \Leftrightarrow \sigma'(\psi/A) = \zeta'\big(Z(\psi/A)\big)Z'(\psi/A) > 0$ for all $\psi/A \in (0, \bar{z})$

Under **A2,** $\mathcal{E}_{\zeta}(z) > 0 > \mathcal{E}_{s}(z)$ for all $z \in (0, \bar{z})$. Hence, **A1** is ensured under **A2.** This assumption means that the price elasticity of demand, $\zeta\big(p_{\psi}/A\big)$, is strictly increasing in its price, p_{ψ} for a fixed A, which each firm takes as given. It is thus equivalent to Marshall's 2nd Law of demand. Under **A2**, $\zeta\big(Z(\psi/A)\big) = \sigma(\psi/A)$ is a strictly increasing function of $\psi/A.$ It means that $\mathcal{E}_{\zeta/(\zeta-1)}(z) = \mathcal{E}_{\zeta}(z)/\mathcal{E}_{s}(z) < 0$, hence $\mathcal{E}_{\mu}(\psi/A) = \mathcal{E}_{[\zeta/(\zeta-1)]\circ Z}(\psi/A) =$ $\mathcal{E}_{\zeta/(\zeta-1)}\big(Z(\psi/A)\big)\rho(\psi/A) < 0$ and

$$\rho\left(\frac{\psi}{A}\right) = \mathcal{E}_{Z}\left(\frac{\psi}{A}\right) = 1 + \mathcal{E}_{\mu}\left(\frac{\psi}{A}\right) < 1,$$

so that less productive firms have lower markup rates and that the price responds less than proportionately to a change in the marginal cost (**Incomplete Pass-Through).** Furthermore,

$$\frac{\partial \ln p_{\psi}}{\partial \ln A} = \frac{\partial \ln(Z(\psi/A)A)}{\partial \ln A} = 1 - \frac{d \ln\big(Z(\psi/A)\big)}{d \ln(\psi/A)} = 1 - \mathcal{E}_{Z}\left(\frac{\psi}{A}\right) = 1 - \rho\left(\frac{\psi}{A}\right) > 0.$$

Thus, the firm reduces its price (and its markup rate) in response to more competitive pressures, a lower A, which occurs either when other firms reduce their prices (**Strategic complementarity in pricing**) or when more firms enter **(Procompetitive entry**).[31]

For further exploration, let us reformulate the definitions of log-super(sub)modularity specifically for our context. A positive-valued C^2-function f of a single variable, $\psi/A > 0$, $f(\psi/A)$, when viewed as a function of the two variables, ψ and A, is strictly *log-super(sub)modular* in ψ and A if $\partial^2 \ln f(\psi/A)/\partial\psi\partial A > (<)0$. Or, we sometimes say, more simply, that $f(\psi/A)$ is strictly *log-super(sub)modular,* when this condition holds.

[31] As pointed out in Matsuyama and Ushchev (2020a), the 2nd law of demand (or incomplete pass-through) is in general neither sufficient nor necessary for procompetitive entry (or strategic complementarity in price), since the former is about the property of the individual demand curve, while the latter is about the property of the entire demand system. They are equivalent under H.S.A., since the single aggregator A, which captures all the interaction across firms, enters the price elasticity function only as ψ/A, so that a change in A is isomorphic to a change in ψ, acting as a magnifier of firm heterogeneity

The log-super(sub)modularity of a decreasing function $f(\psi/A)$ means that more competitive pressures, a lower A, causes a disproportionately larger (smaller) decline for a higher ψ. The next lemma offers a simple way of verifying the log-super(sub)modularity of $f(\psi/A)$.

Lemma 5: For any positive-valued C^2-function f of a single variable, $\psi/A > 0$,

$$sgn\left\{\frac{\partial^2 \ln f(\psi/A)}{\partial\psi\partial A}\right\} = -sgn\left\{\mathcal{E}_f'\left(\frac{\psi}{A}\right)\right\} = -sgn\left\{\frac{d^2 \ln f\left(e^{\ln(\psi/A)}\right)}{(d\ln(\psi/A))^2}\right\}.$$

The proof of this lemma is straightforward and hence omitted. This lemma states that $f(\psi/A)$ is strictly *log-super(sub)modular* in ψ and A if and only if $\mathcal{E}_f(\cdot)$ is strictly decreasing(increasing), that is, if and only if $\ln f(e^x) = \ln f(\psi/A)$ is strictly concave (convex) in $x \equiv \ln(\psi/A)$. Since $\mathcal{E}_\pi(\psi/A) = 1 - \sigma(\psi/A) < 0$ is strictly decreasing in ψ/A under **A2**, Lemma 5 immediately tells us that the profit, $\pi(\psi/A)L$, is strictly log-supermodular in ψ and A.

To summarize these implications of **A2,**

Proposition 2 (Cross-Sectional Implications of 2nd Law): Under **A2,**

2a **(Incomplete pass-through)**:

$$\mathcal{E}_\mu\left(\frac{\psi}{A}\right) < 0 \Leftrightarrow 0 < \rho\left(\frac{\psi}{A}\right) = 1 + \mathcal{E}_\mu\left(\frac{\psi}{A}\right) = 1 - \mathcal{E}_{1/\mu}\left(\frac{\psi}{A}\right) < 1.$$

2b **(Procompetitive effect/strategic complementarity):**

$$\frac{\partial \ln p_\psi}{\partial \ln A} = 1 - \rho\left(\frac{\psi}{A}\right) = -\mathcal{E}_\mu\left(\frac{\psi}{A}\right) = \mathcal{E}_{1/\mu}\left(\frac{\psi}{A}\right) > 0.$$

2c **(Strictly log-supermodular profit)**:

$$\mathcal{E}_\pi'\left(\frac{\psi}{A}\right) = -\sigma'\left(\frac{\psi}{A}\right) < 0 \Leftrightarrow \frac{\partial^2 \ln \pi(\psi/A)L}{\partial\psi\partial A} > 0.$$

Because $\pi(\psi/A)$ is strictly log-supermodular, more competitive pressures, a lower A, causes a proportionately larger decline in the profit among higher-ψ firms. Because higher-ψ firms have lower profits, this implies that more competitive pressures lead to a larger dispersion of profits across firms with the profit density shifting toward lower-ψ firms. Figure 3a illustrates this by plotting the graphs of log-profit, $\ln \Pi_\psi$, as a function of log-marginal cost, $\ln\psi$. The graph is always downward-sloping, and it is strictly concave under A2. The effect of a lower A, for a fixed L, is captured by a parallel leftward shift of the graph, which means a larger downward shift for high-ψ due to the concavity. Thus, higher-ψ firms experience proportionately larger decline in the profit.[32]

4.2. Cross-Sectional Implications of the 3rd Law of Demand

A2 alone does not ensure log-supermodularity nor log-submodularity of $Z(\psi/A), r(\psi/A)L$ or $\ell(\psi/A)L$, since the monotonicity of $\sigma(\cdot)$ alone does not imply the monotonicity of $\mathcal{E}_Z(\cdot) = \rho(\cdot)$; $\mathcal{E}_r(\cdot) = [1-\sigma(\cdot)]\rho(\cdot)$; and $\mathcal{E}_\ell(\cdot) = 1-\rho(\cdot)\sigma(\cdot)$. Partially motivated by this, we now consider the following assumption.

A3: For all $z \in (0, \bar{z})$,

$$\mathcal{E}'_{\zeta/(\zeta-1)}(z) = -\frac{d}{dz}\left(\frac{z\zeta'(z)}{[\zeta(z)-1]\zeta(z)}\right) \geq 0 \iff \rho'\left(\frac{\psi}{A}\right) \geq 0$$

A3 means that the pass-through rate is weakly increasing in ψ, which we shall call t**he 3rd Law of demand.** In particular, we call it the weak 3rd Law of demand or simply **the weak A3** when the inequality in **A3** holds weakly, and the strong 3rd Law of demand or simply **the strong A3**, when the inequality in **A3** holds strictly and hence the pass-through rate is strictly increasing in ψ. Among the three parametric families of H.S.A. discussed in Appendix D, Generalized Translog satisfies A2 but violates even the weak A3; Constant Pass-Through (CoPaTh) satisfies A2 and the weak A3, but violates the strong A3; and Power Elasticity of Markup Rates (PEM) satisfies both A2 and the strong A3.

Then, using Lemma 5, we have the following proposition:

Proposition 3 (Cross-Sectional Implications of 3rd Law):

3a (Weak (strict) log-submodular price and markup rate): Under **the weak (strong) A3,**

$$\mathcal{E}'_Z\left(\frac{\psi}{A}\right) = \rho'\left(\frac{\psi}{A}\right) \geq (>) < 0 \iff \frac{\partial^2 \ln(Z(\psi/A)A)}{\partial\psi\partial A} = \frac{\partial^2 \ln \mu(\psi/A)}{\partial\psi\partial A} \leq (<) 0,$$

3b (Strict log-supermodular revenue): Under **A2** and **the weak A3,**

[32] Figure 3a also depicts the effect of a higher L for a fixed A as a parallel upward shift of the graph. In Proposition 6, it will be shown that a higher L always leads to a lower A. Thus, if A declines due to a higher L, the full impact of a higher L on the profit is captured by a combination of the parallel upward shift (the positive direct effect) and the parallel leftward shift (the indirect effect due to a lower A). Notice that the positive direct effect is uniform across firms, while the negative indirect effect is disproportionately smaller for low-ψ firms under A2. In Proposition 7a, it will be shown that the combined effect leads to a clockwise rotation of the graph, as depicted in Figure 3a, around the pivot point, which is located strictly below the cutoff ψ_c. This means that a higher L causes the profits to go up among low-ψ firms and to go down among high-ψ firms, generating what Mrázová-Neary (2017; 2019) dubbed as The Matthew Effect, "to those who have, more shall be given."

$$\mathcal{E}_r'\left(\frac{\psi}{A}\right) = \left[1-\sigma\left(\frac{\psi}{A}\right)\right]\rho'\left(\frac{\psi}{A}\right) - \sigma'\left(\frac{\psi}{A}\right)\rho\left(\frac{\psi}{A}\right) < 0 \Leftrightarrow \frac{\partial^2 \ln r(\psi/A)}{\partial\psi\partial A} > 0$$

3c (Strict log-supermodular employment): Under **A2** and **the weak A3,**

$$\mathcal{E}_\ell'\left(\frac{\psi}{A}\right) = -\sigma'\left(\frac{\psi}{A}\right)\rho\left(\frac{\psi}{A}\right) - \sigma\left(\frac{\psi}{A}\right)\rho'\left(\frac{\psi}{A}\right) < 0 \Leftrightarrow \frac{\partial^2 \ln \ell(\psi/A)}{\partial\psi\partial A} > 0.$$

Proposition 3a states that the price, $p_\psi = Z(\psi/A)A$, the markup rate, $\mu_\psi = \mu(\psi/A)$, and the relative price, $Z(\psi/A)$, are all weakly (strictly) log-submodular in ψ and A under the weak (strong) A3. More competitive pressures thus cause a markup rate decline, proportionately no larger (strictly smaller) among higher-ψ firms. Since their markup rates are lower under A2, this also implies no larger (strictly smaller) dispersion of the markup rate across firms. Figure 3b illustrates this by plotting the graphs of log-markup rate, $\ln \mu_\psi$, as a function of log-marginal cost, $\ln \psi$. The graph is downward-sloping under A2, and it is strictly convex under strong A3. The effect of a decline in A is captured by a parallel leftward shift of the graph, which means a larger downward shift for low-ψ due to the convexity. Thus, lower-ψ firms experience proportionately larger decline in the markup rate.

Proposition 3b states that the revenue, $r(\psi/A)L$, is strictly log-supermodular in ψ and A under A2 and the weak A3. This means that a lower A, causes a proportionately larger decline in the revenue among higher-ψ firms. Since their revenues are lower, this also implies that more competitive pressures lead to a larger dispersion of revenues across firms with the profit density shifting toward lower-ψ firms. Thus, $R_\psi = r(\psi/A)L$ under A2 and the weak A3 share the same properties with $\Pi_\psi = \pi(\psi/A)L$ under A2, as depicted in Figure 3a.[33] This theoretical finding, a shift of the revenue density from the less productive/smaller firms with lower markup rates to the more productive/larger firms with higher markup rates, echoes the calibration findings by Baqaee, Fahri, and Sangani (2021) and Edmond, Midrigan, and Xu (2021).

Proposition 3c states that the employment, $\ell(\psi/A)L$, is also strictly log-supermodular in ψ and A under A2 and the weak A3. However, its strict log-supermodularity has different implications from that of the profit $\pi(\psi/A)L$ and the revenue $r(\psi/A)$. This is because the

[33] If A declines due to a higher L, the full impact of a higher L on the revenue is captured by a combination of the parallel upward shift (the direct effect) and the parallel leftward shift (the indirect effect of a lower A). Again, the positive direct effect is uniform across firms, while the negative indirect effect is disproportionately smaller for low-ψ firms under A2 and the weak A3. In Proposition 7b, it will be shown that the combined effect leads to a clockwise rotation of the graph, as depicted in Figure 3a, generating the Matthew effect in revenue. Unlike the case of the profit under A2, however, the pivot point for the revenue under A2 and the weak A3 may be above the cutoff ψ_c. If so, all firms below the cutoff would experience an increase in their revenue.

employment $\ell(\psi/A)L$ is hump-shaped in ψ/A under A2 and the weak A3. To see this, we first prove in Appendix C.1:

Lemma 6: Under A2 and the weak A3, $\lim_{\psi/A\to 0}\rho(\psi/A)\sigma(\psi/A) < 1 < \lim_{\psi/A\to\bar{z}}\rho(\psi/A)\sigma(\psi/A)$.

Since $\mathcal{E}_\ell(\psi/A) = 1-\rho(\psi/A)\sigma(\psi/A)$ is globally decreasing, Lemma 6 implies that there exists a unique $\hat{\psi} > 0$, such that $\mathcal{E}_\ell(\psi/A) > 0$ for $\psi < \hat{\psi}$ and $\mathcal{E}_\ell(\psi/A) < 0$ for $\psi > \hat{\psi}$. Thus,

Proposition 4: Under **A2** and **the weak A3,** the employment function, $\ell(\psi/A) = r(\psi/A)/\mu(\psi/A)$ is hump-shaped, with its unique peak is reached at, $\hat{z} \equiv Z(\hat{\psi}/A) < \bar{z}$, where

$$\mathcal{E}_{s(\zeta-1)/\zeta}(\hat{z}) = 0 \Leftrightarrow \frac{\hat{z}\zeta'(\hat{z})}{\zeta(\hat{z})} = [\zeta(\hat{z})-1]^2 \Leftrightarrow \mathcal{E}_\ell\left(\frac{\hat{\psi}}{A}\right) = 0 \Leftrightarrow \rho\left(\frac{\hat{\psi}}{A}\right)\sigma\left(\frac{\hat{\psi}}{A}\right) = 1.$$

Figure 3c illustrates Propositions 3c and 4 by plotting the log-employment as a function of the log-marginal cost, which is not only strictly concave (Proposition 3c) but also hump-shaped (Proposition 4). Thus, there are three generic equilibrium configurations; all firms are below the peak if $\psi_c < \hat{\psi}$, firms are on both sides of the peak if $\underline{\psi} < \hat{\psi} < \psi_c$, or all firms are above the peak if $\hat{\psi} < \underline{\psi}$. The following corollary shows the underlying condition for each of these three cases, whose derivation is straightforward and hence omitted.

Corollary of Proposition 4: Employments across active firms are

- increasing in ψ if $\psi_c < \hat{\psi} \Leftrightarrow F/L = \pi(\psi_c/A) > \pi(\hat{\psi}/A) = \pi(Z^{-1}(\hat{z}))$;
- hump-shaped in ψ if $\underline{\psi} < \hat{\psi} < \psi_c \Leftrightarrow F/L < \pi(\hat{\psi}/A) = \pi(Z^{-1}(\hat{z}))$ & $A > \underline{\psi}/Z^{-1}(\hat{z})$.
- decreasing in ψ, if $\hat{\psi} < \underline{\psi} \Leftrightarrow A < \underline{\psi}/Z^{-1}(\hat{z})$, which is possible only if $\underline{\psi} > 0$.

In the first case, the employments are inversely related to productivity across all active firms. This occurs if $F/L > \pi(Z^{-1}(\hat{z}))$, i.e., when the overhead is high enough relative to market size. In the second case, the employments are inversely related among the relatively productive firms. In the third case, the employments are positively related to firm productivity. This can occur only if $\underline{\psi} > 0$.

Figure 3c also depicts the effect of a decline in A by a parallel leftward shift of the graph, and that of a higher L by a parallel upward shift of the graph. Due to its hump-shape, a decline in A alone causes a crossing of the graphs before and after the change. Thus, the employments of low-ψ firms go up due to more competitive pressures, a lower A, even if market size is unchanged.[34] This never happens for the profit and revenue; a lower A, always reduces the profit and revenue for all firms, unless it is caused by an increase in market size.

For the pass-through rate function, we prove in Appendix C.2.,

Proposition 5: Suppose that A2 and the strong A3 hold, so that $0 < \rho(\psi/A) < 1$ and $\rho(\psi/A)$ is strictly increasing. Then, $\rho(\psi/A)$ is strictly log-submodular for all $\psi/A < \overline{z}$ with a sufficiently small $\overline{z}$.

Figure 3d illustrates Proposition 5. It states that, under the strong A3, a lower A (more competitive pressures) causes a proportionately smaller increase in the pass-through rate for lower-ψ firms for a sufficiently small $\overline{z} > 0$.

5. Melitz under H.S.A.: General Equilibrium Comparative Statics

In Section 4, we studied how a change in competitive pressures, A, an endogenous variable, has differential effects on heterogeneous firms without specifying underlying exogenous shocks that cause it. We now study the general equilibrium effects of exogenous shocks to the entry cost F_e, the overhead F, and market size L. The recursive structure of the model allows us to proceed in two steps. First, we study the effects on competitive pressures, A and the cutoff, ψ_c, in section 5.1. and explore some of the implications in sections 5.2 and 5.3. Then, we study the effects on M and $MG(\psi_c)$ in section 5.4.

5.1. General Equilibrium Effects of F_e, F, and L on ψ_c, ψ_c/A and A

Recall that A and ψ_c are uniquely determined by the cutoff rule, eq.(5), and the free-entry condition, eq.(6). Hence, the general equilibrium effects of a change in F_e, F, and L on ψ_c, ψ_c/A and A can be obtained by totally differentiating eq.(5) and eq.(6). To simplify the expressions and to facilitate the interpretation, it is useful to introduce the following three variables. First,

$$\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right) \equiv \frac{\int_{\underline{\psi}}^{\psi_c} \sigma(\psi/A)\pi(\psi/A)\,dG(\psi)/G(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} \pi(\psi/A)\,dG(\psi)/G(\psi_c)} = 1 + \frac{\int_{\underline{\psi}}^{\psi_c} \ell(\psi/A)\,dG(\psi)/G(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} \pi(\psi/A)\,dG(\psi)/G(\psi_c)} > 1,$$

[34] This occurs whenever $\ell(\psi/A)$ is hump-shaped, for which A2 and the weak A3 is a sufficient but not a necessary condition. Generalized Tranlog in Appendix D.1. offers such an example for $\eta < 1$.

the profit-weighted average of the price elasticity across the active firms. Notice that $\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1 > 0$ is equal to the average employment relative to the average profit across the active firms. Second,

$$f_x \equiv \frac{FG(\psi_c)}{F_e + FG(\psi_c)} = \frac{FG(\psi_c)}{L\int_{\underline{\psi}}^{\psi_c} \pi(\psi/A) dG(\psi)} = \frac{\pi(\psi_c/A)}{\int_{\underline{\psi}}^{\psi_c} \pi(\psi/A)\, dG(\psi)/G(\psi_c)} < 1,$$

the share of the expected overhead cost in the total expected fixed cost. In equilibrium, it is equal to the profit of the cut-off firm relative to the average profit across the active firms. Third,

$$\delta \equiv \frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}{\sigma(\psi_c/A) - 1} = \frac{\pi(\psi_c/A)}{\ell(\psi_c/A)} \frac{\int_{\underline{\psi}}^{\psi_c} \ell(\psi/A)\, dG(\psi)/G(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} \pi(\psi/A)\, dG(\psi)/G(\psi_c)} = f_x \frac{\int_{\underline{\psi}}^{\psi_c} \ell(\psi/A)\, dG(\psi)/G(\psi_c)}{\ell(\psi_c/A)},$$

the profit/employment ratio of the cut-off firm relative to the average profit/the average employment ratio across the active firms. Then,

Proposition 6:

$$\left[\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1\right] \frac{dA}{A} = (1 - f_x)\left(\frac{dF_e}{F_e}\right) - \frac{dL}{L} + f_x\left(\frac{dF}{F}\right);$$

$$\left[\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1\right]\left(\frac{d\psi_c}{\psi_c} - \frac{dA}{A}\right) = \delta\left(\frac{dL}{L} - \frac{dF}{F}\right);$$

$$\left[\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1\right] \frac{d\psi_c}{\psi_c} = (1 - f_x)\left(\frac{dF_e}{F_e}\right) - (1 - \delta)\left(\frac{dL}{L}\right) + (f_x - \delta)\left(\frac{dF}{F}\right).$$

The derivation is straightforward and hence omitted. To summarize the qualitative impacts

Corollary of Proposition 6:

6a **(Entry Cost):** $\frac{dA}{dF_e} > 0, \frac{d(\psi_c/A)}{dF_e} = 0$, and $\frac{d\psi_c}{dF_e} > 0$.

