One Conjecture on Stationary Firm Distribution under Firm Dynamic Setting

By Shu HU *

I. Introduction and Motivations

This paper provides a conjecture about the competitive equilibrium description of the stationary firm distribution and Pareto distribution under classic firm dynamic settings.

The so-called Power Law Distribution or Pareto distribution is commonly used to study many size distributions in both Economics and non-Economics. For example, Zipf's distribution, the discrete Pareto distribution, was used to describe the distribution of city size (Gabaix, 1999).

Pareto distributions have a unique feature[1] . If a random variable X is Pareto distributed, then its counter cumulative probability function (counter CDF) or survival function would have the form: $\mathbb{P}\{X > x\} = Cx^{-d}$, where $d > 0$ is a parameter and $C > 0$ is a constant.

Champernowne (1953) and Simon (1955) proved that Pareto distributions arise naturally if the time-series behavior of the size variable satisfies the Gibrat's law (1931): this variable's expected rate of growth nor the variance of its growth rate does not depend on its current position. Nevertheless, Gibrat's law often does not hold. Córdoba (2008) showed that whenever the size variable follows a unique Markov process[2], it satisfies the Gibrat's law.

Firm size distributions are significant to theoretical development and practical application. For example, some distortions can misallocate resources, and firm size distribution can help us understand the effect of such distortions on productivity. Using a simple model which aggregates firms with different sizes, Hsieh and Klenow (2008) documented empirically that such misallocation takes account for a significant proportion of the difference in aggregate productivity between the US, China, and India.

The size distribution of firms has been shown to follow a striking pattern. Using U.S. data, Axtell (2001) found that this distribution approximately follows a Pareto distribution. Before Axtell, the fit of firm size distributions has attracted many economists' attention. Dating back to Gibrat (1931)'s seminal work, they

* HU: Research School of Economics, Australian National University. Email: shu.hu@anu.edu.au
Grateful thanks are due to John Stachurski for his ideas, instructions and supervision. Thank Simon Grant and Ronald Stauber for their meaningful comments.

[1] We discuss it more in section 4.B.2

[2] We introduce this Markov process in Assumption 1 and Proposition 3 of section 3.B.

found that the industrial firms follow the lognormal distribution in many countries, such as France (Gibrat, 1931) and UK (Florence, 1953). These approximations before Axtell were problematic because their data only cover small firms, but ignored large firms.

In addition to fitting the data, many researchers have also tried to provide explanations relating the shape of the observed size distribution to stochastic models of firm entry, random growth, and exit with purely idiosyncratic shocks. Hart and Praise (1956), Simon and Bonini (1958), Adelman (1958), Steindl (1965), Ijiri and Simon (1964) were successful in fitting empirical models from the data by specifying unique processes for firm size. Lucas and Prescott (1971) firstly built a stochastic competitive equilibrium model [3] of industry evolution by considering price, aggregate output, and price. Brock (1972) firstly introduced entry and exit into the dynamic model. The so-called idiosyncratic productivity shocks[4] were firstly introduced to the firm exit-entry dynamic model by Jovanovic (1982). Pakes and Ericson (1998) built another dynamic model whose firm's production is affected by investments with uncertain outcomes.

Both Jovanovic's and Pakes's models focused on firm-level dynamics, Hopenhayn1992 firstly developed the concept of a stationary equilibrium, which provides a more tractable but more straightforward framework to tackle firm dynamic questions from the aggregate aspect. One of the critical assumptions in Hopenhayn's model is that firm productivity is stationary, which guarantees a stationary industry equilibrium but not a resulting size distribution to look like a Pareto distribution. Many papers have focused on aggregate fluctuations in the firm dynamics framework Campbell and Fisher (2004), Lee and Mukoyama (2008, 2015), Clementi and Palazzo (2016) and Bilbiie et al. (2012)). The time-varying nature of aggregate volatility and its link with the cross-sectional distribution of firms in the Hopenhayn setting has been explicitly studied by Bloom (2018). Carvalho and Grassi (2019) firstly introduced the micro-origins of aggregate fluctuations to Hopenhayn's setting. One of the key assumptions in Carvalho and Grassi's model is that both the productivity of incumbent firms and entrant firms follows a possibly different Pareto distribution across time. This assumption guarantees a stationary productivity distribution and therefore, a stationery firm size distribution[5] to look a Pareto in the upper tail. However, empirical evidence (Axtell, 2001) supports that incumbent firms follow a Pareto distribution, but there is no empirical evidence indicating whether the entrant's firm size or productivity is also Pareto distributed.

With the tools from the finite Markov chain[6], this paper studies the competitive equilibrium description of firm distributions and Pareto distribution in Hopen-

[3] In this model, firms have identical size and in the limit there is no entry and exit

[4] The idiosyncratic productivity shocks are shocks drawn from a distribution with unknown mean but known variance. The mean is specific to a firm and realizations are independent across firms.

[5] Because stationary firm size is pinned down by stationary firm-level productivity. We can reach this conclusion from the Model's setting. Discuss it more in section 3.D.