6b **(Market Size)**: $\frac{dA}{dL} < 0, \frac{d(\psi_c/A)}{dL} > 0$, and $\frac{d\psi_c}{dL} \gtreqless 0 \Leftrightarrow \delta \gtreqless 1 \Leftrightarrow \frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right)}{\sigma(\psi_c/A)} \gtreqless 1$

In particular, $\frac{d\psi_c}{dL} < 0$ holds globally if $\sigma'(\cdot) > 0$, i.e., under A2.

6c **(Overhead Cost):** $\frac{dA}{dF} > 0, \frac{d(\psi_c/A)}{dF} < 0$, and $\frac{d\psi_c}{dF} \gtreqless 0 \Leftrightarrow \frac{f_x}{\delta} \gtreqless 1 \Leftrightarrow \frac{\ell(\psi_c/A)}{\int_{\underline{\psi}}^{\psi_c} \ell(\psi/A) dG(\psi)/G(\psi_c)} \gtreqless$

1. In particular, $\frac{d\psi_c}{dF} > 0$ holds globally if $\ell'(\cdot) > 0$.

Figures 4a-4c illustrate Corollary of Proposition 6.

Figure 4a shows the effects of **a decline in $\boldsymbol{F_e}$.** A smaller entry cost makes the entry more attractive, while keeping an incentive to stay in the market after the entry unaffected. Thus, it shifts the free entry condition down and to the left, while keeping the cutoff rule unchanged. Hence, it leads to a decline in both ψ_c and A at the same rate, resulting in more competitive pressures and a tougher selection.

Figure 4b shows the effects of **an increase in *L*.** A larger market size has two different effects. On one hand, it makes the entry more attractive, thus shifting the free entry condition down and to the left. On the other hand, it gives more incentive to stay in the market after the entry at each level of competitive pressures, thus rotating the cutoff rule counter-clockwise. The intersection thus unambiguously moves to the left, causing a smaller A. To determine the impact on ψ_c, which depends on the relative magnitudes of the two effects, eliminate L from eq.(5) and eq.(6) to obtain:

$$\int_{\underline{\psi}}^{\psi_c} \left[\frac{\pi(\psi/A)}{\pi(\psi_c/A)} - 1\right] dG(\psi) = \frac{F_e}{F}.$$

As L changes, the intersection moves along the locus defined by this equation. Its LHS is globally strictly increasing in ψ_c. It is also strictly decreasing in A, wherever $\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) < \sigma(\psi_c/A)$ holds:[35] that is, whenever the profit-weighted average price elasticity across the active firms is lower than the price elasticity at the cutoff firm. This condition holds globally, if $\sigma(\cdot)$ is strictly increasing, i.e., **A2,** in which case the locus is globally upward-sloping, as depicted by the dotted line in Figure 4b. Thus, under **A2**, a higher L always causes a decline in both ψ_c and A, with ψ_c/A going up.[36]

Figure 4c shows the effects of **a decline in *F***: Similar to a higher L, a smaller overhead cost has two different effects. It not only makes the entry more attractive, thus shifting the free entry condition down and to the left, but also gives more incentive to stay in the market after the entry, thus rotating the cutoff rule ray counter-clockwise. The intersection thus unambiguously moves to the left, causing a decline in A. To determine the impact on ψ_c, eliminating F from eq.(5) and eq.(6) yields:

$$\int_{\underline{\psi}}^{\psi_c} \left[\pi\left(\frac{\psi}{A}\right) - \pi\left(\frac{\psi_c}{A}\right)\right] dG(\psi) = \frac{F_e}{L}.$$

[35] This can be verified by differentiating the LHS with respect to A and making use of $\mathcal{E}_\pi(\psi/A) = 1 - \sigma(\psi/A)$.

[36] Since A2 implies the log-supermodularity of $\pi(\psi/A)$, as shown in Proposition 2, $\pi(\psi/A)/\pi(\psi_c/A)$ is strictly decreasing in A for $\psi < \psi_c$, and so is the integrand of the LHS. Under the opposite of A2, $\sigma'(\cdot) < 0$, the locus would be negatively-sloped and a higher L would lead to an increase in ψ_c. CES is the borderline case, with the horizontal locus, hence a change in L has no effect on ψ_c.

As F changes, the intersection moves along the locus defined by this equation. Its LHS is globally strictly increasing in ψ_c. It is also strictly decreasing in A, wherever $f_x > \delta$, or equivalently $\ell(\psi_c/A) > \int_{\underline{\psi}}^{\psi_c} \ell(\psi/A)\, dG(\psi)/G(\psi_c)$ holds.[37] that is, whenever the average employment across the active firms is lower than the employment by the cutoff firm. This condition holds globally if $\ell(\cdot)$ is strictly increasing. As shown in Corollary of Proposition 4, this occurs under A2 and the weak A3 when the overhead cost is sufficiently large relative to market size. In this case, the locus is globally upward-sloping, as depicted by the dotted curve in Figure 4c. Hence a lower F always causes a decline in both ψ_c and A, with ψ_c/A going up.

5.2. Market Size Effect on Profit, $\boldsymbol{\Pi_\psi \equiv \pi(\psi/A)L}$ and Revenue, $\boldsymbol{R_\psi \equiv r(\psi/A)L}$

As we suggested in section 4, the full impacts of a higher L on the profit (under A2) and of the revenue (under A2 and the weak A3) are captured by a combination of the parallel upward shift (the direct effect) and the parallel leftward shift (the indirect effect due to a lower A) of the graph in Figure 3a. Because the positive direct effect is uniform across firms, while the negative indirect effect is smaller for low-ψ firms, the combined effect could result in a clockwise rotation of the graph, such that a higher L, accompanied by a lower A, leads to an increase in the profit and the revenue among low-ψ firms. We are now ready to state this result formally in Propositions 7a and 7b, whose proof is in Appendix C.3.

Proposition 7a: Under **A2,** there exists a unique $\psi_0 \in (\underline{\psi}, \psi_c)$ such that $\sigma\left(\frac{\psi_0}{A}\right) = \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right)$ with

$$\frac{d \ln \Pi_\psi}{d \ln L} > 0 \Leftrightarrow \sigma\left(\frac{\psi}{A}\right) < \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) \text{ for } \psi \in \left(\underline{\psi}, \psi_0\right),$$

and

$$\frac{d \ln \Pi_\psi}{d \ln L} < 0 \Leftrightarrow \sigma\left(\frac{\psi}{A}\right) > \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) \text{ for } \psi \in (\psi_0, \psi_c).$$

Proposition 7b: Under **A2** and the weak **A3,** there exists $\psi_1 > \psi_0$, such that

$$\frac{d \ln R_\psi}{d \ln L} > 0 \text{ for } \psi \in \left(\underline{\psi}, \psi_1\right).$$

Furthermore, $\psi_1 \in (\psi_0, \psi_c)$ and

$$\frac{d \ln R_\psi}{d \ln L} < 0 \text{ for } \psi \in (\psi_1, \psi_c),$$

[37] This can be verified by differentiating the LHS with respect to A and making use of $(\psi/A)\pi'(\psi/A) = \pi(\psi/A)\mathcal{E}_\pi(\psi/A) = \pi(\psi/A)[1 - \sigma(\psi/A)] = \pi(\psi/A) - r(\psi/A) = -\ell(\psi/A)$

for a sufficiently small F.

Figures 5a-5c graphically put together the main implications of Propositions 2, 3, 6, and 7 under A2 and the weak A3 for the effects on the log-markup rates, the log-profits, and log the revenues, of more competitive pressures (a lower A) and a tougher selection (a lower ψ_c), when they are caused by a decline in F_e, an increase in L and a decline in F (with $\ell'(\cdot) > 0$). In all three cases, log-profit is decreasing, and concave in log-marginal cost due to A2 (Proposition 2) and log-markup rate (log-revenue) is decreasing, and convex (concave) in log-marginal cost due to A2 and the weak A3 (Proposition 3).

5.3. Average Markup and Pass-Through Rates: The Composition Effect

In all three cases illustrated in Figures 4a-4c and Figures 5a-5c, the shocks that cause a decline in A, more competitive pressures, also cause a decline in ψ_c, a tougher selection. This creates non-trivial composition effects.

- Under A2, more competitive pressures (a decline in A) cause all firms that continue to operate in the market to lower their markup rates $\mu(\psi/A)$. However, a decline in ψ_c means that high-ψ firms with lower $\mu(\psi/A)$ drop out.
- Under strong A3, more competitive pressures (a decline in A) cause all firms that continue to operate in the market to increase their pass-through rates $\rho(\psi/A)$. However, a decline in ψ_c means that high-ψ firms with higher $\rho(\psi/A)$ drop out.

Due to this composition effect, the average markup and pass-through rates can go in either direction. Propositions 8a and 8b identify some conditions that determine the direction of the changes in the average values in the case where the shock keeps ψ_c/A unchanged.[38] The proof is in Appendix C.4.

Proposition 8a. *Assume $\underline{\psi} = 0$. Consider a shock, which causes a proportional decline in A and ψ_c, so that ψ_c/A remains constant. Then, for any weighting function $w(\psi/A)$,*

i) the weighted average of any monotonically decreasing (increasing) $f(\psi/A)$,

$$\frac{\int_{\underline{\psi}}^{\psi_c} f(\psi/A)w(\psi/A)dG(\psi)}{\int_{\underline{\psi}}^{\psi_c} w(\psi/A)dG(\psi)},$$

[38]Examples include a decline in F_e, as well as a decline in L for the limit case of $F \to 0$ with $\bar{z} < \infty$, discussed in section 5.5. Moreover, part i) of Proposition 8a and Proposition 8b offer some conditions for the direction of the change for any combination of the shocks, as long as its effect on ψ_c/A is small enough.

decreases (increases) if $\mathcal{E}_g'(\cdot) < 0$ *and increases (decreases) if* $\mathcal{E}_g'(\cdot) > 0$.

ii) *the weighted average of any* $f(\psi/A)$*, monotonic or not, remains constant, if* $\mathcal{E}_g'(\cdot) = 0$.

Note that this result holds for any weight function used to calculate the average, as long as it is a function of ψ/A, such as the revenue $r(\psi/A)$, the profit $\pi(\psi/A)$, or the employment $\ell(\psi/A)$, or by the marginal cost, ψ, for that matter.[39] Part ii) of Proposition 8a also implies that a decline in F_e under Pareto-productivity, $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}$ offers a knife-edge case, where the average markup and pass-through rates remain unchanged.

Note that the conditions in Proposition 8a are stated in terms of the sign of $\mathcal{E}_g'(\cdot)$. For the weaker conditions using the signs of $\mathcal{E}_G'(\cdot)$, we have the following result on the average markup rate weighted by the monotonically increasing employment, $\ell(\psi/A)$.

Proposition 8b. *Assume that A2 holds,* $\underline{\psi} = 0$*, and* $\ell(\psi/A)$ *is increasing in* ψ/A *for all* $\psi/A \in (0, \psi_c/A)$*. Consider a shock, which causes a proportional decline in A and* ψ_c *so that* ψ_c/A *remains constant. Then, the* $\ell(\cdot)$*-weighted average markup rate decreases if* $\mathcal{E}_G'(\cdot) < 0$*; remains constant if* $\mathcal{E}_G'(\cdot) = 0$*; and increases if* $\mathcal{E}_G'(\cdot) > 0$.

These results suggest that a decline in the entry cost, which leads to more competitive pressures and more concentration of high productivity firms, might lead to the rise, not the fall, of the average markup rate due to the composition effect, and hence an increase in the average markup rate should not be interpreted as the *prima-facie* evidence for reduced competitive pressures.[40]

5.4. Comparative Statics on $M, MG(\psi_c), M/L,\ MG(\psi_c)/L$

The next proposition looks at the effects on the mass of entrants as well as the mass of active firms. The proof is in Appendix C.5.

Proposition 9a (The Effects of F_e on M and $MG(\psi_c)$)

$$\frac{dM}{dF_e} < 0; \qquad \mathcal{E}_G'(\psi) \gtreqless 0, \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \frac{d[MG(\psi_c)]}{dF_e} \gtreqless 0$$

Proposition 9b (The Effects of L on M and $MG(\psi_c)$): *Under* **A2***,*

[39] Qualitatively at least. Which weights are used may matter quantitatively, as argued by Edmond, Midrigan, and Xu (2021).

[40] Indeed, Autor et.al. (2020) and De Loecker, Eeckhout, and Unger (2020) pointed out that much of the recent rise in the average markup is due to reallocation from the low markup firms to the high markup firms.

$\frac{dM}{dL} > 0;$ $\qquad$ $\mathcal{E}'_G(\psi) \le 0, \forall\psi \in (\underline{\psi}, \overline{\psi}) \Rightarrow \frac{d[MG(\psi_c)]}{dL} > 0.$
Proposition 9c (The Effects of L on M/L and $MG(\psi_c)/L$): *Under* **A2**, $G(\psi) = (\psi/\overline{\psi})^{\kappa} \Rightarrow \frac{d}{dL}\left(\frac{M}{L}\right) > 0;$ $\qquad$ $\mathcal{E}'_G(\psi) \ge 0, \forall\psi \in (\underline{\psi}, \overline{\psi}) \Rightarrow \frac{d}{dL}\left(\frac{MG(\psi_c)}{L}\right) < 0.$
Proposition 9d (The Effects of F on M and $MG(\psi_c)$): *If* $\ell'(\cdot) > 0$, $\frac{dM}{dF} < 0;$ $\qquad$ $\mathcal{E}'_G(\psi) \le 0, \forall\psi \in (\underline{\psi}, \overline{\psi}) \Rightarrow \frac{d[MG(\psi_c)]}{dF} < 0.$

Proposition 9 suggests that the results on the impacts of the masses of firms usually depend on whether $\mathcal{E}_G(\psi)$ is increasing or decreasing, with $\mathcal{E}'_G(\psi) = 0$, power-distributed marginal costs, i.e., Pareto-distributed productivity, being often the knife-edge case.

5.5. The Limit Case of $F \to 0$ with $\bar{z} < \infty$.

Before proceeding to a multi-market extension, we briefly look at a limit case, $F \to 0$, with $\bar{z} < \infty$. In this limit case, there is no overhead cost and the cutoff firms supply with zero markup, i.e., at that marginal cost.[41] The equilibrium can be described by eq.(5) and eq.(6), which now become simply:

Cutoff Rule:
$$\pi\left(\frac{\psi_c}{A}\right) = 0 \Leftrightarrow \frac{\psi_c}{A} = Z\left(\frac{\psi_c}{A}\right) = \bar{z} = \pi^{-1}(0)$$

Free Entry Condition:
$$\frac{F_e}{L} = \int_{\underline{\psi}}^{\psi_c} \pi\left(\bar{z}\frac{\psi}{\psi_c}\right) dG(\psi) = \int_{\underline{\psi}}^{\bar{z}A} \pi\left(\frac{\psi}{A}\right) dG(\psi).$$

Notice that the cutoff rule alone determines $\psi_c/A = \bar{z}$; it is independent of F_e and L. And the free-entry condition uniquely determines $\psi_c = \bar{z}A$ as C^2 functions of F_e/L with the interior solution, $0 < G(\psi_c) < 1$, guaranteed for

$$0 < \frac{F_e}{L} < \int_{\underline{\psi}}^{\overline{\psi}} \pi\left(\bar{z}\frac{\psi}{\overline{\psi}}\right) dG(\psi).$$

Simple algebra can verify that

[41] Although one of the advantages of the Melitz model under H.S.A. is that it is tractable with $F > 0$, we look at this case in order to facilitate a comparison with some existing studies, such as Melitz and Ottaviano (2008) and Arkolakis et.al. (2019), which assume $F = 0$ in the presence of the choke price.

$$\frac{d\psi_c}{\psi_c} = \frac{dA}{A} = \frac{1}{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}\left(\frac{dF_e}{F_e} - \frac{dL}{L}\right) = \frac{\int_{\underline{\psi}}^{\psi_c} \pi(\psi/A) dG(\psi)}{\int_{\underline{\psi}}^{\psi_c} \ell(\psi/A) dG(\psi)}\left(\frac{dF_e}{F_e} - \frac{dL}{L}\right),$$

which can be also obtained from Proposition 6 by setting $f_x = \delta = 0$. Thus, a decline in F_e/L causes both ψ_c and A to decline at the same rate, with ψ_c/A unchanged, as shown in Figure 6a.

One could also show that

$$\frac{dM}{d(F_e/L)} < 0;\ \ \mathcal{E}_G'(\psi) \lesseqqgtr 0 \Leftrightarrow \frac{d[MG(\psi_c)]}{d(F_e/L)} \lesseqqgtr 0,$$

whose proof is omitted because it follows the same line with the proof of Proposition 9a and 9b.

Figure 6b illustrates the impacts on the markup rate, the profit and the revenue. While a decline in F_e causes the profit and the revenue of all surviving firms to decline with proportionately larger impacts on low-ψ firms, an increase in L causes the profit and revenue to go up among low-ψ firms. The profit and revenue always go down among high-ψ firms, with the clockwise rotation of the profit and revenue schedule, whose pivot point (ψ_0 for the profit; ψ_1 for the revenue) is always located below the cutoff ψ_c, because the cutoff firms always earn zero revenue and profit.

6. Sorting of Heterogeneous Firms: A Multi-Market Extension

6.1. A Multi-Market Setting

We now extend the model to have $J \geq 2$ markets, indexed as $j = 1,2, \dots, J$, from which firms need to choose. The structure of each market is as before; it produces a single consumption good with the H.S.A. technology to assemble market-specific differentiated intermediate inputs supplied by monopolistically competitive producers. The only source of the heterogeneity across markets is market size. The aggregate expenditure for good-j is L_j, with $\sum_{j=1}^{J} L_j = L$, so that $\beta_j = L_j/L > 0$ is its expenditure share. One possible interpretation is that the representative household has the Cobb-Douglas preferences over J consumption goods, $U = \sum_{j=1}^{J} \beta_j \ln C_j$, to be maximized subject to the budget constraint, $\sum_{j=1}^{J} P_j C_j = L$. Another possible interpretation is that there are J different types of households, with $\beta_j = L_j/L$ being the fraction of type-j households who consume only good-j. Here, the types of consumers can be based on the difference in their tastes or their locations. With their expenditure shares being the only exogenous source of heterogeneity, we index the markets such that $L_1 > L_2 >, \dots, > L_J > 0$, without further loss of generality. Labor is fully mobile across the markets, which we continue to use as the numeraire.

As before, each entrant must pay the entry cost, $F_e > 0$, to draw its marginal cost, ψ. Then, after learning its marginal cost, they decide which market to enter and produce with an overhead cost, $F > 0$, or exit without producing. If ψ-firms choose not to exit, they would enter the market that gives the highest profit to earn

$$\Pi_\psi = \max\{\Pi_{1\psi}, \dots, \Pi_{J\psi}\},$$

where

$$\Pi_{j\psi} \equiv \frac{s\left(Z(\psi/A_j)\right)}{\zeta\left(Z(\psi/A_j)\right)} L_j \equiv \frac{r(\psi/A_j)}{\sigma(\psi/A_j)} L_j = \pi\left(\frac{\psi}{A_j}\right) L_j$$

is the profit earned by ψ-firms by entering market-j and A_j is the inverse measure of competitive pressures in market-j. The free entry condition is then

$$\int_{\underline{\psi}}^{\overline{\psi}} \max\{\Pi_\psi - F, 0\}\, dG(\psi) = F_e.$$

6.2. Positive Assortative Matching Between Firms and Markets under A2

We now show a positive assortative matching between firms and markets under **A2** in the sense that more productive firms self-select into larger markets. Specifically, we show that there is a sequence of monotonically increasing cutoffs, $\underline{\psi} = \psi_0 < \psi_1 < \psi_2 < \cdots < \psi_J < \overline{\psi}$, such that firms with $\psi \in (\psi_{j-1}, \psi_j)$ enter market-j, and those with $\psi \in (\psi_J, \overline{\psi})$ do not enter any market.

First, we prove that A_j is strictly monotone in j. Suppose the contrary, so that, for some j, $L_j > L_{j+1}$ and $A_j \geq A_{j+1}$. Because $\pi(\cdot)$ is strictly decreasing, this would mean that, for all ψ,

$$\pi(\psi/A_j) \geq \pi(\psi/A_{j+1}) \Longrightarrow \Pi_{j\psi} = \pi(\psi/A_j)L_j > \pi(\psi/A_{j+1})L_{j+1} = \Pi_{(j+1)\psi}$$

which would imply that no firm would enter market-$(j+1)$, and hence $A_{j+1} = \infty$, a contradiction. Thus, $0 < A_1 < A_2 < \cdots < A_J < \infty$, and $\pi(\psi/A_1) < \pi(\psi/A_2) < \cdots < \pi(\psi/A_J)$ for all ψ.

Second, for $j = 1,2, \dots, J-1$, consider the following ratio:

$$\frac{\Pi_{j\psi}}{\Pi_{(j+1)\psi}} = \frac{\pi(\psi/A_j)L_j}{\pi(\psi/A_{j+1})L_{j+1}}.$$

As a function of ψ, this ratio has to be greater than one for some ψ and less than one for other ψ, to ensure that a positive measure of firms would enter both market-j and market-$(j+1)$. Since **A2** implies that $\pi(\psi/A)$ is strictly log-supermodular in ψ and A (Proposition 2c), this ratio is strictly decreasing in ψ because $A_j < A_{j+1}$. Thus, there exists ψ_j such that

$$\psi \lesseqgtr \psi_j \Leftrightarrow \frac{\Pi_{j\psi}}{\Pi_{(j+1)\psi}} = \frac{\pi\big(\psi/A_j\big)L_j}{\pi\big(\psi/A_{j+1}\big)L_{j+1}} \gtreqless \frac{\pi\big(\psi_j/A_j\big)L_j}{\pi\big(\psi_j/A_{j+1}\big)L_{j+1}} \equiv 1.$$

In other words, all firms with $\psi < \psi_j$ strictly prefer entering market-j to entering market-$(j+1)$, all firms with $\psi > \psi_j$ strictly prefer entering market-$(j+1)$ to entering market-j, and all firms with $\psi = \psi_j$ are indifferent between the two markets. For $j = J$, let ψ_J be defined by $\pi\big(\psi_J/A_J\big)L_J \equiv F$. Then, only the firms with $\psi \in [\psi_{j-1}, \psi_j]$ enter market-j. This also means that ψ_j is strictly monotone in j, because $\psi_{j-1} \geq \psi_j$ would imply that a positive measure of firms would not enter market-j, which is a contradiction. See Figure 7. Thus, the mass of the active firms in market-j is equal to $M[G(\psi_j) - G(\psi_{j-1})]$, and the mass of the firms that enter and choose not to stay in any market is $M[1 - G(\psi_J)]$.

The free entry condition can now be rewritten as:

$$\sum_{j=1}^{J} \int_{\psi_{j-1}}^{\psi_j} \left\{\pi\left(\frac{\psi}{A_j}\right)L_j - F\right\} dG(\psi) = F_e \tag{10}$$

The adding up constraint in market-j is given by:

$$M \int_{\psi_{j-1}}^{\psi_j} r\left(\frac{\psi}{A_j}\right) dG(\psi) = 1, \tag{11}$$

where the cutoff rules are:

$$\frac{\pi\big(\psi_j/A_j\big)L_j}{\pi\big(\psi_j/A_{j+1}\big)L_{j+1}} = 1, \tag{12}$$

for $j = 1,2, \dots, J-1$, and

$$\pi\left(\frac{\psi_J}{A_J}\right)L_J \equiv F, \tag{13}$$

for $j = J$. Altogether, these $2J+1$ conditions in eqs.(10)-(13) determine $2J+1$ endogenous variables, which are M, $\{A_j\}_{j=1}^{J}$ and $\{\psi_j\}_{j=1}^{J}$, $0 < A_1 < A_2 < \cdots < A_J < \infty$; $\underline{\psi} = \psi_0 < \psi_1 < \psi_2 < \cdots < \psi_J < \overline{\psi}$. To summarize,

Proposition 10: Positive Assortative Matching between Firm Productivity and Market Size
Suppose that J markets differ only in market size, as $L_1 > L_2 > \cdots > L_J > 0$. In equilibrium, large markets are characterized by more competitive pressures, $0 < A_1 < A_2 < \cdots < A_J < \infty$. And under **A2,** firms with $\psi \in (\psi_{j-1}, \psi_j)$ enter market-j for $j = 1, 2, \ldots, J$, and firms with $\psi \in (\psi_J, \overline{\psi})$ exit, with $\underline{\psi} = \psi_0 < \psi_1 < \psi_2 < \cdots < \psi_J < \overline{\psi}$, where the two strictly increasing sequences, $\{\psi_j\}_{j=1}^J$ and $\{A_j\}_{j=1}^J$, and M, the mass of entrant, are given by eqs.(10)-(13).

Note that **A2** is crucial for this result. Under the opposite of **A2**, $\pi(\psi/A)$ would be strictly log-submodular in ψ and A, so that $\pi(\psi/A_j)L_j/\pi(\psi/A_{j+1})L_{j+1}$ would be strictly *increasing* in ψ. Hence the equilibrium would feature a strictly *decreasing* sequence, $\underline{\psi} = \psi_J < \cdots < \psi_2 < \psi_1 < \psi_0 < \overline{\psi}$, such that the firms with $\psi \in (\psi_j, \psi_{j-1})$ enter market-j, and those with $\psi \in (\psi_0, \overline{\psi},)$ exit. Thus, there would be a *negative* assortative matching with more productive firms self-selecting into smaller markets. Under CES, $\pi(\psi/A_j)L_j/\pi(\psi/A_{j+1})L_{j+1}$ is *independent* of ψ, hence the model does not predict any sorting. Indeed, in equilibrium, this ratio has to be equal to one so that firms are indifferent across all markets, and the equilibrium distribution would be *indeterminate*.

The Melitz model under H.S.A. thus offers a new mechanism for the positive assortative matching between firm productivity and the city size: Marshall's 2nd law. This demand-side mechanism complements the supply-side mechanisms studied in the literature. For example, what generates the positive assortative matching in Behrens, Duranton, and Robert-Nicoud (2014) and Gaubert (2018), both of which use CES, is the assumption on the firm technology that more productive firms are better at leveraging local agglomeration externalities in larger cities, similar to what Davis and Dingel (2019) assumed in the context of sorting of workers across the cities.[42]

6.3. Cross-Sectional, Cross-Market Patterns

[42]Baldwin and Okubo (2006) also considered sorting of heterogeneous firms in a spatial context under the CES demand. The positive assortative matching in their model is due to their equilibrium selection criterion based on the protocol that larger firms choose in which markets to locate earlier, which they argue is plausible because larger firms gain more (but not proportionately) from moving to the larger markets. Some criticize this protocol as ad hoc, because smaller firms may move faster since they are more agile. Our analysis suggests that such a criticism is unwarranted because, if we consider their CES demand as a limit of the H.S.A. demand under A2, the same equilibrium will be selected.

Figures 8a-8d illustrate the patterns of the profit, the revenue, the markup rates, and the pass-through rates across firms that emerge in equilibrium as more productive firms sort themselves into larger markets.

The profit schedule, $\Pi_{\psi} = \max_{j}\{\pi(\psi/A_j)L_j\}$, shown in Figure 8a, is obtained by the upper envelope of $\pi(\psi/A_j)L_j$. It is globally continuous and strictly decreasing in ψ, with the kink at the cutoff point, ψ_j. It is continuous at each cutoff, ψ_j, because the lower markup rate in market-j cancels out its larger market size, keeping ψ_j-firms indiffierent btw market-j & market-$(j+1)$.

The revenue schedule, R_{ψ}, shown in Figure 8b, is continuously decreasing in ψ within each market. However, it exhibits a downward jump at the cutoff ψ_j $(j = 1, 2, \dots, J-1)$, as

$$\frac{r(\psi_j/A_j)L_j}{r(\psi_j/A_{j+1})L_{j+1}} = \frac{\sigma(\psi_j/A_j)\pi(\psi_j/A_j)L_j}{\sigma(\psi_j/A_{j+1})\pi(\psi_j/A_{j+1})L_{j+1}} = \frac{\sigma(\psi_j/A_j)}{\sigma(\psi_j/A_{j+1})} > 1.$$

This is because, if ψ_j-firms switch from market-$(j+1)$ to larger-but-more-competitive market-j, they need to lower the markup rate, so that they need to earn higher revenue in market-j than in market-$(j+1)$ to keep them indiffierent between the two markets. In spite of these discontinuities, R_{ψ}, is globally strictly decreasing in ψ.

On the other hand, the markup rate schedule, μ_{ψ}, shown in Figure 8c, is not globally monotonic in ψ. It is continuously decreasing in ψ within each market. At the cutoff ψ_j $(j = 1, 2, \dots, J-1)$, however, it jumps upward. This is because $A_j < A_{j+1}$ so that switching from market-j to smaller-but-less-competitive market-$(j+1)$ allows ψ_j-firms to increase the markup rates from $\mu(\psi_j/A_j)$ to $\mu(\psi_j/A_{j+1})$. The markup rate, μ_{ψ}, thus exhibits a sawtooth pattern.

Likewise, the pass-through rate schedule, ρ_{ψ}, is not generally monotonic. Figure 8d shows the schedule under the strong A3. It is continuously increasing in ψ within each market. At the cutoff ψ_j $(j = 1, 2, \dots, J-1)$, however, it jumps downward. This is because $A_j < A_{j+1}$ so that switching from market-j to smaller-but-less-competitive market-$(j+1)$ allows ψ_j-firms to reduce the pass-through rates from $\rho(\psi_j/A_j)$ to $\rho(\psi_j/A_{j+1})$. The pass-through rate, ρ_{ψ}, thus exhibits a sawtooth pattern.

6.4. Average Markup and Pass-Through Rates in a Multi-Market Model: The Composition Effect

Under A2, more productive firms have higher markup rates than less productive firms if they face the same level of competitive pressures. However, more productive firms sort themselves into large and hence more competitive markets. This generates the sawtooth pattern in Figure 8c. Due to this composition effect, the average markup rates in large and hence more competitive markets be higher. Likewise, under A2 and the strong A3, more productive firms have lower pass-through rates than less productive firms if they face the same level of competitive pressures. However, more productive firms also sort themselves into large and hence more competitive markets, which generates the sawtooth pattern in Figure 8d. Due to this composition effect, the average pass-through rates in larger and hence more competitive markets might be higher, as demonstrated in Proposition 11a. Proposition 11b also demonstrates the possiblity that, due to an exogenous shock that causes all markets to become more competitive, the average markup rates to go up and the average pass-through rates to go down in all markets due to the shift in the composition. The proofs of these propositions are in Appendix C.6.

Proposition 11a: *Suppose A2 and* $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}$. *There exists a sequence,* $L_1 > L_2 > \cdots > L_J > 0$*, such that, in equilibrium, the weighted average of* $f\left(\psi/A_j\right)$ *across firms operating at market-*j *are increasing (decreasing) in* j *even though* $f(\cdot)$ *is increasing (decreasing) and hence* $f\left(\psi/A_j\right)$ *is decreasing (increasing) in* j.

Proposition 11a suggests an example with $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}$, in which the average markup rates are *higher* under A2 (and the average pass-through rates are *lower* under Strong A3) in larger markets.

Proposition 11b: *Suppose A2 and* $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}$. *Then, a change in* F_e *keeps*

i) *the ratios* $a_j \equiv \psi_{j-1}/\psi_j$ *and* $b_j \equiv \psi_j/A_j$

and

ii) *the weighted average of* $f\left(\psi/A_j\right)$ *across firms operating at market-*j*, for any weighting function* $w\left(\psi/A_j\right)$,

unchanged for all $j = 1,2,\ldots,J$.

Proposition 11b suggests that a decline in F_e under $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}$ offers a knife-edge case, where the average markup and pass-through rates of all markets remain unchanged.

Propositions 11a and 11b thus suggests a caution when testing A2 and A3 by comparing the average markup & pass-through rates across space and time.

7. Concluding Remarks

In this paper, we apply the H.S.A. (*Homotheticity with a Single Aggregator*) class of demand systems to the Melitz (2003) model of monopolistic competition with firm heterogeneity. H.S.A. is tractable due to its homotheticity and to its single aggregator that serves as a sufficient statistic for competitive pressures. The single aggregator property makes it possible to prove the existence and uniqueness of the free-entry equilibrium and to conduct general equilibrium comparative static analysis, using simple diagrams. It is also flexible enough to allow for the choke price, the 2nd law of demand, and what we call the 3rd law of demand. Furthermore, because the single aggregator enters all firm-specific variables proportionately with the firm-specific marginal cost, and hence acting as a magnifieer of firm heteroneity, we are able to characterize, by taking advantage of log-supermodularity, how a change in competitive pressures, whether due to a change in the entry cost, market size, or in the overhead cost, affects heterogeneous firms differently, and thereby causing reallocation across firms, and hence selection of firms, and sorting of firms across different markets.

References:

Amiti, M., O. Itskhoki, and J. Konings (2014), "Importers, Exporters, and Exchange Rate Disconnect. American Economic Review, 104(7), 1942–1978.

Amiti, M., O. Itskhoki, and J. Konings (2019), "International Shocks, Variable Markups, and Domestic Prices," The Review of Economic Studies.

Arkolakis, C., A. Costinot, D. Donaldson, and Rodríguez-Clare (2019), "The Elusive Pro-Competitive Effects of Trade, The Review of Economic Studies, 86 (1), 46–80.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen, "The Fall of the Labor Share and the Rise of Superstar Firms," Quarterly Journal of Economics, 135 (2020), 645–709.

Baldwin, R.E., and T. Okubo (2006). "Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting." Journal of Economic Geography 6 (3): 323–46.

Bagnoli, M., and T. Bergstrom (2005), "Log-concave probability and its applications," Economic Theory, 26, 445-469.

Baqaee, D., E. Fahri, and K. Sangani (2021), "The Darwinian Returns to Scale," UCLA and Harvard.
Behrens, K., G. Duranton, and F. Robert-Nicoud (2014), "Productive Cities: Sorting, Selection, and Agglomeration." Journal of Political Economy 122 (3): 507–53.
Behrens, K., G. Mion, Y. Murata, and J. Südekum (2020), "Quantifying the gap between equilibrium and optimum under monopolistic competition", Quarterly Journal of Economics, 135(4), 2299-2360.
Behrens, K. and Y. Murata, (2007), "General Equilibrium Models of Monopolistic Competition: A New Approach," *Journal of Economic Theory* 136, 776-787.
Berman, N., P. Martin, and T. Mayer (2012), "How Do Different Exporters React to Exchange Rate Changes? The Quarterly Journal of Economics, 127(1), 437–492.
Burstein, A., and G. Gopinath, G. (2014), "International Prices and Exchange Rates," in G. Gopinath, E. Helpman, and K. Rogoff (eds.), Handbook of International Economics , vol.4, 391-451, Amsterdam: Elsevier.
Campbell, Jeffrey R., and Hugo A. Hopenhayn, "Market size matters," The Journal of Industrial Economics, March 2005, Vol.53, 1-25.
Costinot, A. (2009): An Elementary Theory of Comparative Advantage, Econometrica, 77, 1165-1192.
Costinot, A., and J. Vogel (2010): Matching and Inequality in the World Economy, Journal of Political Economy, 118, 747-786.
Costinot, A., and J. Vogel (2015): Beyond Ricardo: Assignment Models in International Trade, Annual Review of Economics, 7, 31-62.
Davis, D., and J. Dingel (2019). "A Spatial Knowledge Economy," American Economic Review.
De Loecker, J., and P.K. Goldberg (2014), "Firm Performance in a Global Market," Annual Review of Economics, 6(1), 201–227.
De Loecker, J., J. Eeckhout, and G. Unger, (2020), "The Rise of Market Power and the Macroeconomic Implications," Quarterly Journal of Economics, 135, 561–644.
Dhingra, S., and J. Morrow (2019), "Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity," Journal of Political Economy, vol.127, no.1, 196-232.
Dixit, Avinash K., Joseph E. Stiglitz, "Monopolistic Competition and Optimal Product Diversity," American Economic Review, June 1977, 67(3), 297-308.
Edmond, C., V. Midrigan, and D. Y. Xu, (2021), "How Costly are Markups?" Melbourne, NYU, and Duke.
Feenstra, Robert C., "A Homothetic utility function for monopolistic competition models, without constant price elasticity," Economics Letters 78 (2003), 79-86.
Feenstra, Robert C. and David E. Weinstein, "Globalization, Markups, and US Welfare," Journal of Political Economy, 2017, 125(4), 1040-1074.
Fujiwara, I., and K. Matsuyama (2022), "Competition and the Phillips Curve," Keio and Northwestern.
Gaubert, C. (2018), "Firm Sorting and Agglomeration," American Economic Review, 108(11): 3117-3153.
Grossman, G. M., E. Helpman, and H. Luillier (2021), "Supply Chain Resilience: Should Policy Promote Diversification or Reshoring," Princeton and Harvard University.
Head, K., T. Mayer, M. Theonig, "Welfare and Trade without Pareto," AER PP, May 2014.
Kasahara, H., and Y. Sugita, (2020), "Nonparametric Identification of Production Function, Total Factor Productivity, and Markup from Revenue Data," UBC and Hitotsubashi.

Kimball, M. (1995), “The Quantitative Analytics of the Basic Neomonetarist Model,” Journal of Money, Credit and Banking, 27 (4), 1241–77.
Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade, Journal of International Economics 9 (4), 469{479.
Latzer, H., K. Matsuyama, and M. Parenti, (2019) “Reconsidering the Market Size Effect in Innovation and Growth,” CEPR DP #14250.
Matsuyama, K., and P. Ushchev (2017), “Beyond CES: Three Alternative Classes of Flexible Homothetic Demand Systems,” CEPR DP #12210.
Matsuyama, K., and P. Ushchev (2020a), “When Does Procompetitive Entry Imply Excessive Entry?” CEPR DP #14991.
Matsuyama, K., and P. Ushchev (2020b), “Constant Pass-Through,” CEPR DP #15475.
Matsuyama, K., and P. Ushchev (2022), “Destabilizing Effects of Market Size in the Dynamics of Innovation,” Journal of Economic Theory,
Melitz, M.J. (2003), “The Impact of Trade on Intra-Industry Reallocations, and Aggregate Industry Productivity,” Econometrica, 71, 6, 1695–1725.
Melitz, M.J., (2018), “Competitive Effects of Trade: Theory and Measurement,” Review of World Economics, 154 (1), 1–13.
Melitz, M.J., and G.I.P. Ottaviano (2008), “Market size, trade, and productivity,” The Review of Economic Studies, 75 (1), 295–316.
Melitz, M.J., and S. J. Redding (2014), “Heterogeneous Firms and Trade.” In *Handbook of International Economics*. Vol. 4, edited by Elhanan Helpman, Kenneth Rogoff, and Gita Gopinath, 1–54. Amsterdam: Elsevier North Holland.
Mrázová, M., and J. P. Neary (2017), “Not So Demanding: Demand Structure and Firm Behavior,” American Economic Review, 107(12): 3835–3874
Mrázová, M., and J. P. Neary (2019), “Selection Effects with Heterogeneous Firms,” Journal of the European Economic Association 17(4):1294–1334.
Nocke, V. (2006), “A Gap for Me: Entrepreneurs and Entry.” Journal of the European Economic Association 4 (5): 929–56.
Ottaviano, G.I.P., T.Tabuchi, and J.-F. Thisse, (2002). “Agglomeration and trade revisited,” International Economic Review, 43, 409–435.
Parenti, M., P. Ushchev, and J-F. Thisse, (2017), “Toward a Theory of Monopolistic Competition,” *Journal of Economic Theory,* 167, 86-115.
Thisse, J.-F., and P. Ushchev (2018), Monopolistic Competition without Apology, Ch 5 in *Handbook of Game Theory and Industrial Organization*, Vol. I, ed by L. Cochon and M. Marini, Edgar Publication.
Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and Jacques-Francois Thisse, “Monopolistic Competition: Beyond the Constant Elasticity of Substitution,” *Econometrica,* November 2012, Vol.80, No.6, 2765-2784.

Appendices for

Selection and Sorting of Heterogeneous Firms through Competitive Pressures

Kiminori Matsuyama and Philip Ushchev

Date: 2022-03-06, Time: 11:43 AM

Appendix A: Firm type distributions and their elasticities

Let the distribution of the marginal cost, ψ, be given by its cdf, $G(\psi)$, with the support, $\left(\underline{\psi},\overline{\psi}\right) \subseteq (0,\infty)$, and hence that of productivity, $\varphi = 1/\psi$, be given by its cdf, $F(\varphi) = 1 - G(1/\varphi)$, with the support, $\left(\underline{\varphi},\overline{\varphi}\right) = \left(1/\overline{\psi}, 1/\underline{\psi}\right) \subseteq (0,\infty)$. We assume that these cdfs are thrice continuously differentiable, C^3, and hence that their pdfs satisfy, $G'(\psi) = g(\psi) > 0$ on $\left(\underline{\psi},\overline{\psi}\right)$ and $F'(\varphi) = f(\varphi) > 0$ on $\left(\underline{\varphi},\overline{\varphi}\right)$ and are twice continuously differentiable, C^2, so that $\mathcal{E}_G(\psi) \equiv \psi g(\psi)/G(\psi) \in C^2$, $\mathcal{E}_g(\psi) \equiv \psi g'(\psi)/g(\psi) \in C^1$ and $\mathcal{E}_F(\varphi) \equiv \varphi f(\varphi)/F(\varphi) \in C^2$, $\mathcal{E}_f(\varphi) \equiv \varphi f'(\varphi)/f(\varphi) \in C^1$ It is straightforward to show that:

$$\varphi f(\varphi) = \psi g(\psi);$$

$$\mathcal{E}_f(\varphi) + \mathcal{E}_g(\psi) = -2;$$

and

$$\varphi \mathcal{E}_f'(\varphi) = \psi \mathcal{E}_g'(\psi).$$

We also assume that the mean productivity is finite:

$$\int_{\underline{\varphi}}^{\overline{\varphi}} \varphi f(\varphi) d\varphi = \int_{\underline{\psi}}^{\overline{\psi}} \psi^{-1} g(\psi) d\psi < \infty.$$

This is guaranteed if $\underline{\psi} > 0 \Leftrightarrow \overline{\varphi} < \infty$. If $\underline{\psi} = 0 \Leftrightarrow \overline{\varphi} = \infty$, a sufficient condition for the finite mean productivity is given by:

$$-\lim_{\psi\to 0} \mathcal{E}_g(\psi) = \lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) + 2 < 0.$$

To see this, note that $\lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) + 2 < 0 \Leftrightarrow \lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) + 1 < -1$ implies that $\varphi f(\varphi)$ decreases faster than $1/\varphi$ as $\varphi \to \infty$, $\int_{\underline{\varphi}}^{\infty} \varphi f(\varphi) d\varphi < \infty$.[43]

Lemma 1:

$$\mathcal{E}_g'(\psi) < 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \mathcal{E}_G'(\psi) < 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right).$$

Furthermore, if $\underline{\psi} = 0$ *and* $\lim_{\psi\to 0} \psi g(\psi) = 0$,

$$\mathcal{E}_g'(\psi) \geq 0, \forall\psi \in \left(0, \overline{\psi}\right) \Rightarrow \mathcal{E}_G'(\psi) \geq 0, \forall\psi \in \left(0, \overline{\psi}\right).$$

Proof:[44] $\mathcal{E}_g'(\psi) \gtreqless 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right)$ implies

$$\left[\mathcal{E}_g(\psi) + 1\right] G(\psi) = \left[\mathcal{E}_g(\psi) + 1\right] \int_{\underline{\psi}}^{\psi} g(\xi) d\xi \gtreqless \int_{\underline{\psi}}^{\psi} \left[\mathcal{E}_g(\xi) + 1\right] g(\xi) d\xi$$

$$= \int_{\underline{\psi}}^{\psi} [\xi g'(\xi) + g(\xi)] d\xi = \int_{\underline{\psi}}^{\psi} d[\xi g(\xi)] = \psi g(\psi) - \lim_{\psi\to\underline{\psi}} \psi g(\psi),$$

which in turn implies

$$\mathcal{E}_G'(\psi) = \frac{d}{d\psi}\left[\frac{\psi g(\psi)}{G(\psi)}\right] = \frac{[\psi g'(\psi) + g(\psi)] G(\psi) - \psi [g(\psi)]^2}{[G(\psi)]^2}$$

$$= \frac{g(\psi)}{[G(\psi)]^2}\left\{\left[\mathcal{E}_g(\psi) + 1\right] G(\psi) - \psi g(\psi)\right\} \gtreqless -\frac{g(\psi)}{[G(\psi)]^2}\left[\lim_{\psi\to\underline{\psi}} \psi g(\psi)\right].$$

Hence, the first part always holds, while the second part holds because $\lim_{\psi\to 0} \psi g(\psi) = 0$.