[6] This paper gives a brief introduction to Markovian process in section 2.

hayn's dynamic model (1992) with a finite but possibly large number of firms (Carvalho and Grassi, 2019). We put forward a conjecture about the stationary firm distribution by relaxing Carvalho and Grassi's distribution assumption on entrants. We also provide a uniqueness proof of and implement some simulations about the upper tail of the stationary firm distribution in the conjecture.

The rest of the report is structured as follows: Section 2 presents definitions and theorems about the Markov process. Section 3 describes the firm dynamic model and one of its theoretical results. Section 4 provides a conjecture, proof, and simulations. Section 5 introduces the plan for the next step of the research. All formal proofs and figures are deferred to the Appendix.

II. Markov Chain and Stationary Distribution

In this section, we briefly introduce the definitions and theorems about the Markov process, which provides proper tools for analyzing the firm dynamic model in section 3. These definitions and theorems follow Stachurski (2009) closely[7].

A. Stochastic Markov/Chain

Markov (stochastic) Matrix An $n \times n$ square matrix $P = (p_{i,j})_{n\times n}$ is said to be a Markov matrix (or stochastic matrix) if (i) Each element of P is nonnegative ($p_{i,j} \geq 0$); (ii) Each row of P sums to 1, i.e. $\sum_{j=1}^{n} p_{i,j} = 1, \forall i \in \{1, ...n\}$.

State Space and State Values Let S be a finite set with n elements $\{x_1, \cdots, x_n\}$, where the set S is called the state space and $x_1, \cdots, x_n$ are the state values.

Markov Chain A Markov chain $\{X_t\}$ on S is a sequence of random variables on S (from probability sample space Ω) that have the Markov property, that is, for any data t and any state $y \in S$, $Pr\{X_{t+1} = y|X_t\} = Pr\{X_{t+1}|X_t, X_{t-1}, \dots\}$

Dynamics of a Markov Chain The dynamics of a Markov chain $\{X_t\}$ are fully determined by the set of values

$$P(x, y) = Pr\{X_{t+1} = y|X_t = x\}(x, y \in S) \tag{1}$$

where P^i is the i power of P and $P^i(x, y)$ is the (x, y) element of the Markov matrix P.

B. Irreducibility and Aperiodicity

Two States Communicate Let P be a fixed stochastic matrix. Two states x and y communicate with each other if there exists a positive integer n such that

$$P^n(x, y) > 0 \; and \; P^n(y, x) > 0 \tag{2}$$

[7]For more details, see `https://python.quantecon.org/finite_markov.html`

where $P^i(x, y)$ is the (x, y)-th element of the matrix P^i.

Irreducibility A stochastic matrix P is called irreducible if all its states communicate with each other.

Period The period of a state x is the greatest common divisor of the set of integers

$$D(x) = \{j \geq 1 | P^j(x, x) > 0\} \tag{3}$$

Aperiodicity A stochastic matrix is aperiodic if the period of its every state is 1.

C. Stationary Distribution

Distributions of Markov Chain/Matrix Suppose that $\{X_t\}$ is a Markov chain with stochastic matrix P and the distribution of X_t is known to be ψ_t. Then the distribution of X_{t+m} is updated by

$$\psi_{t+m} = \psi_t P^m \tag{4}$$

where P^m is the m-th power of P.

Stationary Distribution A distribution ψ^* on S is stationary for P if

$$\psi^* = \psi^* P \tag{5}$$

For the distribution of Markov matrix, we have the following two propositions:

Proposition 1.
Every stochastic matrix P has at least one stationary distribution.

Proof. This proposition can be proved by Brouwer Fixed-Point Theorem. For a detailed proof, see Stachurski (2009)'s theorem 4.3.5. ■

Proposition 2.
If a Markov matrix P is both aperiodic and irreducible, then

- *P has exactly one stationary distribution ψ^*;*
- *For any initial distribution ψ_0, we have as $t \to \infty$*

$$\|\psi_0 P^t - \psi^*\| \to 0 \tag{6}$$

Proof. See Häggström (2002). ■

III. Standard Firm Dynamic Model with Entry and Exit

In this section, we introduce a standard firm dynamics model (Hopenhayn, 1992) with a finite but possibly large number of firms (Carvalho and Grassi, 2019)[8].

The basic setup, assumptions, and the law of motion follow Hopenhayn (1992), Carvalho and Grassi (2019) closely. Our model closes with a detailed description of this economy's steady-state equilibrium from Carvalho and Grassi (2019).

A. Basic Setups

1. Incumbents and Entrants

There are infinite periods, that is, $t \in \{0, 1, 2, ...\}$.

There is an exogenously given, finite but large number of incumbent firms, $N_t = N$ and potential entrant firms, $M_t = M$ in period t.