This completes the proof. ■

---

[43]Equivalently, $-\lim_{\psi\to 0} \mathcal{E}_g(\psi) < 0 \Leftrightarrow \lim_{\psi\to 0} \mathcal{E}_g(\psi) - 1 > -1$, implies that $\psi^{-1} g(\psi)$ increases slower than ψ^{-1} as $\psi \searrow 0$, hence $\int_0^{\overline{\psi}} \psi^{-1} g(\psi) d\psi < \infty$. Even though this condition for the finite mean productivity is sufficient but not necessary, it is close to being necessary in the sense that the mean productivity is infinite if $-\lim_{\psi\to 0} \mathcal{E}_g(\psi) = \lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) + 2 > 0$. The case of $-\lim_{\psi\to 0} \mathcal{E}_g(\psi) = \lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) + 2 = 0$ would require case-by-case scrutiny.

[44] For the second part of Lemma 1, we consider only the case of $\underline{\psi} = 0$, since $\underline{\psi} > 0$ and $\lim_{\psi\to\underline{\psi}} \psi g(\psi) = 0$ would imply $\lim_{\psi\to\underline{\psi}} g(\psi) = 0$, so that $\lim_{\psi\to\underline{\psi}} \mathcal{E}_g(\psi) = \infty$, hence $\mathcal{E}_g'(\psi) < 0$ for ψ close to $\underline{\psi} > 0$. Thus, it would be impossible to satisfy $\mathcal{E}_g'(\psi) \geq 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right)$. It is also worth noting that the second part would fail if $\underline{\psi} > 0$ and $\lim_{\psi\to\underline{\psi}} \psi g(\psi) > 0$. [An example is a truncated power, for which $\mathcal{E}_g'(\cdot) = 0$ but $\mathcal{E}_G'(\cdot) \neq 0$.] Lemma 1 can also be obtained as a corollary of Theorem 1 and Theorem 2 of Bagnoli and Bergstrom (2005) by noting that $\mathcal{E}_G'(\cdot) < (>) 0$ if and only if the cdf of $\theta \equiv \ln\psi$, $G\left(e^{\theta}\right)$, is log-concave (log-convex) and that $\mathcal{E}_g'(\cdot) < (>) 0$ if and only if the density of θ, $e^{\theta} g\left(e^{\theta}\right)$, is log-concave (log-convex).

The following lemma states how a change in ψ_c shifts the distribution of $\xi \equiv \psi/\psi_c$, the marginal cost relative to the cutoff marginal cost, ψ_c, among surviving firms. It shows that, if $\mathcal{E}_g(\cdot)$ is increasing (decreasing), an increase in ψ_c causes a shift to the right (left) in the sense of the monotone likelihood ratio ordering; and that, if $\mathcal{E}_G(\cdot)$ is increasing (decreasing), an increase in ψ_c causes a shift to the right (left) in the sense of the first-order stochastic dominance.

Lemma 2: *Define* $\xi \equiv \psi/\psi_c \in \left(\underline{\xi}, 1\right)$, *where* $\underline{\xi} \equiv \underline{\psi}/\psi_c$. *Consider a cdf,*

$$\tilde{G}(\xi; \psi_c) \equiv \frac{G(\psi_c \xi)}{G(\psi_c)},$$

and its density function,

$$\tilde{g}(\xi; \psi_c) \equiv \frac{d\tilde{G}(\xi; \psi_c)}{d\xi} = \frac{\psi_c g(\psi_c \xi)}{G(\psi_c)},$$

whose support is $\left(\underline{\xi}, 1\right)$ *with* $\tilde{G}\left(\underline{\xi}; \psi_c\right) = 0$ *and* $\tilde{G}(1; \psi_c) = 1$. Then,

$$\mathcal{E}_g'(\xi) \gtreqless 0, \forall \xi \in \left(\underline{\xi}, 1\right) \Rightarrow \frac{\partial^2 \ln \tilde{g}(\xi; \psi_c)}{\partial \xi \partial \psi_c} \gtreqless 0, \forall \xi \in (\underline{\xi}, 1)$$

and

$$\mathcal{E}_G'(\xi) \gtreqless 0, \forall \xi \in \left(\underline{\xi}, 1\right) \Rightarrow \frac{\partial \tilde{G}(\xi; \psi_c)}{\partial \psi_c} \lesseqgtr 0, \forall \xi \in \left(\underline{\xi}, 1\right).$$

Proof: The first statement follows from

$$\frac{\partial^2 \ln \tilde{g}(\xi; \psi_c)}{\partial \xi \partial \psi_c} = \frac{\partial^2 \ln g(\psi_c \xi)}{\partial \xi \partial \psi_c} = \mathcal{E}_g'(\psi_c \xi) \gtreqless 0, \qquad \forall \xi \in \left(\underline{\xi}, 1\right).$$

The second statement follows from

$$\frac{\partial \ln \tilde{G}(\xi; \psi_c)}{\partial \ln \psi_c} = \frac{\partial \ln[G(\psi_c \xi)/G(\psi_c)]}{\partial \ln \psi_c} = \mathcal{E}_G(\psi_c \xi) - \mathcal{E}_G(\psi_c) \lesseqgtr 0, \forall \xi \in (\underline{\xi}, 1),$$

if $\mathcal{E}_G'(\xi) \gtreqless 0$. This completes the proof. ■

The signs of $\mathcal{E}_g'(\cdot)$ and of $\mathcal{E}_G'(\cdot)$ play critical roles for some of the comparative statics results. Thus, we now list some parametric families of distributions (widely used in the literature), for which the sign of $\mathcal{E}_g'(\cdot)$ never changes over the support, which also means, from Lemma 1, that the sign of $\mathcal{E}_G'(\cdot)$ never changes over the support, either.

Example 1: Pareto (or power) distribution. The cdfs are given by

$$F(\varphi) = 1 - \left(\varphi/\underline{\varphi}\right)^{-\kappa} \Leftrightarrow G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa};$$

for $\varphi > \underline{\varphi} > 0 \iff 0 < \psi < \overline{\psi} < \infty$. The pdfs satisfy:

$$\varphi f(\varphi) = \kappa\left(\varphi/\underline{\varphi}\right)^{-\kappa} = \kappa\left(\psi/\overline{\psi}\right)^{\kappa} = \psi g(\psi)$$

Hence, $\mathcal{E}_f(\varphi) = -\kappa - 1$ and $\mathcal{E}_g(\psi) = \kappa - 1$, so that $\mathcal{E}'_f(\varphi) = \mathcal{E}'_g(\psi) = 0$. The condition for the finite mean productivity is given by $\kappa > 1$.

Example 2: Generalized Pareto (Power) distribution. The generalized Pareto (Power) family nests Pareto (Power) as a special case and allows all the three possibilities for $sgn\{\mathcal{E}'_f(\cdot)\} = sgn\{\mathcal{E}'_g(\cdot)\}$ to depend on the parameter values. The cdfs are given by

$$F(\varphi) = 1 - \left(1 + \frac{\varphi - \underline{\varphi}}{\lambda}\right)^{-\kappa}, \qquad \varphi > \underline{\varphi} > 0, \qquad \lambda > 0.$$

$$G(\psi) = \left(1 + \frac{1/\psi - 1/\overline{\psi}}{\lambda}\right)^{-\kappa}, \qquad 0 < \psi < \overline{\psi} < \infty, \qquad \lambda > 0.$$

Hence, the pdfs satisfy:

$$\varphi f(\varphi) = \frac{\varphi\kappa}{\lambda}\left(1 + \frac{\varphi - \underline{\varphi}}{\lambda}\right)^{-\kappa-1} = \frac{\kappa}{\psi\lambda}\left(1 + \frac{1/\psi - 1/\overline{\psi}}{\lambda}\right)^{-\kappa-1} = \psi g(\psi)$$

from which

$$\mathcal{E}_f(\varphi) = -(1+\kappa)\left(\frac{\varphi}{\lambda - \underline{\varphi} + \varphi}\right) = -(1+\kappa)\left(\frac{1/\psi}{\lambda - 1/\overline{\psi} + 1/\psi}\right) = -\mathcal{E}_g(\psi) - 2.$$

Clearly, the standard Pareto (Power) distribution is a special case with $\lambda = \underline{\varphi} = 1/\overline{\psi}$. More generally, one can readily verify that:

$$\psi\mathcal{E}'_g(\psi) = \varphi\mathcal{E}'_f(\varphi) = -(1+\kappa)\frac{\varphi\left(\lambda - \underline{\varphi}\right)}{\left(\lambda - \underline{\varphi} + \varphi\right)^2} \gtreqless 0 \iff \lambda \lesseqgtr \underline{\varphi} = 1/\overline{\psi}.$$

Example 3: Lognormal distribution. Since $\ln\varphi = -\ln\psi$, productivity is distributed lognormally if and only if the marginal cost is distributed lognormally. In this case, the support is $(0, \infty)$. For all $\varphi > 0$ and for all $\psi > 0$, the pdfs can be represented by

$$f(\varphi) = \frac{1}{\varphi\tilde{\sigma}\sqrt{2\pi}}\exp\left\{-\frac{(\log\varphi - \mu)^2}{2\tilde{\sigma}^2}\right\},$$

$$g(\psi) = \frac{1}{\psi\tilde{\sigma}\sqrt{2\pi}} \exp\left\{-\frac{(\log\psi + \mu)^2}{2\tilde{\sigma}^2}\right\},$$

where $\mu \in \mathbb{R}$ and $\tilde{\sigma} > 0$. The mean productivity is:

$$\int_0^\infty \varphi f(\varphi)d\varphi = \int_0^\infty \psi^{-1} g(\psi)d\psi = \exp\left\{\mu + \frac{\tilde{\sigma}^2}{2}\right\} < \infty.$$

The elasticities of the pdfs are strictly decreasing, because

$$\mathcal{E}_f(\varphi) = \frac{\mu - \log\varphi}{\tilde{\sigma}^2} - 1 = \frac{\mu + \log\psi}{\tilde{\sigma}^2} - 1 = -\mathcal{E}_g(\psi) - 2$$

$$\Rightarrow \varphi\mathcal{E}_f'(\varphi) = \psi\mathcal{E}_g'(\psi) = -\frac{1}{\tilde{\sigma}^2} < 0.$$

Hence, from Lemma 1, the elasticities of the cdfs are also strictly decreasing.

Example 4: Fréchet and Weibull distributions. The parametric families of Fréchet and Weibull distributions both belong to the class of extreme-value distributions.[45] When the distribution of φ is Fréchet (respectively, Weibull) if and only if that of $\psi = 1/\varphi$ is Weibull (respectively, Fréchet). Therefore, we consider the case of φ being Fréchet and omit the case of φ being Weibull.

For all $\varphi > 0$ and for all $\psi > 0$, the cdf of the Fréchet productivity distribution F and the corresponding Weibull cost distribution G are given, respectively, by

$$F(\varphi) = \exp\{-\varphi^{-\alpha}\}, \qquad G(\psi) = 1 - \exp\{-\psi^{\alpha}\},$$

where $\alpha > 0$. The pdfs are given by

$$f(\varphi) = \alpha\varphi^{-(1+\alpha)} \exp\{-\varphi^{-\alpha}\}, \qquad g(\psi) = \alpha\psi^{\alpha-1} \exp\{-\psi^{\alpha}\}.$$

Hence,

$$\mathcal{E}_f(\varphi) = -(1+\alpha) + \alpha\varphi^{-\alpha}, \;\; \mathcal{E}_g(\psi) = \alpha - 1 - \alpha\psi^{\alpha}$$

$$\Rightarrow \varphi\mathcal{E}_f'(\varphi) = -\alpha^2\varphi^{-\alpha} = -\alpha^2\psi^{\alpha} = \psi\mathcal{E}_g'(\psi) < 0,$$

so that the elasticities of the pdfs are strictly decreasing, and so are the elasticities of the cdfs from Lemma 1. The mean productivity is finite if and only if

$$-\lim_{\psi\to 0}\mathcal{E}_g(\psi) = \lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) + 2 = \alpha - 1 < 0 \Leftrightarrow \alpha > 1.$$

and given by:

[45] The third parametric family belonging to the class of extreme-value distributions is the Gumbel distribution. However, without any modification (e.g., truncation), it is not a legitimate distribution for φ or ψ since its support includes negative real numbers.

$$\int_0^\infty \varphi f(\varphi) d\varphi = \int_0^\infty \psi^{-1} g(\psi) d\psi = \Gamma\left(1 - \frac{1}{\alpha}\right) < \infty,$$

where $\Gamma(x)$ is the Gamma function.

$$\Gamma(x) \equiv \int_0^\infty y^{x-1} \exp\{-y\}\, dy.$$

Appendix B: A Sufficient Condition under which the equilibrium is well-defined.

For the equilibrium to be well-defined, the integrals in the free entry condition and the adding-up constraint must be both well-defined. Since

$$\pi\left(\frac{\psi}{A}\right) = \frac{r(\psi/A)}{\sigma(\psi/A)} < r\left(\frac{\psi}{A}\right),$$

it suffices to show that

$$\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi) < \infty.$$

First, we introduce the following lemma.

Lemma 3: $If\ \zeta(0) < \infty,\ \lim_{z\to 0} \frac{z\zeta'(z)}{\zeta(z)} = \lim_{z\to 0} \mathcal{E}_\zeta(z) = 0.$

Proof: This follows from $1 < \zeta(z) = \zeta(0) \exp\left[\int_0^z \frac{\xi\zeta'(\xi)}{\zeta(\xi)} \frac{d\xi}{\xi}\right] = \zeta(0) \exp\left[\int_0^z \mathcal{E}_\zeta(\xi) \frac{d\xi}{\xi}\right] < \infty$. ■

Lemma 4. *The above integral is finite and hence the equilibrium is well-defined and uniquely exists either if* $\underline{\psi} > 0 \Leftrightarrow \overline{\varphi} < \infty$ *or*

$$1 \leq \lim_{z\to 0} \zeta(z) < 2 + \lim_{\psi\to 0} \mathcal{E}_g(\psi) = -\lim_{\varphi\to\infty} \mathcal{E}_f(\varphi) < \infty,$$

for $\underline{\psi} = 0 \Leftrightarrow \overline{\varphi} = \infty$.

Proof. Clearly, the integral is well-defined if $\underline{\psi} > 0$. Now suppose $\underline{\psi} = 0$, and $1 \leq \lim_{z\to 0} \zeta(z) \equiv \zeta(0) < 2 + \lim_{\psi\to 0} \mathcal{E}_g(\psi) < \infty$. First, $1 \leq \zeta(0) < \infty$ implies $\lim_{z\to 0} \mathcal{E}_\zeta(z) = 0$ from Lemma 3.

Second, because

$$\frac{\partial \ln\left[r\left(\frac{\psi}{A}\right)g(\psi)\right]}{\partial \ln \psi} = \mathcal{E}_g(\psi) + \frac{\partial \ln\left[\pi\left(\frac{\psi}{A}\right)\right]}{\partial \ln \psi} + \frac{\partial \ln\left[\sigma\left(\frac{\psi}{A}\right)\right]}{\partial \ln \psi} = \mathcal{E}_g(\psi) - \frac{\left[\sigma\left(\frac{\psi}{A}\right) - 1\right]^2}{\sigma\left(\frac{\psi}{A}\right) - 1 + \mathcal{E}_\zeta\left(Z\left(\frac{\psi}{A}\right)\right)},$$

$$\lim_{\psi\to 0}\frac{\partial \ln\left[r\left(\frac{\psi}{A}\right)g(\psi)\right]}{\partial \ln \psi} = \lim_{\psi\to 0}\mathcal{E}_g(\psi) - \zeta(0) + 1 > -1,$$

where use has been made of $\lim_{z\to 0}\mathcal{E}_\zeta(z) = 0$ and $\zeta(0) < 2 + \lim_{\psi\to 0}\mathcal{E}_g(\psi)$. This inequality means that, for every finite $\psi_c > 0$, there exist $\Lambda(\psi_c) > 0$ and $\delta > 0$ such that,

$$\int_0^{\psi_c} r\left(\frac{\psi}{A}\right)g(\psi)d\psi < \int_0^{\psi_c}\Lambda(\psi_c)\psi^{\delta-1}d\psi = \Lambda(\psi_c)\frac{{\psi_c}^{\delta}}{\delta} < \infty.$$

This completes the proof. ■

It should be noted that the mean productivity is neither sufficient nor necessary for the existence of equilibrium. The equilibrium exists even when the mean productivity is infinite, if

$$1 \le \lim_{z\to 0}\zeta(z) < 2 + \lim_{\psi\to 0}\mathcal{E}_g(\psi) = -\lim_{\varphi\to\infty}\mathcal{E}_f(\varphi) < 2,$$

while the equilibrium fails to exist even when the mean productivity is finite if

$$\lim_{z\to 0}\zeta(z) > 2 + \lim_{\psi\to 0}\mathcal{E}_g(\psi) = -\lim_{\varphi\to\infty}\mathcal{E}_f(\varphi) > 2.$$

For example, $\zeta(z) = \sigma > 1$ under CES, and $\mathcal{E}_g(\psi) = \kappa - 1$ under a Power (Pareto), so that the equilibrium exists if $1 < \sigma < \kappa + 1$, and the mean productivity is finite if $\kappa > 1$. Hence, the equilibrium exists even when the mean productivity is infinite, if $1 < \sigma < \kappa + 1 < 2$, while the equilibrium fails to exist even when the mean productivity is finite, if $\sigma > \kappa + 1 > 2$.

Appendix C: Technical Proofs

C.1. Proof of Lemma 6

Lemma 6: *Under A2 and the weak A3,* $\lim_{\psi/A\to 0}\rho(\psi/A)\sigma(\psi/A) < 1 < \lim_{\psi/A\to \bar{z}}\rho(\psi/A)\sigma(\psi/A)$.

Proof: The proof proceeds in two steps.
Step 1: **A2** and the weak **A3** jointly imply

$$\lim_{\psi/A\to 0}\rho\left(\frac{\psi}{A}\right) < 1 \Leftrightarrow \lim_{z\to 0}\frac{z\zeta'(z)/\zeta(z)}{\zeta(z) - 1} > 0.$$

From Lemma 3, the numerator goes to zero, hence, $\lim_{z\to 0}\zeta(z) = \lim_{\psi/A\to 0}\sigma(\psi/A) = 1$, which proves $\lim_{\psi/A\to 0}\rho(\psi/A)\sigma(\psi/A) < 1$.

Step 2: For $\overline{z} < \infty$,

$$\lim_{z\to \overline{z}}\zeta(z) = \lim_{\psi/A\to \overline{z}}\sigma(\psi/A) = \infty \Longrightarrow \lim_{\psi/A\to \bar{z}}\rho(\psi/A)\sigma(\psi/A) = \infty.$$

For $\overline{z} = \infty$, if $\lim_{\psi/A\to \infty}\rho(\psi/A) = 1$,

$$\lim_{\psi/A\to \bar{z}}\rho(\psi/A)\sigma(\psi/A) = \lim_{\psi/A\to \bar{z}}\sigma(\psi/A) > 1.$$

On the other hand, if $\lim_{\psi/A\to\infty}\rho(\psi/A) < 1 \Leftrightarrow \lim_{z\to\infty}\frac{z\zeta'(z)/\zeta(z)}{\zeta(z)-1} > 0 \Leftrightarrow \lim_{z\to\infty}\frac{z\zeta'(z)}{\zeta(z)} > 0$,

$$\lim_{\psi/A\to\infty}\sigma\left(\frac{\psi}{A}\right) = \lim_{z\to\infty}\zeta(z) = \zeta(z')\exp\left[\int_{z'}^{\infty}\frac{\xi\zeta'(\xi)}{\zeta(\xi)}\frac{d\xi}{\xi}\right] = \infty \Longrightarrow \lim_{\psi/A\to\bar{z}}\rho\left(\frac{\psi}{A}\right)\sigma\left(\frac{\psi}{A}\right) = \infty.$$

Thus, in all of these cases,

$$\lim_{\psi/A\to \bar{z}}\rho(\psi/A)\sigma(\psi/A) > 1.$$

This completes the proof. ■

C.2. Proof of Proposition 5

To prove Proposition 5, we first need the following two lemmas.

Lemma 7:

$$\mathcal{E}_\rho\left(\frac{\psi}{A}\right) = \epsilon\left(Z\left(\frac{\psi}{A}\right)\right), where\ \epsilon(z) \equiv -\frac{z\theta'(z)}{[1+\theta(z)]^2}.$$

Proof: Straightforward from $0 < \rho(\psi/A) = \mathcal{E}_Z(\psi/A) = \frac{1}{1+\theta(Z(\psi/A))}$.

Lemma 8: For $0 \le \rho(0) < \infty, \lim_{z\to 0}\epsilon(z) = 0$.

Proof: From $\rho(\psi/A) = \frac{1}{1+\theta(Z(\psi/A))}$,

$$\rho\left(\frac{\psi}{A}\right) - \rho\left(\frac{\psi_0}{A}\right) = \frac{1}{1+\theta\big(Z(\psi/A)\big)} - \frac{1}{1+\theta\big(Z(\psi_0/A)\big)} = \int_{Z(\psi_0/A)}^{Z(\psi/A)}\frac{d}{d\xi}\left[\frac{1}{1+\theta(\xi)}\right]d\xi$$

$$= \int_{Z(\psi_0/A)}^{Z(\psi/A)}\left[-\frac{\theta'(\xi)}{[1+\theta(\xi)]^2}\right]d\xi \equiv \int_{Z(\psi_0/A)}^{Z(\psi/A)}\frac{\epsilon(\xi)}{\xi}d\xi,$$

for any $\psi_0 > 0$. From $0 \le \rho(0) < \infty$, the RHS remains bounded as $z_0 = Z(\psi_0/A) \to 0$. Hence,

$$\int_0^z \frac{\epsilon(\xi)}{\xi} d\xi < \infty,$$

which implies $\lim_{z \to 0} \epsilon(z) = 0$.

Proposition 5: Suppose that A2 and the strong A3 hold, so that $0 < \rho(\psi/A) < 1$ and $\rho(\psi/A)$ is strictly increasing. Then, $\rho(\psi/A)$ is strictly log-submodular for all $\psi/A < \bar{z}$ with a sufficiently small $\bar{z}$.

Proof: Under **A2**, $\rho(\psi/A) < 1$ for all $\psi/A < \bar{z}$, hence the condition for Lemma 8 holds and $\lim_{z \to 0} \epsilon(z) = 0$. Under the strong **A3**, $\epsilon(z) \equiv -z\theta'(z)/[1+\theta(z)]^2 > 0$ for all $z > 0$. Thus, $\epsilon(\cdot) > 0$ is increasing for a sufficiently small $z > 0$. Hence, from Lemma 7, $\mathcal{E}_\rho(\psi/A)$ is strictly increasing in ψ/A for $\psi/A < Z(\psi/A) < \bar{z}$, with a sufficiently small $\bar{z}$. Hence, from Lemma 5, $\rho(\psi/A)$ is strictly log-submodular for any $\psi/A < Z(\psi/A) < \bar{z}$.

C.3. Proof of Proposition 7a and 7b

Proposition 7a (Market Size Effect on Profit, $\boldsymbol{\Pi_\psi \equiv \pi(\psi/A)L}$): Under **A2,** there exists a unique $\psi_0 \in (\underline{\psi}, \psi_c)$ such that $\sigma\left(\frac{\psi_0}{A}\right) = \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right)$ with

$$\frac{d \ln \Pi_\psi}{d \ln L} > 0 \Leftrightarrow \sigma\left(\frac{\psi}{A}\right) < \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) \text{ for } \psi \in \left(\underline{\psi}, \psi_0\right),$$

and

$$\frac{d \ln \Pi_\psi}{d \ln L} < 0 \Leftrightarrow \sigma\left(\frac{\psi}{A}\right) > \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) \text{ for } \psi \in (\psi_0, \psi_c).$$

Proof:

From Proposition 6, $\frac{d \ln A}{d \ln L} = \frac{1}{1-\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right)}$. Hence, using $\mathcal{E}_\pi\left(\frac{\psi}{A}\right) = 1 - \sigma\left(\frac{\psi}{A}\right)$,

$$\frac{d \ln \Pi_\psi}{d \ln L} = 1 + \frac{\partial \ln \pi(\psi/A)}{\partial \ln A}\frac{d \ln A}{d \ln L} = 1 - \mathcal{E}_\pi\left(\frac{\psi}{A}\right)\frac{d \ln A}{d \ln L} = \frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - \sigma(\psi/A)}{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}.$$

Thus,

$$\frac{d \ln \Pi_\psi}{d \ln L} \gtreqless 0 \Leftrightarrow \sigma\left(\frac{\psi}{A}\right) \lesseqgtr \mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right).$$

Since $\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right)$ is the (profit-weighted) average of $\sigma(\psi/A)$ over $\left(\underline{\psi}, \psi_c\right)$ and $\sigma(\psi/A)$ is strictly increasing under A2, there exists a unique $\psi_0 \in (\underline{\psi}, \psi_c)$ such that $\sigma(\psi_0/A) =$

$\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)$, and $\sigma(\psi/A) < \mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)$ for $\psi \in \left(\underline{\psi},\psi_0\right)$ and $\sigma(\psi/A) > \mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)$ for $\psi \in (\psi_0,\psi_c)$. This completes the proof. ■

Proposition 7b (Market Size Effect on Revenue, $\boldsymbol{R_\psi \equiv r(\psi/A)L}$): Under **A2** and the weak **A3,** there exists $\psi_1 > \psi_0$, such that

$$\frac{d\ln R_\psi}{d\ln L} > 0 \text{ for } \psi \in \left(\underline{\psi},\psi_1\right).$$

Furthermore, $\psi_1 \in (\psi_0,\psi_c)$ and

$$\frac{d\ln R_\psi}{d\ln L} < 0 \text{ for } \psi \in (\psi_1,\psi_c),$$

for a sufficiently small F.[46]

Proof:

From Proposition 6, $\frac{d\ln A}{d\ln L} = \frac{1}{1-\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)}$. Hence, using $\mathcal{E}_r\left(\frac{\psi}{A}\right) = \rho\left(\frac{\psi}{A}\right)\left[1-\sigma\left(\frac{\psi}{A}\right)\right]$,

$$\frac{d\ln R_\psi}{d\ln L} = 1 + \frac{\partial \ln r(\psi/A)}{\partial \ln A}\frac{d\ln A}{d\ln L} = 1 - \mathcal{E}_r\left(\frac{\psi}{A}\right)\frac{d\ln A}{d\ln L} = 1 - \rho\left(\frac{\psi}{A}\right)\left[\frac{\sigma(\psi/A)-1}{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}\right].$$

Thus,

$$\frac{d\ln R_\psi}{d\ln L} \gtreqless 0 \Leftrightarrow \rho\left(\frac{\psi}{A}\right) \lesseqgtr \frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma(\psi/A)-1}.$$

Since $\sigma(\psi/A)$ is strictly increasing under A2 and $\rho(\psi/A)$ is non-decreasing under the weak A3, the above inequality changes the sign at most once at $\psi_1 \le \bar{\psi}$, so that

$$\frac{d\ln R_\psi}{d\ln L} > 0 \text{ for all } \psi \in \left(\underline{\psi},\psi_1\right)$$

and $\psi_1 > \psi_0 > \underline{\psi}$, because A2 implies

$$\frac{d\ln R_\psi}{d\ln L} = \frac{d\ln \zeta_\psi}{d\ln L} + \frac{d\ln \Pi_\psi}{d\ln L} > \frac{d\ln \Pi_\psi}{d\ln L} \ge 0 \text{ for all } \psi \in \left(\underline{\psi},\psi_0\right].$$

We now prove $\rho\left(\frac{\psi_c}{A}\right) > \frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma(\psi_c/A)-1}$ and hence $\psi_1 < \psi_c$ for a sufficiently small F by showing

$$\lim_{F\to 0}\rho\left(\frac{\psi_c}{A}\right) = \lim_{\psi_c/A\to\bar{z}}\rho\left(\frac{\psi_c}{A}\right) > \lim_{\psi_c/A\to\bar{z}}\left[\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma(\psi_c/A)-1}\right] = \lim_{F\to 0}\left[\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma(\psi_c/A)-1}\right]$$

We divide the proof of this inequality into the following three cases.

[46]We conjecture whether $\psi_c < \psi_1 \le \bar{\psi}$ and $\frac{d\ln R_\psi}{d\ln L} > 0$ for all $\psi \in \left(\underline{\psi},\psi_c\right)$ for a sufficiently large F.

Case 1: $0 < \lim_{\psi_c/A \to \bar{z}} \rho\left(\frac{\psi_c}{A}\right) < 1$ and $\bar{z} < \infty$. Then, $\lim_{\psi/A \to \bar{z}} \sigma\left(\frac{\psi}{A}\right) = \infty \Rightarrow \lim_{\psi_c/A \to \bar{z}} \left[\frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}{\sigma(\psi_c/A) - 1}\right] = 0$.

Case 2: $0 < \lim_{\psi_c/A \to \bar{z}} \rho\left(\frac{\psi_c}{A}\right) < 1$ and $\bar{z} = \infty$. Then, $\lim_{\psi/A \to \infty} \rho\left(\frac{\psi}{A}\right) < 1 \Leftrightarrow \lim_{z \to \infty} \frac{z\zeta'(z)/\zeta(z)}{\zeta(z) - 1} > 0 \Leftrightarrow$

$\lim_{z \to \infty} \frac{z\zeta'(z)}{\zeta(z)} > 0$, so that $\lim_{\psi/A \to \infty} \sigma\left(\frac{\psi}{A}\right) = \lim_{z \to \infty} \zeta(z) = \zeta(z') \exp\left[\int_{z'}^{\infty} \frac{\xi\zeta'(\xi)}{\zeta(\xi)} \frac{d\xi}{\xi}\right] = \infty \Rightarrow$

$\lim_{\psi_c/A \to \bar{z}} \left[\frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}{\sigma(\psi_c/A) - 1}\right] = 0$.

Case 3: $\lim_{\psi_c/A \to \bar{z}} \rho\left(\frac{\psi_c}{A}\right) = 1$. Then, $\lim_{\psi_c/A \to \bar{z}} \left[\frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}{\sigma(\psi_c/A) - 1}\right] = \lim_{F \to 0} \left[\frac{\mathbb{E}_\sigma\left(\underline{\psi}, \psi_c\right) - 1}{\sigma(\psi_c/A) - 1}\right] < 1$.

This completes the proof. ■

C.4. Proof of Propositions 8a and 8b

Proposition 8a. *Assume $\underline{\psi} = 0$. Consider a shock, which causes a proportional decline in A and ψ_c, so that ψ_c/A remains constant. Then, for any weighting function $w(\psi/A)$,*

iii) the weighted average of any monotonically decreasing (increasing) $f(\psi/A)$,

$$\frac{\int_{\underline{\psi}}^{\psi_c} f(\psi/A) w(\psi/A) dG(\psi)}{\int_{\underline{\psi}}^{\psi_c} w(\psi/A) dG(\psi)}$$

decreases (increases) if $\mathcal{E}_g'(\cdot) < 0$ and increases (decreases) if $\mathcal{E}_g'(\cdot) > 0$.

iv) *the weighted average of any $f(\psi/A)$, monotonic or not, remains constant, if $\mathcal{E}_g'(\cdot) = 0$.*

Proof. Setting $\xi \equiv \psi/\psi_c$, so that $\psi/A = b\xi$, where $\psi_c/A = b > 0$, a constant. By denoting the density function of $\tilde{G}(\xi; \psi_c) \equiv G(\xi\psi_c)/G(\psi_c)$ by $\tilde{g}(\xi; \psi_c)$, the weighted average becomes:

$$\frac{\int_0^1 f(b\xi) w(b\xi) \tilde{g}(\xi; \psi_c) d\xi}{\int_0^1 w(b\xi) \tilde{g}(\xi; \psi_c) d\xi}.$$

i) Since

$$\frac{w(b\xi_1)\tilde{g}(\xi_1; \psi_c) / \int_0^1 w(b\xi')\tilde{g}(\xi'; \psi_c) d\xi'}{w(b\xi_2)\tilde{g}(\xi_2; \psi_c) / \int_0^1 w(b\xi')\tilde{g}(\xi'; \psi_c) d\xi'} = \left[\frac{w(b\xi_1)}{w(b\xi_2)}\right] \frac{\tilde{g}(\xi_1; \psi_c)}{\tilde{g}(\xi_2; \psi_c)}$$

for any ξ_1 and ξ_2, the density transformation

$$\tilde{g}(\xi; \psi_c) \to \frac{w(b\xi)\tilde{g}(\xi; \psi_c)}{\int_0^1 w(b\xi')\tilde{g}(\xi'; \psi_c) d\xi'}$$

preserves the MLR ordering. Hence, from Lemma 2, tougher selection (a lower ψ_c) shifts the distribution given by the density $\frac{w(b\xi)\tilde{g}(\xi;\psi_c)}{\int_0^1 w(b\xi')\tilde{g}(\xi';\psi_c)d\xi'}$ to the right (left) in the MLR ordering if $\mathcal{E}_g'(\cdot) < (>)0$. Since MLR implies FSD, and $f(\cdot)$ is strictly monotone, this completes the proof of part i).

ii) For $\mathcal{E}_g'(\cdot) = 0$, the expression for the weighted average becomes

$$\frac{\int_0^1 f(b\xi)w(b\xi)\xi^{\kappa-1}d\xi}{\int_0^1 w(b\xi)\xi^{\kappa-1}d\xi},$$

which is independent of ψ_c and hence independent of A. This completes the proof. ■

Proposition 8b. *Assume that A2 holds, $\underline{\psi} = 0$, and $\ell(\psi/A)$ is increasing in ψ/A for all $\psi/A \in (0, \psi_c/A)$. Consider a shock, which causes a proportional decline in A and ψ_c, so that ψ_c/A remains constant. Then, the $\ell(\cdot)$-weighted average markup rate decreases if $\mathcal{E}_G'(\cdot) < 0$; remains constant if $\mathcal{E}_G'(\cdot) = 0$; and increases if $\mathcal{E}_G'(\cdot) > 0$.*

Proof. Setting $\xi \equiv \psi/\psi_c$, so that $\psi/A = b\xi$, where $\psi_c/A = b > 0$, a constant. The employment-weighted average markup is given by

$$\frac{\int_0^1 \mu(b\xi)\ell(b\xi)d\tilde{G}(\xi;\psi_c)}{\int_0^1 \ell(b\xi)d\tilde{G}(\xi;\psi_c)} = \frac{\int_0^1 r(b\xi)d\tilde{G}(\xi;\psi_c)}{\int_0^1 \ell(b\xi)d\tilde{G}(\xi;\psi_c)},$$

where $r(\cdot)$ is the revenue function. Since $r'(\cdot) < 0 < \ell'(\cdot)$, $\mathcal{E}_G'(\cdot) < 0(> 0)$ implies from Lemma 2 that the numerator increases (decreases) and the denominator decreases (increases) in response to tougher selection (a lower ψ_c). This completes the proof. ■

C.5. Proof of Propositions 9a, 9b, 9c, and 9d

To prove Proposition 9, we will need two lemmas.

Lemma 9. *Any shock shifting ψ_c and A in the same direction shifts M in the opposite direction.*

Proof. This follows from the RHS of

$$M = \left[\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi)\right]^{-1}$$

being strictly decreasing both in ψ_c and A. This completes the proof. ■

To formulate another lemma, recall first that, from the adding up constraint,

$$\frac{1}{MG(\psi_c)} = \mathcal{J}\left(\frac{\psi_c}{A}, \psi_c\right) \equiv \int_{\underline{\xi}}^{1} r\left(\frac{\psi_c}{A}\xi\right) \frac{dG(\psi_c \xi)}{G(\psi_c)} \equiv \int_{\underline{\xi}}^{1} r\left(\frac{\psi_c}{A}\xi\right) d\tilde{G}(\xi;\psi_c).$$

Using integration by parts,

$$\frac{1}{MG(\psi_c)} = \mathcal{J}\left(\frac{\psi_c}{A}, \psi_c\right) = r\left(\frac{\psi_c}{A}\xi\right)\tilde{G}(\xi;\psi_c)\Big|_{\xi=\underline{\xi}}^{\xi=1} - \frac{\psi_c}{A}\int_{\underline{\xi}}^{1} \tilde{G}(\xi;\psi_c) r'\left(\frac{\psi_c}{A}\xi\right) d\xi$$

$$= r\left(\frac{\psi_c}{A}\right) - \frac{\psi_c}{A}\int_{\underline{\xi}}^{1} \tilde{G}(\xi;\psi_c) r'\left(\frac{\psi_c}{A}\xi\right) d\xi.$$

Lemma 10:
a): $\frac{\partial \mathcal{J}(\psi_c/A,,\psi_c)}{\partial(\psi_c/A)} < 0$;

b): $\mathcal{E}_G'(\psi) \gtreqless 0,\ \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right) \Longrightarrow \frac{\partial \mathcal{J}(\psi_c/A,,\psi_c)}{\partial \psi_c} \lesseqgtr 0.$

Proof.

a): This follows from

$$\mathcal{J}\left(\frac{\psi_c}{A}, \psi_c\right) \equiv \int_{\underline{\xi}}^{1} r\left(\frac{\psi_c}{A}\xi\right) d\tilde{G}(\xi;\psi_c),$$

and $r\left(\frac{\psi_c}{A}\xi\right)$ is strictly decreasing in $\frac{\psi_c}{A}$.

b): Since $\mathcal{J}(a, \psi_c) = r(a) - a\int_{\underline{\xi}}^{1} \tilde{G}(\xi;\psi_c) r'(a\xi) d\xi$, by setting $a = \psi_c/A$,

$$\frac{\partial \mathcal{J}(a,\psi_c)}{\partial \psi_c} = -a\int_{\underline{\xi}}^{1} \frac{\partial \tilde{G}(\xi;\psi_c)}{\partial \psi_c} r'(a\xi) d\xi$$

Thus, from $r'(\cdot) < 0$ and Lemma 2,

$$\mathcal{E}_G'(\psi) \gtreqless 0, \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right) \Longrightarrow \frac{\partial \tilde{G}(\xi;\psi_c)}{\partial \psi_c} \lesseqgtr 0, \forall \xi \in (\underline{\xi}, 1) \ \Longrightarrow \frac{\partial \mathcal{J}\left(\frac{\psi_c}{A}, \psi_c\right)}{\partial \psi_c} \lesseqgtr 0.$$

Proposition 9a (The Effects of F_e on M and $MG(\psi_c)$)

$$\frac{dM}{dF_e} < 0; \qquad \mathcal{E}_G'(\psi) \gtreqless 0, \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right) \Longrightarrow \frac{d[MG(\psi_c)]}{dF_e} \gtreqless 0$$

Proof of Proposition 9a

We first prove the effect of F_e on M. From Corollary 6a of Proposition 6, $d\psi_c/dF_e > 0$ and $dA/dF_e > 0$. Therefore, from Lemma 9, $dM/dF_e < 0$.

We now prove the effect of F_e on $MG(\psi_c)$. From Corollary 6a of Proposition 6, $\frac{d(\psi_c/A)}{dF_e} = 0$ and $\frac{d\psi_c}{dF_e} > 0$, and **Lemma 10**b, $\mathcal{E}_G'(\psi) \gtreqless 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right)$ implies

$$\frac{\partial \mathcal{J}(a,\psi_c)}{\partial a}\frac{d(\psi_c/A)}{dF_e} + \frac{\partial \mathcal{J}(a,\psi_c)}{\partial \psi_c}\frac{d\psi_c}{dF_e} = \frac{\partial \mathcal{J}(a,\psi_c)}{\partial \psi_c}\frac{d\psi_c}{dF_e} \lesseqgtr 0 \Leftrightarrow \frac{d[MG(\psi_c)]}{dF_e} \gtreqless 0$$

This completes the proof. ■

Proposition 9b (The Effects of L on M and $MG(\psi_c)$): *Under* **A2**,

$$\frac{dM}{dL} > 0; \qquad \mathcal{E}_G'(\psi) \le 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \frac{d[MG(\psi_c)]}{dL} > 0.$$

Proof of Proposition 9b

We first prove the effect of L on M. By Corollary 6b of Proposition 6, $d\psi_c/dL < 0$ under A2, and $dA/dL < 0$. Therefore, from Lemma 9, $dM/dL > 0$.