Every firm varies in its productivity levels, which is assumed to follow a discrete Markovian process over the idiosyncratic state space $\Phi = \{\varphi^1, \ldots, \varphi^S\}$, where $\frac{\varphi^{s+1}}{\varphi^s} = \varphi$. A firm is in state s when its idiosyncratic productivity is equal to φ^s. Each incumbent's productivity level is assumed to follow a monotone Markov chain with a transition matrix P. Every entrant has access to a signal about their potential productivity next period. The entrants' signals are distributed by a discrete distribution supported over Φ, $G = (G(q))_{q\in\{1,\ldots,S\}}$ that is, $0 < G(q) \leq 1, \forall q$ and $\sum_{q=1}^{S} G(q) = 1$. Given the current period's idiosyncratic productivity φ^s, the conditional distribution of the next period's idiosyncratic productivity $\varphi^{s'}$ is $F(\cdot|\varphi^s)$.

The distribution of firms, described by a $(S \times 1)$ vector μ_t, gives the number of firms at each productivity level s at time t. We assume that $N_t = N = \sum_{s=1}^{S} \mu_{t,s}$ and $M_t = M = \sum_{s=1}^{s^*-1} \mu_{t,s}$, which restricts the relative sizes of incumbents and entrants such that $M \leq N$[9]. There are $MG(q)$ potential entrants for each signal level φ^q.

2. Static Profits and Firm Size

For the following analysis, we abstract from explicit time t notation and assume that μ is the given aggregate state and φ^s is a given idiosyncratic productivity level.

Both incumbents and potential entrants can use labor, denoted by n, only to produce a unique good, denoted by y, in a perfectly competitive market with a decreasing returns-to-scale technology, described by α, which is described by the

[8] In this setting, Carvalho and Grassi (2019) characterize the law of motion for any finite number of firms.

[9] This is because we assume that $s^* \geq S$, which we discuss in section 3.A.3. Further discussion about the relative size of incumbents and entrants is required.

production function $y = \varphi^s n^\alpha$. Incumbents have to pay the operating cost c_f in per unit of output if they continue to stay at every period. Potential entrants have to pay an entry cost c_e once if they want to enter at next period.

The incumbent/entrant maximises the static profit:

$$\pi^*(\mu, \varphi^s) = \max_n \{\varphi^s n^\alpha - w(\mu)n - c_f\} \tag{7}$$

where $0 < \alpha < 1$ is the decreasing returns-to-scale technology, n is the labor input, $w(\mu)$ is the wage for a given aggregate state μ, c_f is the operating cost to be paid in unit of output every period. For entrants, they also need to pay the entry cost c_e only once in the period before they enter.

By solving for optimal n^* from the first-order condition of (7), we obtain the output of a firm by plugging n^* into the production $y = \varphi^s n^\alpha$:

$$y(\mu, \varphi^s) = (\varphi^s)^{\frac{1}{1-\alpha}} (\frac{\alpha}{w(\mu)})^{\frac{\alpha}{1-\alpha}} \tag{8}$$

It is also easy to show that π^* is increasing in φ^s and decreasing in w for a given aggregate state μ [10].

Note that the size of a firm refers to its output level if not specified.

3. Profits and Entry/Exit Problems

Incumbents and potential entrants maximise expected discounted profits with perfect foresight on future prices.

Given μ and φ^s, the present discounted value of an incumbent is the Bellman equation

$$V(\mu, \varphi^s) = \pi^*(\mu, \varphi^s) + \max\{0, V^i(\mu, \varphi^s)\} \tag{9}$$

where $V^i(\mu, \varphi^s) = \beta \int_{\mu' \in \Lambda} \sum_{\varphi^{s'} \in \Phi} V(\mu', \varphi^{s'}) F(\varphi^{s'}|\varphi^s) \Gamma(d\mu'|\mu)$ is the present discounted future profit of an incumbent; β is the discount factor; $\Gamma(\cdot|\mu)$ is the conditional distribution of μ', tomorrow's aggregate state; $F(\cdot|\varphi^s)$ is the conditional distribution of tomorrow's idiosyncratic productivity for a given today's idiosyncratic productivity of the incumbent; Λ is the set of $(S \times 1)$-vectors whose elements are non-negative.

Given μ and φ^q, the present discounted value of a potential entrant is

$$V^e(\mu, \varphi^q) = \beta \int_{\mu' \in \Lambda} \sum_{\varphi^{q'} \in \Phi} V(\mu', \varphi^{q'}) F(\varphi^{q'}|\varphi^q) \Gamma(d\mu'|\mu) \tag{10}$$

Given μ, the endogenous exit/enter decision for incumbents/entrants is defined

[10] By plugging n^* into (7), we get the value function. By Envelope Theorem, we then can solve for $\frac{\partial \pi^*}{\partial \varphi^s} > 0$ and $\frac{\partial \pi^*}{\partial w} < 0$

by a unique threshold[11] level of idiosyncratic productivity, $\varphi^{s^*(\mu)}/\varphi^{e^*(\mu)}$.

An incumbent exits at the end of this period if $V^i(\mu, \varphi^s) < 0$: for $\varphi^s \geq \varphi^{s^*(\mu)}$, the incumbent continues to operate next period; for $\varphi^s < \varphi^{s^*(\mu)}$, the incumbent decides to exit the next period.