We now prove the effect of L on $MG(\psi_c)$. From **Lemma 10a** and applying **Lemma 10b** for $\mathcal{E}_G'(\psi) \le 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right)$,

$$\frac{\partial \mathcal{J}(a,\psi_c)}{\partial a} < 0; \ \frac{\partial \mathcal{J}(a,\psi_c)}{\partial \psi_c} \ge 0.$$

From Corollary 6b of Proposition 6, $d\psi_c/dL < 0$ under A2, and $d(\psi_c/A)/dL > 0$. Hence,

$$\mathcal{E}_G'(\psi) \le 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \frac{\partial \mathcal{J}(a,\psi_c)}{\partial a}\frac{d(\psi_c/A)}{dL} + \frac{\partial \mathcal{J}(a,\psi_c)}{\partial \psi_c}\frac{d\psi_c}{dL} < 0 \Rightarrow \frac{d[MG(\psi_c)]}{dL} > 0.$$

This completes the proof. ■

Proposition 9c (The Effects of L on M/L, $MG(\psi_c)/L$): *Under* **A2**,

$$G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa} \Rightarrow \frac{d}{dL}\left(\frac{M}{L}\right) > 0; \qquad \mathcal{E}_G'(\psi) \ge 0, \forall\psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \frac{d}{dL}\left(\frac{MG(\psi_c)}{L}\right) < 0.$$

Proof of Proposition 9c

We first prove the effect of L on M/L. From the adding up constraint,

$$\frac{1}{M} = \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi) \Rightarrow \frac{L}{M} = \int_{\underline{\psi}}^{\psi_c} R_{\psi} dG(\psi),$$

hence,

$$\frac{d}{dL}\left(\frac{L}{M}\right) = R_{\psi_c} g(\psi_c)\frac{d\psi_c}{dL} + \int_{\underline{\psi}}^{\psi_c} \frac{dR_\psi}{dL} dG(\psi)$$

$$= r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)\left\{\frac{d\psi_c/\psi_c}{dL/L}\right\} + \int_{\underline{\psi}}^{\psi_c}\left\{\frac{d\ln R_\psi}{d\ln L}\right\} r\left(\frac{\psi}{A}\right) dG(\psi)$$

$$= r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)\left\{\frac{d\psi_c/\psi_c}{dL/L}\right\} + \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi) + \int_{\underline{\psi}}^{\psi_c} \frac{d\ln r\left(\frac{\psi}{A}\right)}{d\ln L} r\left(\frac{\psi}{A}\right) dG(\psi)$$

$$= r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)\left\{\frac{d\psi_c/\psi_c}{dL/L}\right\} + \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi) - \left\{\frac{dA/A}{dL/L}\right\}\int_{\underline{\psi}}^{\psi_c} \frac{d\ln r\left(\frac{\psi}{A}\right)}{d\ln\left(\frac{\psi}{A}\right)} r\left(\frac{\psi}{A}\right) dG(\psi)$$

$$= r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)\left\{\frac{d\psi_c/\psi_c}{dL/L}\right\} + \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi) - \left\{\frac{dA/A}{dL/L}\right\}\int_{\underline{\psi}}^{\psi_c} r'\left(\frac{\psi}{A}\right)\frac{\psi}{A} dG(\psi).$$

Multiplying by $\left[\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right) - 1\right]/\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi)$ and using the expressions for $d\psi_c/\psi_c$ and dA/A from Proposition 6 can verify that the sufficient and necessary conditions are:

$$\frac{d}{dL}\left(\frac{L}{M}\right) \gtreqless 0 \Leftrightarrow$$

$$\left[\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right) - 1}{\sigma\left(\frac{\psi_c}{A}\right) - 1}\right]\frac{r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi)} + \mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right) - 1 + \frac{\int_{\underline{\psi}}^{\psi_c} r'\left(\frac{\psi}{A}\right)\frac{\psi}{A} dG(\psi) - r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right) dG(\psi)} \gtreqless 0.$$

Under $\lim_{\psi\to\underline{\psi}} \psi g(\psi) = 0$, which holds for $G(\psi) = \left(\psi/\overline{\psi}\right)^\kappa$,

$$\int_{\underline{\psi}}^{\psi_c} r'\left(\frac{\psi}{A}\right)\frac{\psi}{A} g(\psi) d\psi = r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c) - \int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)\left[1 + \mathcal{E}_g(\psi)\right] dG(\psi)\,;$$

$$\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right) - 1 = \frac{\int_{\underline{\psi}}^{\psi_c}\left[\sigma\left(\frac{\psi}{A}\right) - 1\right]\pi\left(\frac{\psi}{A}\right) dG(\psi)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right) dG(\psi)} = -\frac{\int_{\underline{\psi}}^{\psi_c}\pi'\left(\frac{\psi}{A}\right)\frac{\psi}{A} dG(\psi)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right) dG(\psi)}$$

$$= \frac{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)\left[1 + \mathcal{E}_g(\psi)\right] dG(\psi) - \pi\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right) dG(\psi)}.$$

Using these expressions to rewrite the 2nd & 3rd terms in the sufficient & necessary condition,

$$\left[\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma\left(\frac{\psi_c}{A}\right)-1}\right]\frac{r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)dG(\psi)}+\frac{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)[1+\mathcal{E}_g(\psi)]dG(\psi)-\pi\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)dG(\psi)}-\frac{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)[1+\mathcal{E}_g(\psi)]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)dG(\psi)}\gtreqless 0.$$

$$\Leftrightarrow\left[\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma\left(\frac{\psi_c}{A}\right)-1}-\frac{\pi\left(\frac{\psi_c}{A}\right)\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)dG(\psi)}{r\left(\frac{\psi_c}{A}\right)\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)dG(\psi)}\right]\frac{r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)dG(\psi)}+\left[\frac{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)[1+\mathcal{E}_g(\psi)]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)dG(\psi)}-\frac{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)[1+\mathcal{E}_g(\psi)]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)dG(\psi)}\right]\gtreqless 0.$$

$$\Leftrightarrow\left[\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma\left(\frac{\psi_c}{A}\right)-1}-\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)}{\sigma\left(\frac{\psi_c}{A}\right)}\right]\frac{r\left(\frac{\psi_c}{A}\right)\psi_c g(\psi_c)}{\int_{\underline{\psi}}^{\psi_c} r\left(\frac{\psi}{A}\right)dG(\psi)}+\left[\frac{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)[1+\mathcal{E}_g(\psi)]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)dG(\psi)}-\frac{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)\sigma\left(\frac{\psi}{A}\right)[1+\mathcal{E}_g(\psi)]dG(\psi)}{\int_{\underline{\psi}}^{\psi_c}\pi\left(\frac{\psi}{A}\right)\sigma\left(\frac{\psi}{A}\right)dG(\psi)}\right]\gtreqless 0.$$

Under $G(\psi)=\left(\psi/\overline{\psi}\right)^{\kappa}$, $\mathcal{E}_g(\psi)$ is constant. Hence, the second bracketed term is zero, and the sufficient and necessary conditions become:

$$\frac{d}{dL}\left(\frac{L}{M}\right)\gtreqless 0\Leftrightarrow\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)-1}{\sigma\left(\frac{\psi_c}{A}\right)-1}\gtreqless\frac{\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)}{\sigma\left(\frac{\psi_c}{A}\right)}\Leftrightarrow\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)\gtreqless\sigma\left(\frac{\psi_c}{A}\right).$$

Under A2,

$$\mathbb{E}_\sigma\left(\underline{\psi},\psi_c\right)<\sigma\left(\frac{\psi_c}{A}\right)\Longrightarrow\frac{d\ln(M/L)}{d\ln L}>0.$$

We now prove the effect of L on $MG(\psi_c)/L$. By setting $a=\psi_c/A$ in the definition of $\mathcal{J}(\psi_c/A\,,\psi_c)$ and the cutoff rule, $\pi(\psi_c/A)L=F$,

$$\frac{L}{MG(\psi_c)}=L\mathcal{J}(a,\psi_c)=F\frac{\mathcal{J}(a,\psi_c)}{\pi(a)}=F\int_{\underline{\xi}}^{1}\frac{r(a\xi)}{\pi(a)}d\tilde{G}(\xi;\psi_c).$$

Since both $\sigma(a\xi)$ and $\pi(a\xi)/\pi(a)$ for any $\xi<1$ are increasing in a under **A2**, so is $r(a\xi)/\pi(a)=\sigma(a\xi)\pi(a\xi)/\pi(a)$ for any $\xi<1$. Thus,

$$\frac{\partial L\mathcal{J}(a,\psi_c)}{\partial a}>0.$$

Furthermore, from **Lemma 10b**,

$$\mathcal{E}_G'(\psi)\geq 0,\qquad\forall\psi\in\left(\underline{\psi},\overline{\psi}\right)\Longrightarrow\frac{\partial\mathcal{J}(a,\psi_c)}{\partial\psi_c}\leq 0$$

From Corollary 6b of Proposition 6, $da/dL=d(\psi_c/A)/dL>0$, and $d\psi_c/dL<0$ under **A2**. Hence,

$$\frac{dL\mathcal{J}(a,\psi_c)}{dL}=\frac{\partial L\mathcal{J}(a,\psi_c)}{\partial a}\frac{d(\psi_c/A)}{dL}+\frac{\partial L\mathcal{J}(a,\psi_c)}{\partial\psi_c}\frac{d\psi_c}{dL}>0\ \Leftrightarrow\frac{d}{dL}\left(\frac{MG(\psi_c)}{L}\right)<0.$$

This completes the proof. ■

Proposition 9d (The Effects of F on M and $MG(\psi_c)$): *If* $\ell'(\cdot) > 0$

$$\frac{dM}{dF} < 0; \qquad \mathcal{E}_G'(\psi) \le 0, \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \frac{d[MG(\psi_c)]}{dF} < 0.$$

Proof of Proposition 9d

We first prove the effect of F on M. From Corollary 6c of Proposition 6, $dA/dF > 0$ and $d\psi_c/dF > 0$. Hence, form Lemma 9, $dM/dF < 0$.

We now prove the effect of F on $MG(\psi_c)$. From **Lemma 10a** and applying **Lemma 10b** for $\mathcal{E}_G'(\psi) \le 0, \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right)$,

$$\frac{\partial \mathcal{J}(a,, \psi_c)}{\partial a} < 0; \; \frac{\partial \mathcal{J}(\psi_c/A\,, \psi_c)}{\partial \psi_c} \ge 0.$$

By Corollary 6c of Proposition 6, $d(\psi_c/A)/dF < 0 < d\psi_c/dF$. Hence,

$$\mathcal{E}_G'(\psi) \le 0, \forall \psi \in \left(\underline{\psi}, \overline{\psi}\right) \Rightarrow \frac{d\mathcal{J}(a, \psi_c)}{dF} = \frac{\partial \mathcal{J}(a,, \psi_c)}{\partial a}\frac{d(\psi_c/A)}{dF} + \frac{\partial \mathcal{J}(a, \psi_c)}{\partial \psi_c}\frac{d\psi_c}{dF} > 0$$
$$\Rightarrow \frac{dMG(\psi_c)}{dF} < 0.$$

This completes the proof. ■

C.6. Proof of Propositions 11a and 11b

To prove Proposition 11, we will need the following lemma.

Lemma 11: *Suppose* $G(\psi) = \left(\psi/\overline{\psi}\right)^{\kappa}$*. Then, the equilibrium conditions can be stated as*

$$\int_{a_j}^{1} r\left(b_j \xi\right) \xi^{\kappa-1} d\xi = a_{j+1}^{-\kappa} \int_{a_{j+1}}^{1} r\left(b_{j+1} \xi\ \right) \xi^{\kappa-1} d\psi\, ;\; a_0 = 0$$

$$L_j \pi\left(b_j\right) = L_{j+1} \pi\left(a_j b_{j+1}\right);\;\; L_J \pi\left(b_J\right) = F.$$

$$\sum_{j=1}^{J} \left(a_2 \ldots a_{j-1}\right)^{-\kappa} \int_{a_{j-1}}^{1} \left[L_j \pi\left(b_j \xi\right) - F\right] \xi^{\kappa-1} d\xi = \left(\frac{\overline{\psi}}{\psi_1}\right)^{\kappa} \frac{F_e}{\kappa},$$

where $a_j \equiv \psi_{j-1}/\psi_j$ *and* $b_j \equiv \psi_j/A_j$*.*

Proof:. First, from the adding-up constraints,

$$\int_{\psi_{j-1}}^{\psi_j} r\left(\frac{\psi}{A_j}\right) \psi^{\kappa-1} d\psi = \int_{\psi_j}^{\psi_{j+1}} r\left(\frac{\psi}{A_{j+1}}\right) \psi^{\kappa-1} d\psi.$$

for $j = 1,2, \dots, J-1$. By setting $\xi \equiv \psi/\psi_j$ in the LHS and $\xi \equiv \psi/\psi_{j+1}$ in the RHS, this can be written as:

$$\int_{\psi_{j-1}/\psi_j}^{1} r\left(\frac{\psi_j}{A_j}\xi\right)\xi^{\kappa-1}d\xi = \left(\frac{\psi_j}{\psi_{j+1}}\right)^{-\kappa}\int_{\psi_j/\psi_{j+1}}^{1} r\left(\frac{\psi_{j+1}}{A_{j+1}}\xi\right)\xi^{\kappa-1}d\psi.$$

Second, the cutoff conditions *for* $j = 1,2, \dots, J-1$ can rewritten as:

$$L_j\pi\left(\frac{\psi_j}{A_j}\right) = L_{j+1}\pi\left(\frac{\psi_j}{A_{j+1}}\right);$$

and

$$L_J\pi\left(\frac{\psi_J}{A_J}\right) = F.$$

Third, the free-entry condition can be written as

$$\sum_{j=1}^{J}\left(\frac{\psi_j}{\psi_1}\right)^{\kappa}\int_{\psi_{j-1}/\psi_j}^{1}\left[L_j\pi\left(\frac{\psi_j}{A_j}\xi\right) - F\right]\xi^{\kappa-1}d\xi = \left(\frac{\overline{\psi}}{\psi_1}\right)^{\kappa}\frac{F_e}{\kappa}.$$

Using $a_j \equiv \psi_{j-1}/\psi_j < 1$ and $b_j \equiv \psi_j/A_j$ for $j = 1,2 \dots, J$, the three conditions can be written as:

$$\int_{a_j}^{1} r(b_j\xi)\xi^{\kappa-1}d\xi = a_{j+1}^{-\kappa}\int_{a_{j+1}}^{1} r(b_{j+1}\xi)\xi^{\kappa-1}d\psi;\ a_0 = 0$$

$$L_j\pi(b_j) = L_{j+1}\pi(a_j b_{j+1});\ \ L_J\pi(b_J) = F.$$

$$\sum_{j=1}^{J}(a_2 \dots a_{j-1})^{-\kappa}\int_{a_{j-1}}^{1}[L_j\pi(b_j\xi) - F]\xi^{\kappa-1}d\xi = \left(\frac{\overline{\psi}}{\psi_1}\right)^{\kappa}\frac{F_e}{\kappa}.$$

This completes the proof. ■

Proposition 11a: *Suppose A2 and* $G(\psi) = (\psi/\overline{\psi})^{\kappa}$. *There exists a sequence,* $L_1 > L_2 > \cdots > L_J > 0$, *such that, in equilibrium, the weighted average of* $f(\psi/A_j)$ *across firms operating at market-j are increasing (decreasing) in* j *even though* $f(\cdot)$ *is increasing (decreasing) and hence* $f(\psi/A_j)$ *is decreasing (increasing) in* j.

Proof: First, consider an equilibrium along which

$$b_j = b = \pi^{-1}\left(\frac{F}{L_J}\right)$$

is constant. Then, the first condition implies that a_j solves the following difference equation, $a_{j+1} = \Phi(a_j)$, defined by:

$$\int_{a_j}^{1} r(b\xi)\xi^{\kappa-1}d\xi \equiv a_{j+1}^{-\kappa}\int_{a_{j+1}}^{1} r(b\xi\ )\xi^{\kappa-1}d\psi.$$

with the initial condition, $a_0 = 0$. The LHS is strictly positive and strictly decreasing in $0 < a_j < 1$ and goes to zero as $a_j \to 1$, while the RHS is positive and strictly decreasing in $0 < a_{j+1} < 1$ and goes to infinity as $a_{j+1} \to 0$ and goes to zero as $a_{j+1} \to 1$. Hence, it has a unique solution, $a_{j+1} = \Phi(a_j)$, which satisfies, for $0 \le a_j < 1, a_j < \Phi(a_j) = a_{j+1} < 1$. Thus, $0 = a_0 < a_1 < \cdots < a_J < 1$. From A2, the second condition is satisfied with

$$\frac{L_j}{L_{j+1}} = \frac{\pi(a_j b)}{\pi(b)} > 1.$$

Furthermore, a_j is monotone increasing in j implies that the weighted average of $f(\psi/A_j) = f(b\,\psi/\psi_j)$,

$$\frac{\int_{a_{j-1}}^{1} f(b\xi)w(b\xi)\xi^{\kappa-1}d\xi}{\int_{a_{j-1}}^{1} w(b\xi)\xi^{\kappa-1}d\xi},$$

is increasing (decreasing) in j if and only if $f(\cdot)$ is increasing (decreasing).

This completes the proof. ■

Proposition 11b: *Suppose* $G(\psi) = (\psi/\overline{\psi})^{\kappa}$. *Then, a change in* F_e *keeps*

iii) *the ratios* $a_j \equiv \psi_{j-1}/\psi_j$ *and* $b_j \equiv \psi_j/A_j$

and

iv) *the weighted average of* $f(\psi/A_j)$ *across firms operating at market-j, for any weighting function* $w(\psi/A_j)$,

unchanged for all $j = 1,2, \ldots, J$.

Proof:

i) The first two equilibrium conditions of Lemma 11 jointly pin down $(a_0, a_1, a_2, \ldots, a_{J-1}; b_1, b_2, \ldots, b_J)$ and hence the LHS of the third condition pins down the RHS. Thus, for all $j = 1,2, \ldots, J$,

$$\frac{d\psi_j}{\psi_j} = \frac{dA_j}{A_j} = \frac{1}{\kappa}\frac{dF_e}{F_e}.$$

ii) Take any firm-specific variable that can be written as a function of ψ/A_j, $f(\psi/A_j)$, for firms operating at market-j, and let $w(\psi/A_j) > 0$ be a weighting function, such as the revenue, profit, or employment within market-j. The weighted average of $f(\psi/A_j)$ for market-j is given by

$$\frac{\int_{\psi_{j-1}}^{\psi_j} f(\psi/A_j)w(\psi/A_j)dG(\psi)}{\int_{\psi_{j-1}}^{\psi_j} w(\psi/A_j)dG(\psi)}.$$

Setting $\xi \equiv \psi/\psi_j$, the weighted average of $f(\psi/A_j)$ across firms operating at market-j becomes:

$$\frac{\int_{\psi_{j-1}/\psi_j}^{1} f\left(\frac{\psi_j}{A_j}\xi\right) w\left(\frac{\psi_j}{A_j}\xi\right)\xi^{\kappa-1}d\xi}{\int_{\psi_{j-1}/\psi_j}^{1} w\left(\frac{\psi_j}{A_j}\xi\right)\xi^{\kappa-1}d\xi} = \frac{\int_{a_j}^{1} f(b_j\xi)w(b_j\xi)\xi^{\kappa-1}d\xi}{\int_{a_j}^{1} w(b_j\xi)\xi^{\kappa-1}d\xi},$$

where $a_j \equiv \psi_{j-1}/\psi_j < 1$ and $b_j \equiv \psi_j/A_j$. Since a_j and b_j remain unchanged in response to a change in F_e by part i), the weighted average of $f(\psi/A_j)$ also remain unchanged in response to a reduction in F_e. This completes the proof. ■

Appendix D: Three Parametric Families of H.S.A.

D.1. Generalized Translog: Matsuyama and Ushchev (2020a, 2022).

For $\sigma > 1$ and $\beta, \eta, \gamma > 0$,

$$s(z) = \gamma\left(1 - \frac{\sigma-1}{\eta}\ln\left(\frac{z}{\beta}\right)\right)^{\eta} = \gamma\left(-\frac{\sigma-1}{\eta}\ln\left(\frac{z}{\bar{z}}\right)\right)^{\eta};\ z < \bar{z} \equiv \beta e^{\frac{\eta}{\sigma-1}}$$

$$\Longrightarrow \zeta(z) = 1 + \frac{\sigma-1}{1 - \frac{\sigma-1}{\eta}\ln\left(\frac{z}{\beta}\right)} = 1 - \frac{\eta}{\ln(z/\bar{z})} > 1,$$

which is strictly increasing in z for all $z \in (0, \bar{z})$, hence satisfying **A2**. In contrast,

$$\frac{z\zeta'(z)}{[\zeta(z)-1]\zeta(z)} = \frac{1}{\eta}\left[1 - \frac{1}{\zeta(z)}\right] = \frac{1}{\eta - \ln(z/\bar{z})}$$

is strictly increasing in z for all $z \in (0, \bar{z})$. Thus, the weak **A3** is violated.[47]

---

[47] Indeed, any H.S.A. satisfying A2 and $\lim_{z\to 0} s(z) = \infty$ violates the weak A3. To see this, under A2, $1 \le \zeta(0) < \zeta(z) < \infty$ for any $\bar{z} > z_0 > z > 0$, hence, $0 < \int_z^{z_0} \frac{\zeta'(\xi)}{\zeta(\xi)} d\xi = \ln\zeta(z_0) - \ln\zeta(z) < \infty$. Moreover, under the weak

Notes:

- CES is the limit case, as $\eta \to \infty$, while holding $\beta > 0$ and $\sigma > 1$ fixed.

$$z < \bar{z} \equiv \beta e^{\frac{\eta}{\sigma-1}} \to \infty$$

$$\zeta(z) = 1 + \frac{\sigma - 1}{1 - \frac{\sigma-1}{\eta}\ln\left(\frac{z}{\beta}\right)} \to \sigma; \;\; s(z) = \gamma\left(1 - \frac{\sigma-1}{\eta}\ln\left(\frac{z}{\beta}\right)\right)^{\eta} \to \gamma\left(\frac{z}{\beta}\right)^{1-\sigma};$$

- Translog is the special case where $\eta = 1$.
- $z = Z\left(\frac{\psi}{A}\right)$ is given as the inverse of $\frac{\eta z}{\eta - \ln(z/\bar{z})} = \frac{\psi}{A}$;
- If $\eta \geq 1$, $\frac{z\zeta'(z)}{\zeta(z)} < \eta z\zeta'(z) = [\zeta(z) - 1]^2$; and employment is globally decreasing in z;
- If $\eta < 1$, employment is hump-shaped with the peak, given by $\eta\zeta(\hat{z}) = 1 \Leftrightarrow \hat{z}/\bar{z} = \frac{\hat{\psi}}{(1-\eta)\bar{z}A} = \exp\left[-\frac{\eta^2}{1-\eta}\right] < 1$, decreasing in η.

D.2. Constant Pass-Through (CoPaTh): Matsuyama and Ushchev (2020a, 2020b)

For $0 < \rho < 1$, $\sigma > 1$, $\beta > 0$, and $\gamma > 0$,

$$s(z) = \gamma\left[\sigma - (\sigma - 1)\left(\frac{z}{\beta}\right)^{\frac{1-\rho}{\rho}}\right]^{\frac{\rho}{1-\rho}} = \gamma\sigma^{\frac{\rho}{1-\rho}}\left[1 - \left(\frac{z}{\bar{z}}\right)^{\frac{1-\rho}{\rho}}\right]^{\frac{\rho}{1-\rho}} \text{ for } z < \bar{z} \equiv \beta\left(\frac{\sigma}{\sigma-1}\right)^{\frac{\rho}{1-\rho}}$$

$$\Rightarrow 1 - \frac{1}{\zeta(z)} = \left(\frac{z}{\bar{z}}\right)^{\frac{1-\rho}{\rho}} < 1 \;\text{ for } z < \bar{z} \equiv \beta\left(\frac{\sigma}{\sigma-1}\right)^{\frac{\rho}{1-\rho}}$$

$$\Rightarrow \mathcal{E}_{1-1/\zeta}(z) = -\mathcal{E}_{\zeta/(\zeta-1)}(z) = \frac{1-\rho}{\rho} > 0.$$

satisfying **A2** and the weak form of A3 (but not the strong form).

Note: CES is the limit case, as $\rho \to 1$, while holding $\beta > 0$ and $\sigma > 1$ fixed:

$$z < \bar{z} \equiv \beta\left(\frac{\sigma}{\sigma-1}\right)^{\frac{\rho}{1-\rho}} \to \infty;$$

$$\zeta(z) = \frac{\sigma}{\sigma - (\sigma-1)\left(\frac{z}{\beta}\right)^{\frac{1-\rho}{\rho}}} \to \sigma;$$

A3, $\theta(z) \equiv \frac{z\zeta'(z)}{[\zeta(z)-1]\zeta(z)} > 0$ is non-increasing because $\theta\big(Z(\psi/A)\big) = \frac{1}{\rho(\psi/A)} - 1$. Thus, $\ln s(z) - \ln s(z_0) = \int_z^{z_0} \frac{\zeta(\xi)-1}{\xi} d\xi = \int_z^{z_0} \frac{1}{\theta(\xi)} \frac{\zeta'(\xi)}{\zeta(\xi)} d\xi \leq \frac{1}{\theta(z_0)} \int_z^{z_0} \frac{\zeta'(\xi)}{\zeta(\xi)} d\xi$, from which $\lim_{z\to 0} s(z) \leq \ln s(z_0) + \frac{1}{\theta(z_0)} \int_0^{z_0} \frac{\zeta'(\xi)}{\zeta(\xi)} d\xi < \infty$.

$$s(z) = \gamma\left[\sigma - (\sigma-1)\left(\frac{z}{\beta}\right)^{\frac{1-\rho}{\rho}}\right]^{\frac{\rho}{1-\rho}} \to \gamma\left(\frac{z}{\beta}\right)^{1-\sigma};$$

because, by applying l'Hôpital's rule for $\Delta = \frac{1-\rho}{\rho}$,

$$\lim_{\rho\nearrow 1}\ln\frac{s(z)}{\gamma} = \lim_{\Delta\searrow 0}\frac{\ln\left[\sigma-(\sigma-1)\left(\frac{z}{\beta}\right)^{\Delta}\right]}{\Delta} = \lim_{\Delta\searrow 0}\frac{(1-\sigma)\left(\frac{z}{\beta}\right)^{\Delta}\ln\left(\frac{z}{\beta}\right)}{\sigma-(\sigma-1)\left(\frac{z}{\beta}\right)^{\Delta}} = (1-\sigma)\ln\left(\frac{z}{\beta}\right).$$

Monopoly Pricing: From the firm's FOC:

$$z_\psi\left[1-\frac{1}{\zeta(z_\psi)}\right] = \frac{\psi}{A}.$$

$$z_\psi \equiv Z\left(\frac{\psi}{A}\right) = (\bar{z})^{1-\rho}\left(\frac{\psi}{A}\right)^{\rho}$$

which features a constant (incomplete) pass-through rate, $0<\rho<1$. Hence, the weak form of **A3** holds, but not the strong form of **A3**. Furthermore,

$$\sigma\left(\frac{\psi}{A}\right) = \zeta\left(Z\left(\frac{\psi}{A}\right)\right) = \frac{1}{1-\left(\frac{\psi}{\bar{z}A}\right)^{1-\rho}} = \frac{1}{1-\left(1-\frac{1}{\sigma}\right)^{\rho}\left(\frac{\psi}{\beta A}\right)^{1-\rho}} > \sigma$$

increasing in ψ/A for $\psi/A < \bar{z}$, while

$$r\left(\frac{\psi}{A}\right) = s\left(Z\left(\frac{\psi}{A}\right)\right) = \gamma\sigma^{\frac{\rho}{1-\rho}}\left[1-\left(\frac{\psi}{\bar{z}A}\right)^{1-\rho}\right]^{\frac{\rho}{1-\rho}} = \gamma\sigma^{\frac{\rho}{1-\rho}}\left[1-\left(1-\frac{1}{\sigma}\right)^{\rho}\left(\frac{\psi}{\beta A}\right)^{1-\rho}\right]^{\frac{\rho}{1-\rho}}$$

$$\pi\left(\frac{\psi}{A}\right) = \frac{r(\psi/A)}{\sigma(\psi/A)} = \gamma\sigma^{\frac{\rho}{1-\rho}}\left[1-\left(\frac{\psi}{\bar{z}A}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}} = \gamma\sigma^{\frac{\rho}{1-\rho}}\left[1-\left(1-\frac{1}{\sigma}\right)^{\rho}\left(\frac{\psi}{\beta A}\right)^{1-\rho}\right]^{\frac{1}{1-\rho}}$$

are decreasing in ψ/A for $\psi/A < \bar{z}$. In contrast,

$$\ell\left(\frac{\psi}{A}\right) = r\left(\frac{\psi}{A}\right) - \pi\left(\frac{\psi}{A}\right) = \gamma\sigma^{\frac{\rho}{1-\rho}}\left(\frac{\psi}{\bar{z}A}\right)^{1-\rho}\left[1-\left(\frac{\psi}{\bar{z}A}\right)^{1-\rho}\right]^{\frac{\rho}{1-\rho}}$$

$$= \gamma\sigma^{\frac{\rho}{1-\rho}}\left(1-\frac{1}{\sigma}\right)^{\rho}\left(\frac{\psi}{\beta A}\right)^{1-\rho}\left[1-\left(1-\frac{1}{\sigma}\right)^{\rho}\left(\frac{\psi}{\beta A}\right)^{1-\rho}\right]^{\frac{\rho}{1-\rho}}$$

increasing in ψ/A for $\psi/A < \hat{\psi}/A \equiv \bar{z}(1-\rho)^{\frac{1}{1-\rho}}$ and decreasing in ψ/A for $\hat{\psi}/A < \psi/A < \bar{z}$. Equivalently, employment is increasing in z for $z < \hat{z} \equiv (\bar{z})^{1-\rho}\left(\hat{\psi}/A\right)^{\rho} = \bar{z}(1-\rho)^{\frac{\rho}{1-\rho}}$ and decreasing in z for $\hat{z} < z < \bar{z}$. Note also that

$$\hat{z}/\bar{z} = (1-\rho)^{\frac{\rho}{1-\rho}} > \hat{\psi}/\bar{z}A = (1-\rho)^{\frac{1}{1-\rho}},$$

which is monotonically decreasing in ρ with $\hat{z}/\bar{z} \to 1$ and $\hat{\psi}/\bar{z}A \to 1$, as $\rho \to 0$, and $\hat{z}/\bar{z} \to 0$ and $\hat{\psi}/\bar{z}A \to 0$, as $\rho \to 1$.

D.3. Power Elasticity of Markup Rate: For $\kappa \geq 0$ and $\lambda > 0$

$$s(z) = \exp\left[\int_{z_0}^{z} \frac{c}{c - \exp\left[-\frac{\kappa\bar{z}^{-\lambda}}{\lambda}\right]\exp\left[\frac{\kappa\xi^{-\lambda}}{\lambda}\right]}\frac{d\xi}{\xi}\right],$$

with either $\bar{z} = \infty$ and $c \leq 1$ or $\bar{z} < \infty$ and $c = 1$. Then,

$$1 - \frac{1}{\zeta(z)} = c\exp\left[\frac{\kappa\bar{z}^{-\lambda}}{\lambda}\right]\exp\left[-\frac{\kappa z^{-\lambda}}{\lambda}\right] < 1$$

$$\Longrightarrow \mathcal{E}_{1-1/\zeta}(z) = -\mathcal{E}_{\zeta/(\zeta-1)}(z) = \kappa z^{-\lambda};$$

satisfying **A2** and the strong **A3** for $\kappa > 0$ and $\lambda > 0$.

CES for $\kappa = 0$; $\bar{z} = \infty$; $c = 1 - \frac{1}{\sigma}$; CoPaTh for $\bar{z} < \infty$; $c = 1$; $\kappa = \frac{1-\rho}{\rho} > 0$, and $\lambda \to 0$**.**

With $z = Z\left(\frac{\psi}{A}\right)$ given implicitly by $c\exp\left[\frac{\kappa\bar{z}^{-\lambda}}{\lambda}\right] z\exp\left[-\frac{\kappa z^{-\lambda}}{\lambda}\right] \equiv \frac{\psi}{A}$,

$$\rho\left(\frac{\psi}{A}\right) = \frac{1}{1+\kappa z^{-\lambda}} \Leftrightarrow \mathcal{E}_{\rho}\left(\frac{\psi}{A}\right) = \frac{\lambda\kappa z^{-\lambda}}{[1+\kappa z^{-\lambda}]^2} > 0.$$

Hence,

$$\frac{\partial^2 \ln\rho\left(\frac{\psi}{A}\right)}{\partial A\partial\psi} \lesseqqgtr 0 \Leftrightarrow \mathcal{E}_{\rho}'\left(\frac{\psi}{A}\right) \gtreqqless 0 \Leftrightarrow \kappa z^{-\lambda} \gtreqqless 1 \Leftrightarrow \frac{\psi}{A} \lesseqqgtr (\kappa)^{\frac{1}{\lambda}} zc\exp\left[\frac{\kappa\bar{z}^{-\lambda}-1}{\lambda}\right].$$

Thus, the pass-through rate is log-submodular among more efficient firms, while log-supermodular among less efficient firms. In particular, if $\bar{z} < (\kappa)^{\frac{1}{\lambda}}$, $\frac{\partial^2 \ln\rho(\psi/A)}{\partial A\partial\psi} < 0$ for all $\psi/A < Z(\psi/A) < \bar{z} < \infty$.

Employment is hump-shaped with the peak at $\hat{z} = Z\left(\frac{\hat{\psi}}{A}\right)$, satisfying $\frac{\hat{z}\zeta'(\hat{z})}{\zeta(\hat{z})} \equiv [\zeta(\hat{z})-1]^2 \Leftrightarrow \rho\left(\frac{\hat{\psi}}{A}\right)\sigma\left(\frac{\hat{\psi}}{A}\right) = 1$. This is given by

$$c\left(1+\frac{\hat{z}^{\lambda}}{\kappa}\right)\exp\left[-\frac{\kappa\hat{z}^{-\lambda}}{\lambda}\right]\exp\left[\frac{\kappa\bar{z}^{-\lambda}}{\lambda}\right] = 1 \Leftrightarrow \left(1+\frac{\hat{z}^{\lambda}}{\kappa}\right)\hat{z} = \frac{\hat{\psi}}{A}.$$

Figures For

Selection and Sorting of Heterogeneous Firms through Competitive Pressures

Kiminori Matsuyama and Philip Ushchev

Date: 2022-03-06, Time: 11:43 AM

Figure 1:

Existence and Uniqueness

Cutoff Rule and Free Entry Condition jointly determine ψ_c and $A = A(\mathbf{p})$ uniquely.

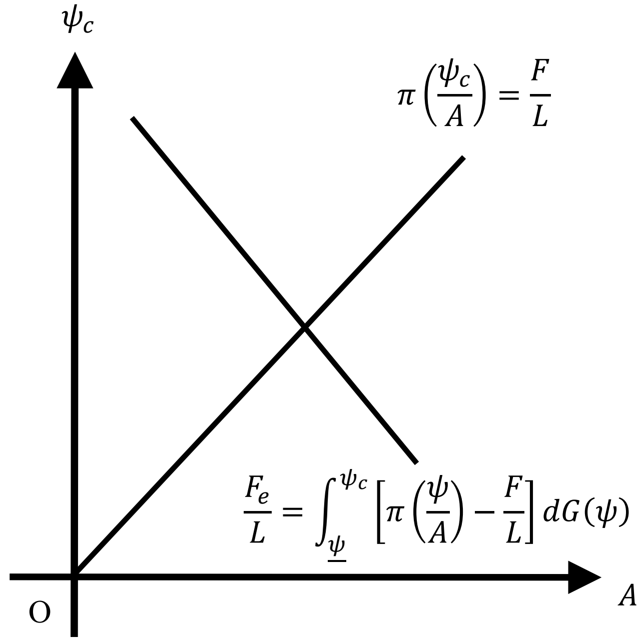


Figure 2: CES Benchmark

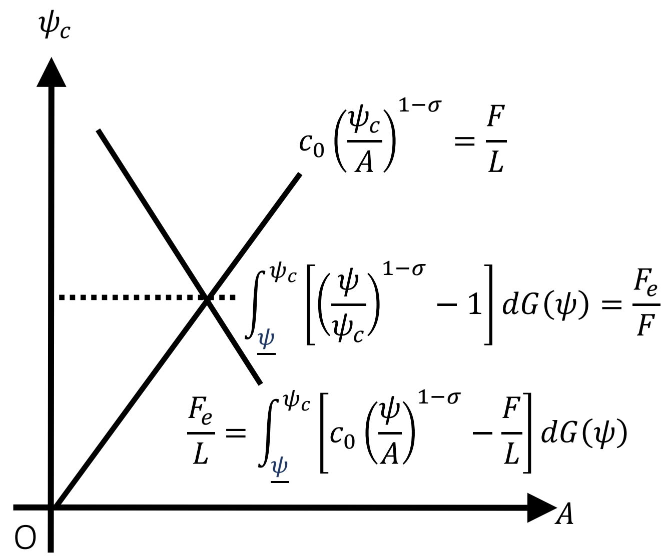


Figure 3: Cross-Sectional Implications of A2 and A3

Figure 3a: Log-Supermodular Profit under A2

Log-profit always downward-sloping and strictly concave under A2. A lower A causes a parallel leftward shift; A higher L causes a parallel upward shift.

[Under the weak A3, the graph of log-revenue has the same properties.]

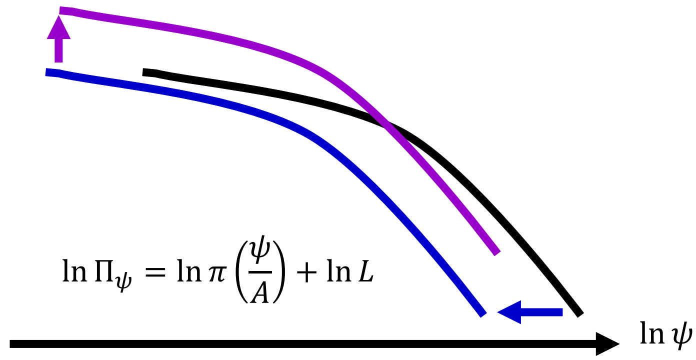


Figure 3b: A2 & A3 and Log-Supermodular Markup Rate

Downward-sloping under A2 and strict(weak)ly convex under strong(weak) A3. A lower A (more competitive pressures) causes a parallel leftward shift.

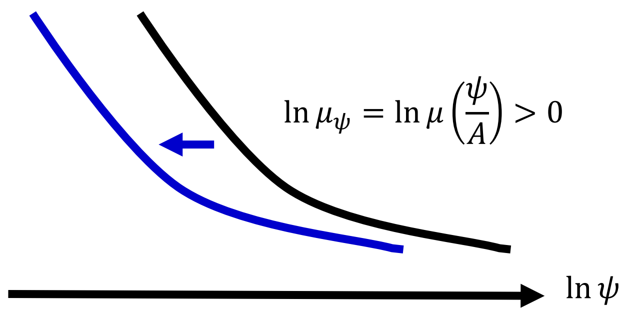


Figure 3c: A2 & the weak A3 and Log-Supermodular Employment

Hump-shaped and strictly concave under A2 and the weak A3. A lower A (more competitive pressures) causes a parallel leftward shift; A higher L (larger market size) causes a parallel upward shift.

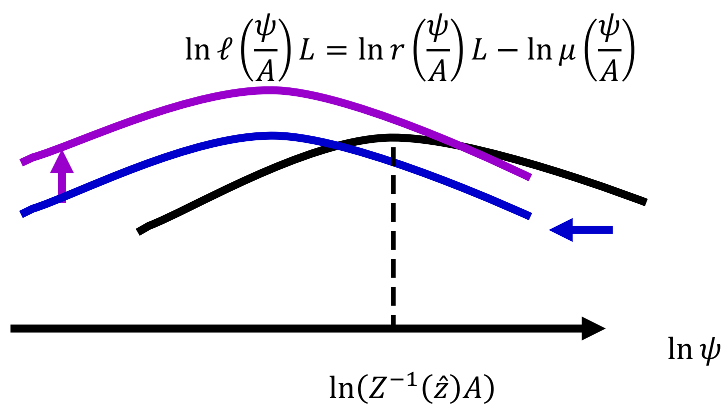


Figure 3d: A2 and strong A3 and Pass-Through Rate

Under A2, $\ln \rho(\psi/A) < 0$;
Under strong A3, strictly increasing;
Under A2 and strong A3, globally strictly convex for a sufficiently small $\bar{z}$:
A lower A (more competitive pressures) causes a parallel leftward shift.

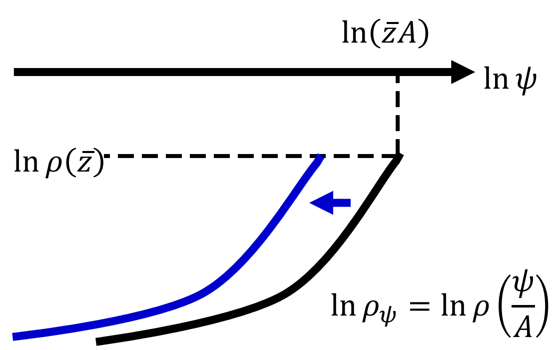


Figure 4: Comparative Statics on ψ_c and A

Figure 4a:

Effects of $F_e \downarrow$

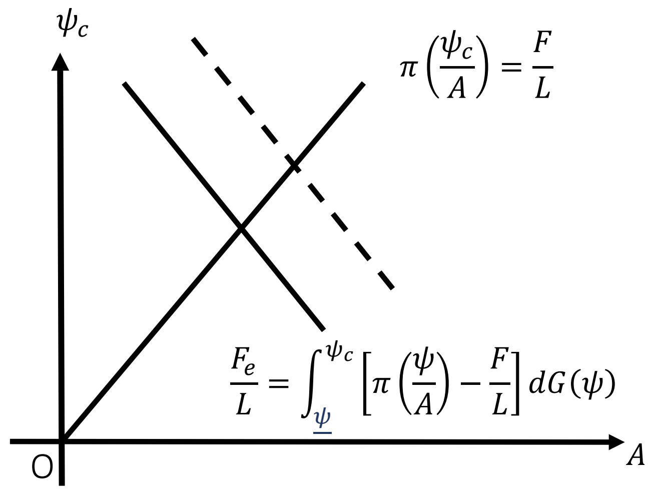


Figure 4b:

Effects of $L \uparrow$
for $\sigma'(\cdot) > 0$, i.e.,
under A2

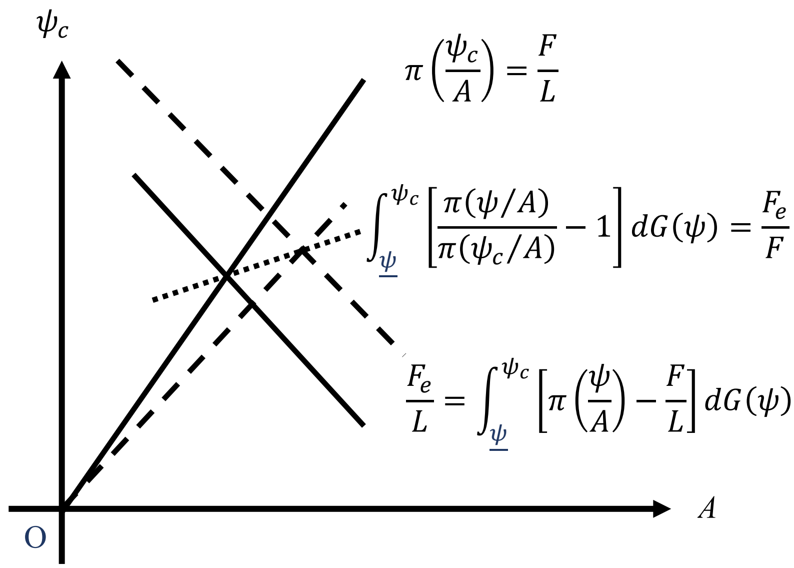
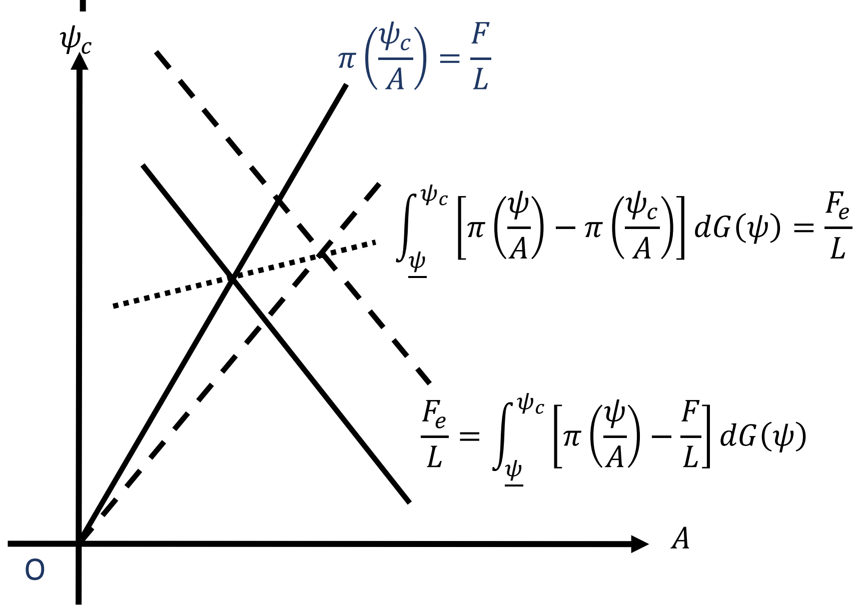


Figure 4c:

Effects of $F \downarrow$

for $\ell'(\cdot) > 0$.