A prospect entrant enters the next period if $V^e(\mu, \varphi^q) \geq c_e$: for $\varphi^q \geq \varphi^{e^*(\mu)}$, the potential entrant enters next period; for $\varphi^q < \varphi^{e^*(\mu)}$, the potential entrant decides not to enter the next period.

For simplification, we assume that $c_e = 0$ or equivalently $\varphi^{e^*(\mu)} = \varphi^{s^*(\mu)}$.

4. Timing of decisions

The timing of decisions for an incumbent is:

- The incumbent first draws its idiosyncratic productivity φ^s from a conditional distribution $F(\cdot|\varphi^s)$[12] at the beginning of the period, pays the operating cost c_f, and hires labor to produce.
- Then it decides whether to exit at the end of the period or to continue as an incumbent the next period.

The timing of decisions for a potential entrant is:

- The entrant first has access to a signal $G(q)$ draw from the distribution G about their potential productivity next period and decides whether to enter at the end of the period or to continue as an entrant in the next period.
- If it decides to enter, then it has to pay an entry cost c_e and hire labor to produce the next period.

5. Labor Market and Equilibrium Wage

Assume that the aggregate labor supply at a given wage w is $L^s(w) = Mw^{\gamma}$[13]. The aggregate labor demand can be derived by aggregating all individual outputs (8) to get aggregate output $Y_t = A_t^{1-\alpha}(L_t^d)^{\alpha}$:

(11)

$$Y_t = A_t^{1-\alpha}(L_t^d)^{\alpha} = \sum_{i=1}^{N_t} y(\mu_{s,t}, \varphi^{s_{i,t}}) = \sum_{i=1}^{N_t}(\varphi^{s_{i,t}})^{\frac{1}{1-\alpha}}(\frac{\alpha}{w_t})^{\frac{\alpha}{1-\alpha}} = (\frac{\alpha}{w_t})^{\frac{\alpha}{1-\alpha}}\sum_{i=1}^{N_t}(\varphi^{s_{i,t}})^{\frac{1}{1-\alpha}}$$

Let $A_t = \sum_{i=1}^{N_t}(\varphi^{s_{i,t}})^{\frac{1}{1-\alpha}} = \sum_{s=1}^{S}\mu_{s,t}(\varphi^s)^{\frac{1}{1-\alpha}} = B'\mu_t$, (11) becomes

$$A_t^{1-\alpha}(L_t^d)^{\alpha} = A_t(\frac{\alpha}{w_t})^{\frac{\alpha}{1-\alpha}} \tag{12}$$

[11] Since the static profit π^* is increasing in the idiosyncratic productivity level φ^s, there is a unique index $s^*(\mu)$ for each aggregate state μ such that $V^i(\mu, \varphi^{s^*}) = 0$ for incumbents. Similarly, there is also a unique index $e^*(\mu)$ for each aggregate state μ such that $V^e(\mu, \varphi^{e^*}) = c_e$ for entrants.

[12] $F(\cdot|\varphi^s)$ is determined by the Markov process P in Assumption 1.

[13] The key point here is to assume that for a given wage, the labor supply function should be a linear function of M, the number of potential entrants. For detail reasons, see Carvalho and Grassi (2019).

Simplifying (12) yields the aggregate labor demand

$$L_t^d = (\frac{\alpha A^{1-\alpha}}{w_t})^{\frac{1}{1-\alpha}} \tag{13}$$

The market-clearing condition that equates labor supply and labor demand, i.e., $L^s(w_t) = L^d(w_t)$ gives us the equilibrium wage

$$w_t = (\alpha^{\frac{1}{1-\alpha}} \frac{B'\mu_t}{M})^{\frac{1-\alpha}{\gamma(1-\alpha)+1}} \tag{14}$$

Note that the wage w_t and aggregate output Y_t is fully pinned down by the distribution μ_t.

B. *Assumptions*

Assumption 1. *Incumbents' firm-level productivity evolves as a Markov Chain on the state space $\Phi = \{\varphi^s\}_{s=1,\cdots,S}$ with transition matrix*

$$P = \begin{pmatrix} a+b & c & 0 & \cdots & \cdots & 0 & 0 \\ a & b & c & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a & b & c \\ 0 & 0 & 0 & \cdots & 0 & a & b+c \end{pmatrix} \tag{15}$$

where

- $0 < a < 1$ *is the probability of productivity-improving;*
- $0 < b < 1$ *is the probability of remaining at the same productivity level;*
- $0 < c < 1$ *is the probability of productivity declining;*
- $a + b + c = 1$.

This Markovian process has been first introduced by Champernowne (1953) and Simon (1955) and studied extensively in Córdoba (2008). For completeness, we introduce the properties of the Markovian process in Assumption 1 as follows:

Proposition 3. *For a given firm i at time t with productivity level $\varphi^{s_{i,t}}$ with $s_{i,t} \neq 1, S$ that follows the Markovian process in Assumption 1, then we have the following:*

- *The conditonal expected growth rate and conditional variance of firm-level productivity are given by*

$$\mathbb{E}[\frac{\varphi^{s_{i,t+1}} - \varphi^{s_{i,t}}}{\varphi^{s_{i,t}}} | \varphi^{s_{i,t}}] = a(\varphi^{-1} - 1) + c(\varphi - 1) \tag{16}$$

$$\mathbb{V}ar[\frac{\varphi^{s_{i,t+1}} - \varphi^{s_{i,t}}}{\varphi^{s_{i,t}}}|\varphi^{s_{i,t}}] = \sigma_e^2 \tag{17}$$

where σ_e^2 is a constant. Both the conditions expected growth rate and the conditional variance are independent of i's productivity level, $\varphi^{s_{i,t}}$.

- *As $t \to \infty$, the probability of firm i having productivity level φ^s is*

$$\mathbb{P}(\varphi^{s,t} = \varphi^s) \to_{t\to\infty} K(\varphi^s)^{-\delta} \tag{18}$$

where $\delta = \frac{\log(a/c)}{\log \varphi}$ and K is a normalization constant.

- *Therefore, the stationary distribution of the Markovian Process in Assumption 1 is Pareto with the tail index $\delta = \frac{\log(a/c)}{\log \varphi}$.*

Proof. See Córdoba (2008)'s Proposition 2. ■

This property shows two essential features of the Markovian process in Assumption 1. The first is that whenever firm-level productivity follows this process, its conditional expected growth rate and its conditional variance are independent of its current level. The Markov chain/matrix in Assumption 1 provides an excellent way to obtain Gibrat's law on a discrete state space (Cordoba 2008). The second is that if the firm-level productivity follows this Markovian process, its stationary distribution is Pareto distributed with tail index $\delta = \frac{\log(a/c)}{\log \varphi}$.

Assumption 2. *Assume that $\varphi^S = ZN^{\frac{1}{\delta}}$, for some constant Z.*

From Newman (2005), the expectation of the maximum value of a sample N of random variables drawn from a power-law distribution with tail index δ is proportional to $N^{\frac{1}{\delta}}$.

This assumption restricts the rate at which the maximum level of productivity scales with the number of firms. Under Assumption 2, for any sample of size N following the Markovian process in Assumption 1, the stationary distribution of this sample is Pareto distributed with a constant tail index δ.

C. Law of Motion

Given S productivity levels and an exogenous finite number of incumbents N, making optimal employment and production decisions with entry and exit decisions, the following proposition holds:

Proposition 4. *Without uncertainty, the aggregate state μ_t follows an evolution*

$$\mu_{t+1} = \tilde{P}_t\mu_t + MG \tag{19}$$

where $\tilde{P}$ is the transition matrix P with first $s^*(\mu_t) - 1$ rows are replaced with 0s.

Proof. For a detailed proof, see Hopenhayn (1992). ■

The first term of (19) gives the contribution of old incumbents to μ_{t+1}, and its second term gives the contribution of new entrants to μ_{t+1}.

In order to use the tools of Markov Chain to analyze this model, we rewrite the law of motion as follows:

Proposition 5. *Let $\hat{\mu}_t = \frac{\mu_t}{N}$. Without uncertainty, the aggregate state $\hat{\mu}_t$ follows an evolution*

$$\hat{\mu}_{t+1} = \hat{\mu}_t \mathbb{Q}_t \tag{20}$$

where $\mathbb{Q}_t$ is the transition matrix $\mathbb{Q}$ such that:

$$\mathbb{Q}(s, s') = P(s, s') \cdot \mathbb{1}\{\varphi^s \geq \varphi^{s^*}\} + G(s')\mathbb{1}\{\varphi^s < \varphi^{s^*}\} \tag{21}$$

- *$\mathbb{Q}(s, s')$ is the (s, s')-th element of $\mathbb{Q}$;*
- *$P(s, s')$ is the (s, s')-th element of P;*
- *$G(s')$ is the s'-th element of G.*

Proof. This follows from Proposition 4 by consider definitions of M and dividing both sides by N. For a detailed proof, see Appendix A1. ■

D. *Carvalho and Grassi's Double Pareto Theorem*

With the setups, the law of motion and assumptions above, Carvalho and Grassi provide a detailed description of the steady-state equilibrium:

Proposition 6. *Under Assumption 1 and Proposition 5. If the potential entrant's productivity distribution is also Pareto (i.e., $G(s) = K_e(\varphi^s)^{-\delta_e}$), then as $S \to \infty$, the stationary firm productivity distribution $\hat{\mu}^* = (\hat{\mu}_1^*, \cdots, \hat{\mu}_S^*)$, given by*

$$\hat{\mu}^* = \hat{\mu}^* \mathbb{Q} \tag{22}$$

will converge point-wise to a distribution whose upper tail is Pareto distributed with a surviving function like

$$\mathbb{P}\{\varphi > \varphi^s\} = \lim_{S \to \infty} \sum_{i>s}^{S} \hat{\mu}_i^* = C \cdot (\varphi^\gamma)^{-\hat{\delta}} \tag{23}$$

where C is a constant, $\gamma > 0$ is a parameter, and $\hat{\delta} > 0$ is a parameter.