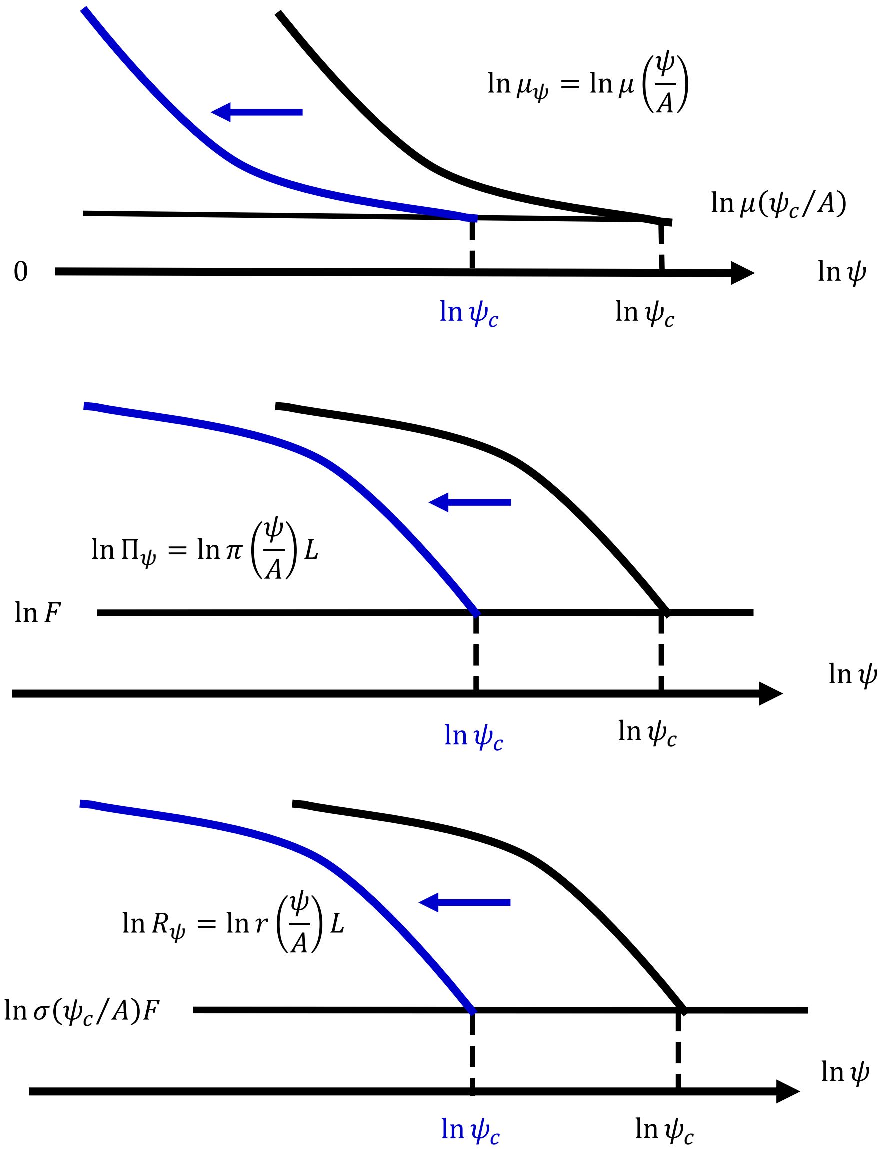


Figure 5a: $F_e \downarrow$ under A2 and the weak A3

From Corollary 6a of Proposition 6, $A \downarrow$, $\psi_c \downarrow$ with ψ_c/A unchanged. Hence, the cutoff firms before the change and those after the change have the same markup rate $\mu(\psi_c/A)$, the same profit $\pi(\psi_c/A)L = F$, and the same revenue, $r(\psi_c/A)L = \sigma(\psi_c/A)\pi(\psi_c/A)L = \sigma(\psi_c/A)F$.

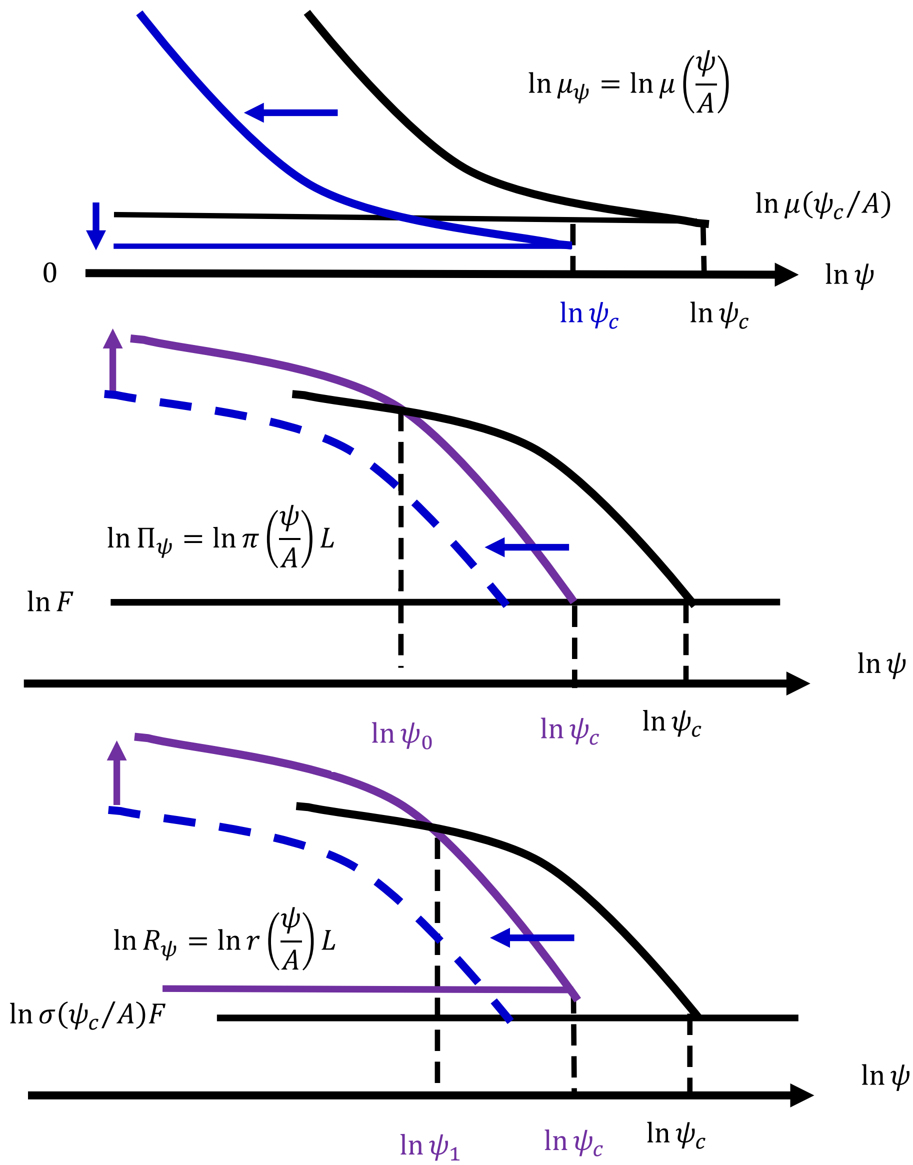


Figure 5b: An increase in L under A2 and the weak A3
From Corollary 6b of Proposition 6, $A \downarrow$, $\psi_c \downarrow$ with $\psi_c/A \uparrow$ and $\sigma(\psi_c/A) \uparrow$. Hence, compared to the cutoff firms before the change, the cutoff firms after the change have a lower markup rate, $\mu(\psi_c/A) \downarrow$, the same profit, $\pi(\psi_c/A)L = F$, and a higher revenue, $r(\psi_c/A)L = \sigma(\psi_c/A)F \uparrow$

From Proposition 7a, the profits are up (down) for $\psi < (>)\psi_0$.
From Proposition 7b, the revenues are up (down) for $\psi < (>)\psi_1$ for a sufficiently small F.

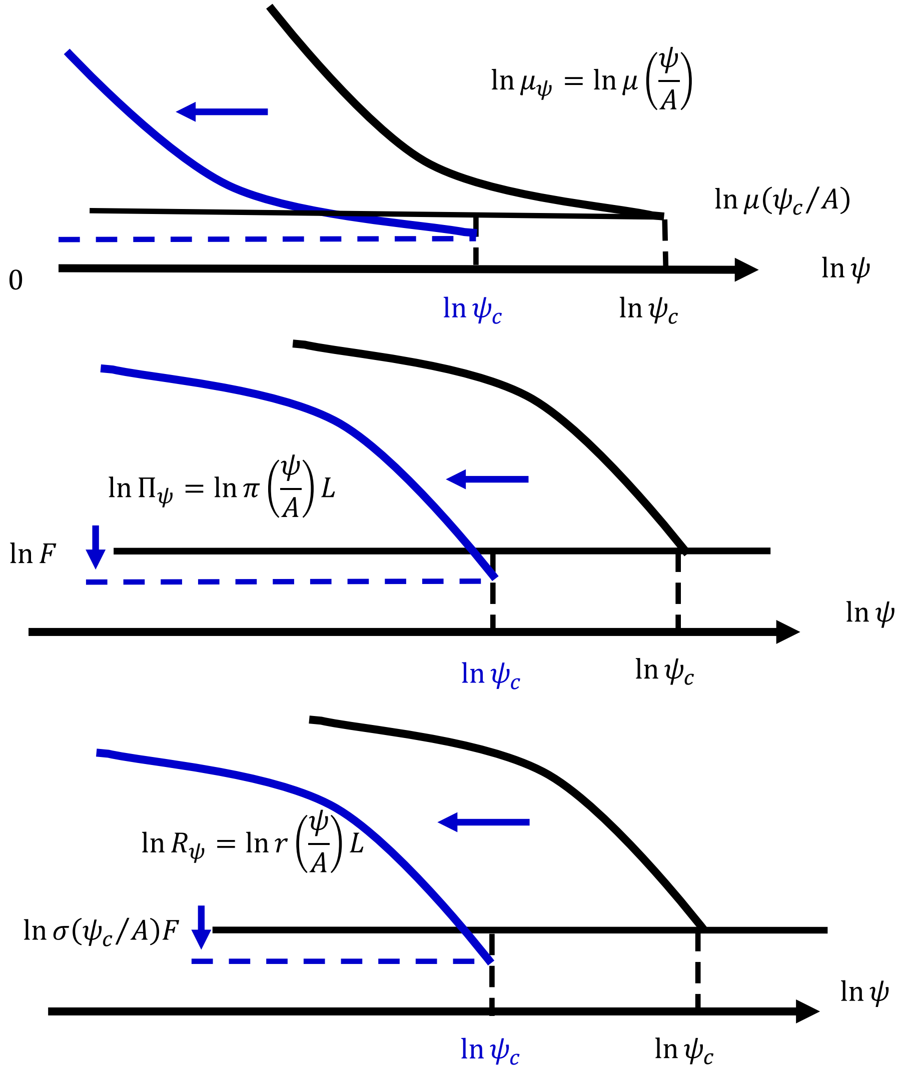


Figure 5c: $F\downarrow$ under A2 and the weak A3 with $\ell'(\cdot) > 0$
From Corollary 6c of Proposition 6, $A\downarrow$, $\psi_c\downarrow$ with $\psi_c/A\uparrow$ and $\sigma(\psi_c/A)\uparrow$. Hence, compared to the cutoff firms before the change, the cutoff firms after the change have a lower markup rate, $\mu(\psi_c/A)\downarrow$, a lower profit, $\pi(\psi_c/A)L = F\downarrow$, and a lower revenue, $r(\psi_c/A)L = \sigma(\psi_c/A)F\downarrow$.

Figure 6: The Limit Case: for $F \rightarrow 0$ with $\bar{z} < \infty$

Figure 6a: $F_e/L \downarrow$ for $F \rightarrow 0$ with $\bar{z} < \infty$

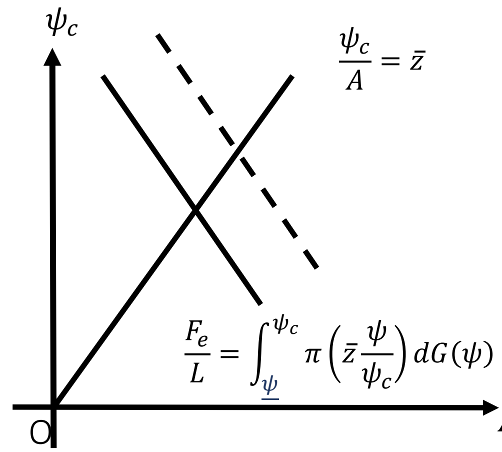


Figure 6b: $F_e/L \downarrow$ for $F \rightarrow 0$ with $\bar{z} < \infty$ under A2 and the weak A3

$A \downarrow, \psi_c \downarrow$ with $\psi_c/A = \bar{z}$ unchanged. Hence, the cutoff firms always (i.e., both before and after the change) have $\mu(\psi_c/A) = 1$ and $\pi(\psi_c/A)L = r(\psi_c/A)L = 0$.

In the middle and bottom panels, Blue indicates the effects of $F_e/L \downarrow$ due to $F_e \downarrow$ and Purple indicates the effects of $F_e/L \downarrow$ due to $L \uparrow$

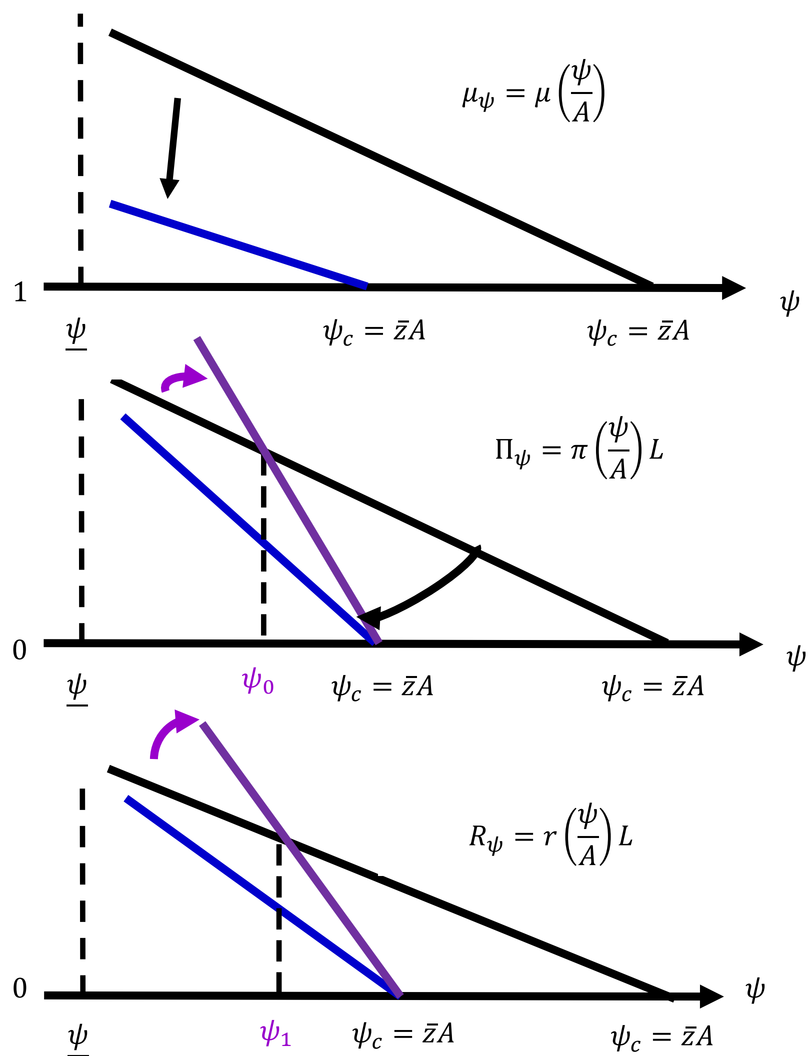


Figure 7:
Logic behind
Sorting

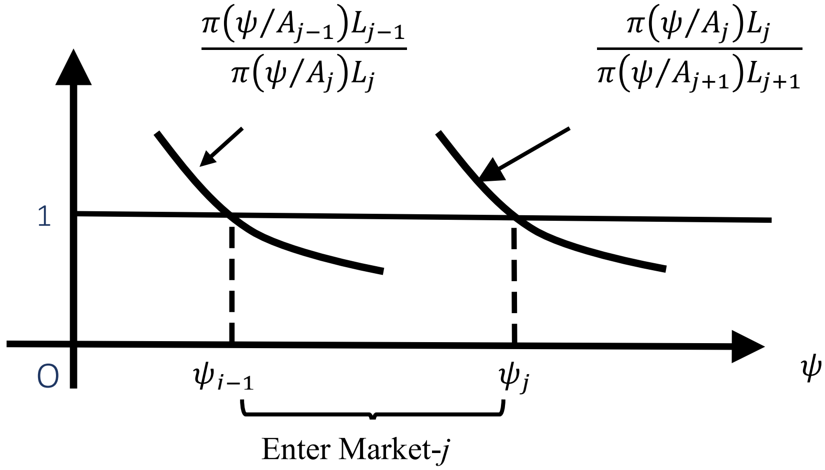


Figure 8: Profit, Revenue, Markup, and Pass-through Schedules across Firms and Markets

Figure 8a:
Profits: Under A2

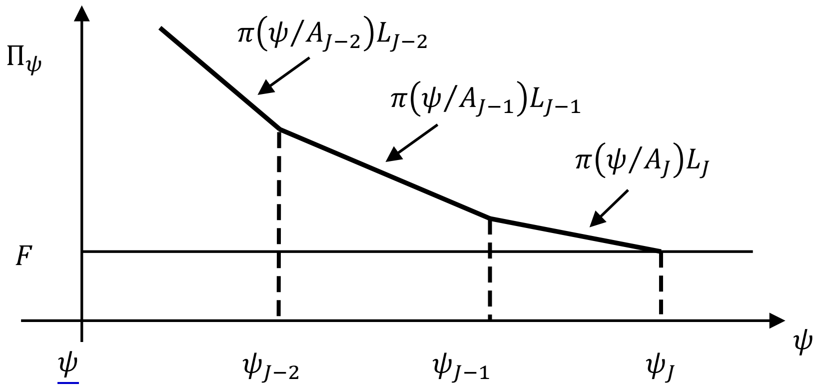


Figure 8b:
Revenues under A2

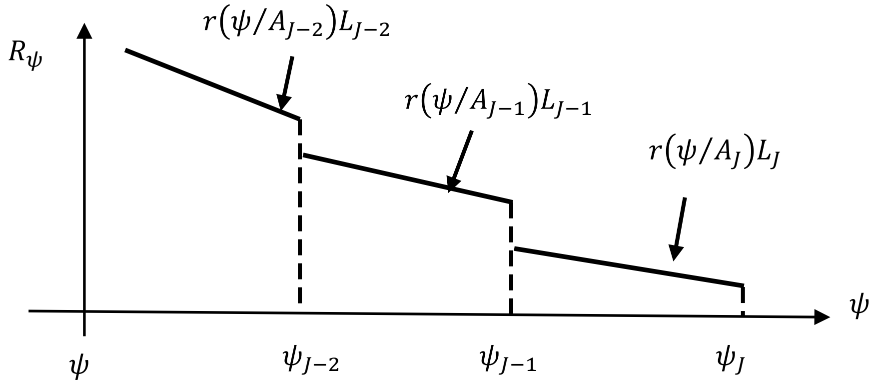


Figure 8c:
Markup rates under A2

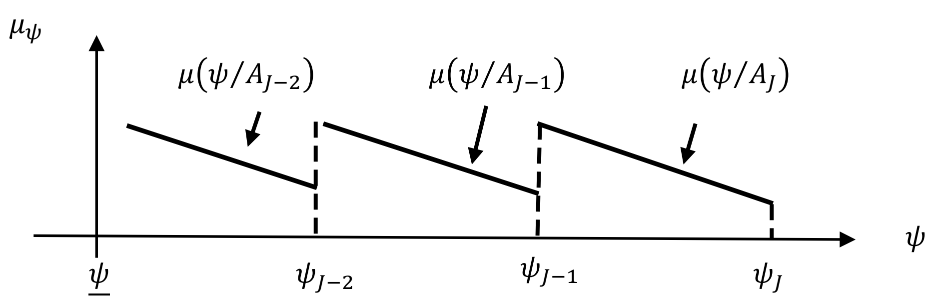


Figure 8d:
Pass-through rates under A2 and the strong A3