Proof. See Carvalho and Grassi (2019). ■

While this proposition characterizes the stationary firm-level productivity distribution, our previous settings imply that we can easily apply these results to the stationary firm size distribution with our setting. Recall that the stationary output of a firm with productivity level φ^s is given by

$$y_s = (\varphi^s)^{\frac{1}{1-\alpha}}(\frac{\alpha}{\bar{w}})^{\frac{\alpha}{1-\alpha}} \tag{24}$$

where $\bar{w}$ is the limit of the steady-state value of the wage when $S \to \infty$[14].

Therefore, in the steady-state, the number of firms with size y_s is given by μ_s and the percentage of firms with size y_s is given by $\hat{\mu}_s$.

IV. One Conjecture, a Proof and Graphical Examinations

In this section, we first put forward a conjecture based on proposition 6. Then we give a proof of the uniqueness of the stationary distribution in the conjecture. Our work is closed by graphically examining the upper tails of the stationary distributions with nine different entrant types.

A. One Conjecture and Its Uniqueness Proof

According to Proposition 6, the potential entrants must also follow a Pareto distribution as incumbents do so that the stationary distribution can converge to a distribution whose upper tail is Pareto when $S \to \infty$.

The condition that potential entrants must also follow a Pareto distribution might be too strict, and there is no substantial empirical evidence indicating that prospect entrants' distribution is Pareto.

By relaxing the distribution restriction on entrants in Proposition 6, we provide a conjecture on the stationary equilibrium.

Conjecture 1. *Under Assumption 1 and Proposition 5. If the potential entrant's productivity distribution follows any distribution with supports over Φ[15], then as $S \to \infty$, the stationary firm productivity distribution $\hat{\mu}^* = (\hat{\mu}^*_1, \cdots, \hat{\mu}^*_S)$, given by*

$$\hat{\mu}^* = \hat{\mu}^* \mathbb{Q} \tag{25}$$

will uniquely converge point-wise to a distribution whose upper tail is Pareto distributed with a surviving function like

$$\lim_{S\to\infty} \sum_{i>s}^{S} \hat{\mu}^*_i = C \cdot (\varphi^\gamma)^{-\hat{\delta}} \tag{26}$$

where C is a constant, $\gamma > 0$, $\hat{\delta} > 0$.

[14] Recall (14) that the wage w is pinned down by the distribution μ

[15] That is, $G(s) > 0, \forall s \in \Phi$.

Proof. I give a sketch of proof here. For a detailed proof, see A2. To prove the uniqueness of the stationary distribution, we should prove its existence firstly and then its uniqueness.

To prove its existence, recall the definition of Markov matrix, we show that $\mathbb{Q}$ is a Markov matrix, regardless of s^* and G. By Proposition 1, $\mathbb{Q}$ has at least one stationary distribution.

To prove its uniqueness, we show that $\mathbb{Q}$ is irreducible and aperiodic, regardless of s^* and G. For irreducibility, recall that I need to show all states communicate, and this happens because firms can both grow and shrink with $\mathbb{Q}$. For aperiodicity, I show that the period of all states is 1 and this happens because firms can remain the same productivity with $\mathbb{Q}$. By Proposition 2, $\mathbb{Q}$ has exactly one stationary distribution. ■

As we stated in section 3.D., the results of the stationary firm-level productivity distribution in Conjecture 1 can be easily applied to stationary firm size distribution.

B. Simulating Stationary Distribution

To verify whether the upper tail of the unique stationary distribution in Conjecture 1 is also Pareto when $S \to \infty$, we first compute the stationary firm distributions with some appropriate value assignments and nine different types of entrants' distributions. And then, we graphically examine whether their upper tails are Pareto distributed.

1. Solving for Stationary Distribution

Recall the proof in section 4.1, we have proved that the matrix $\mathbb{Q}$ in conjecture 1 is a Markov matrix, which is both irreducible and aperiodic.

By Proposition 2, we not only can prove the uniqueness of Markov matrix $\mathbb{Q}$'s stationary distribution $\hat{\mu}^*$, but also are provided with a method of calculating such stationary distribution $\hat{\mu}^*$, given any initial distribution $\hat{\mu}_0$:

$$As\ t \to \infty, \hat{\mu}_0 \mathbb{Q}^t \to \hat{\mu}^* \tag{27}$$

Algorithm. [16]

- **Step 1**
  According to Proposition 5, generate $\mathbb{Q}$ by considers Incumbents' transition matrix P and a given type of Entrants' distribution G.

- **Step 2**
  According to Proposition 2, compute the stationary distribution $\hat{\mu}^*$ with $\mathbb{Q}$ obtained from step 1 and a given initial distribution $\hat{\mu}_0 = G$.

[16] Note that the reasons for conducting steps 4-6 are discussed in the later.

- **Step 3**
  According to the definition, compute the counter cumulative distribution functions from the stationary distribution $\hat{\mu}^*$ obtained in step 2, $\sum_{i>s}^{S} \hat{\mu}_i^*$.
- **Step 4**
  Calculate the log term of the counter CDFs, $\ln \sum_{i>s}^{S} \hat{\mu}_i^*$, and of the firm productivity level, $\log \varphi^s$ with $\sum_{i>s}^{S} \hat{\mu}_i^*$ obtained in step 3 and φ^s from the setups.
- **Step 5**
  Draw the log-log graph with $\ln \sum_{i>s}^{S} \hat{\mu}_i^*$ and $\log \varphi^s$ obtained in step 4.
- **Step 6**
  Zoom in and focus on the right-hand tails of the log-log plot.

Our simulations follow the algorithm above. We firstly solve for its stationary firm-level productivity distribution with a specific type of the entrant's distribution. Then, we plot the survival function of this distribution on a log-log graph. Finally, we zoom in the figure to see the upper tail of the survival function. For each simulation, we produce a log-log graph about the right-hand tail of the stationary firm productivity distribution.

We implement such simulations for nine times, and each simulation is conducted with one of the nine different entrant's distributions, respectively. We label these simulations with numbers 1-9, shown in Table 2. For example, in simulation 1, the entrant's productivity is uniformly distributed.

The value assignments for these nine simulations are shown in Table 1. Since we consider the case when $S \to \infty$, we assign a large number to S. s^* is the threshold level for the exit-entry decisions, which is determined by the market condition. The value assignment for a and c follows Carvalho and Grassi (2019)'s calibration, which is estimated from the U.S.'s firm data. For following simulations, we use $(S, s^*, a, c, \varphi, \alpha, \beta) = (1000, 50, 0.6129, 0.3870, 1.1, 0.7, 0.2)$ from Table 1, if not specified.

Table 1—Value Assignments

Parameters	S	s^*	a	c	h	φ	S_1	S_2	α	β
Values	1000	50	0.6129	0.3870	1.57	1.1	171	86	0.7	0.2

Note: For default, we use $(S, s^*, a, c, \varphi, \alpha, \beta) = (1000, 50, 0.6129, 0.3870, 1.1, 0.7, 0.2)$ from the above table, if not specified.

2. Discussions on Results of Simulations

The graphical method we use for identifying upper tails of Pareto distributions is called the double-log plots. A log-log plot is a two-dimensional graph of nu-

TABLE 2—SIMULATIONS WITH NINE ENTRANT TYPES

Simulation No.	Entrant's Distribution	Probability Mass Function
1	Uniform	$G(s) = \frac{1}{S}$
2	Zipf	$G(s) = \frac{1}{(s)^h} \frac{1}{H_{S,h}}$
3	Logarithmic (series)	$G(s) = -\frac{p^s}{s \log(1-p)}$
4	Binomial	$G(s) = \binom{S-1}{s} p^s (1-p)^{S-1-s}$
5	Poisson	$G(s) = e^{-\lambda} \frac{\lambda^{s-1}}{(s-1)!}$
6	Geometric	$G(s) = (1-p)^{s-1} p$
7	Negative Binomial	$G(s) = \binom{s+S-1}{s} p^S (1-p)^s$
8	Beta-Binomial	$G(s) = \binom{S-1}{s} \frac{B(s+\alpha, S-1-s+\beta)}{B(\alpha,\beta)}$
9	Benford	$G(s) = \log_{S+1}(1 + \frac{1}{s})$

Note: $\forall s \in \Phi$, $0 < p < 1$, $\alpha > 0$ *and* $\beta > 0$, $H_{S,h} = \sum_{n=1}^{S} (\frac{1}{n^h})$ *and* $h \geq 0$, $Sp = \lambda \geq 0$ $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)}$ and $\Gamma(n) = (n-1)!$.

merical data that uses logarithmic scales on both the horizontal and vertical axes (Clauset et al. 2009). Given a distribution, the main idea of this method is that if the log-log graph of its counter cumulative distribution function tends to converge to a downward sloping straight line for large numbers in the x-axis, then we can conclude that the distribution has a Pareto (power-law) tail. Recall that if a random variable X is Pareto distributed, then its counter cumulative distribution function (counter-CDF) or equivalently survival function will have the form

$$F(x) = Pr\{X > x\} = C \cdot (x)^{-\omega} \tag{28}$$

where C is a constant, and $\omega > 0$ is a positive parameter determined by the tail index of the distribution.

Taking logarithmic of both sides of the (28), we have

$$\ln F(x) = \ln Pr\{X > x\} = \ln C - \omega \ln x \tag{29}$$

Since $\omega > 0$, the $\ln F(x)$ or $\ln Pr\{X > x\}$ will be linear in $\ln x$ with a negative slope $-\omega$. Note that although the log-log plot only provides necessary but insufficient evidence for a Pareto (power-law) relationship, it gives us a meaningful way to verify whether a distribution is Pareto or not.

With these expectations, we examine the results of the nine simulations mentioned above. The result for each simulation is a log-log plot of the upper tail of the stationary firm productivity distribution, shown from Figure A1. to Figure A5. We put the results of the simulations 1, 2, 3, 6, 9 in the same graph, Figure A1., and the results of the simulations 4, 5, 7, 8 are presented in Figure A2., Figure A3., Figure A4. and Figure A5., respectively.

From all figures, we can see that the upper tails of those log-log graphs tend to converge to downward-sloping straight lines, even though entrants follow the different distribution. Thus, we can conclude that results from all simulations fulfill our expectations, and the unique stationary productivity distribution in Conjecture 1 might have a Pareto tail. With our previous setting, this conclusion can be applied to the related stationary firm size distribution. Since the graphical examination only supplies a necessary but insufficient condition for a Pareto distribution, further investigations are required.

V. A Plan for the Next Step

For the next step, we are going to mainly focus on three aspects:

- First work on the remaining part of the conjecture and investigate whether the stationary firm distribution is Pareto, given any entrants' distribution.
- Work on other interesting parts of the firm dynamic model.
- Find real-world data, employ empirical tools to analyze quantiative implications of the model and the conjecture.

REFERENCES

Adelman, I. 1958. "A Stochastic Analysis of the Size Distribution of Firms." *American Statistical Association Journal*, 53(284): 893–900.

Axtell, R. L. 2001. "Distribution of U.S. Firm Sizes." *Science*, 293(5536): 1818–1820.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz. 2012. "Endogenous Entry, Product Variety, and Business Cycles." *Journal of Political Economy*, 120(2): 304–345.

Bloom, Nicholas, Max Floetotto Nir Jaimovich Itay Saporta-Eksten, and Stephen J. Terry. 2018. "Really Uncertain Business Cycles." *Econometrica*, 86(3): 1031–1065.

Brock, W. A. 1972. "On Models of Expectations that Arise from Maximizing Behaviour of Economic Agents over Time." *Journal of Economic Theory*, 5(3): 348–376.

Campbell, Jeffrey R., and Jonas D.M. Fisher. 2004. "Idiosyncratic risk and aggregate employment dynamics." *Review of Economic Dynamics*, 7(2): 331–353.

Carvalho, V.M., and B. Grassi. 2019. "Large Firm Dynamics and the Business Cycle." *American Economic Review*, 109(4): 1375–1425.

Champernowne, David G. 1953. "A Model of Income Distribution." *Economic Journal*, 63(250): 318–351.

Clauset Aaron, Shalizi Cosma Rohilla, and Newman M. E. J. 2009. "Power-law distributions in empirical data." *SIAM Review*, 51(4): 661–703.

Clementi, Gian Luca, and Berardino Palazzo. 2016. "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations." *American Economic Journal: Macroeconomics*, 8(3): 1–41.

Córdoba, Juan Carlos. 2008. "A Generalized Gibrat's Law." *International Economic Review*, 49(4): 1463–1468.

Florence, P. S. 1953. *The Logic of British and American Industry: A Realistic Analysis of Economic Structure and Government. Chapel Hill, N.C.*, The University of North Carolina Press.

Gabaix, X. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics*, 114(3): 739–767.

Gibrat, R. 1931. *Les Inegalieies Economiques; Applications: Aux Ineqalities des Richesses, a la Concentartion des Entreprises, Aux Populations des Villes, Aux Statistiques des Families, etc., d'une Loi Nouvelle, La Loi de l'Effet Proportionnel.* Librairie du Recueil Sirey.

Hart, P. E., and S. J. Prais. 1956. "The Analysis of Business Concentration: A Statistical Approach." *Journal of the Royal Statistical Society*, 119(2): 150–191.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5): 1127–1150.

Hsieh, Chang-Tai, and J-Klenow Peter. 2009. "Misallocation and manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403–1448.

Häggström, Olle. 2002. *Finite Markov Chains and Algorithmic Applications. London Mathematical Society Student Texts*, Cambridge University Press.

Ijiri, Y., and H.A. Simon. 1964. "Business Firm Growth and Size." *American Economic Review*, 54(2): 607–617.

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

Lee, Yoonsoo, and Toshihiko Mukoyama. 2008. "Entry, exit and plant-level dynamics over the business cycle." *Federal Reserve Bank of Cleveland Working Paper*, 0718.

Lee, Yoonsoo, and Toshihiko Mukoyama. 2015. "Entry and exit of manufacturing plants over the business cycle." *European Economic Review*, 77(3): 20–27.

Lucas, R., and E. C. Prescott. 1971. "Investment Under Uncertainty." *Econometrica*, 39(5): 659–681.

Newman, Mark E. J. 2005. "Power laws, Pareto distributions and Zipf's law." *Contemporary Physics*, 46(5): 323–351.

Pakes, A., and R. Ericson. 1998. "Selection and the Evolution of Industry." *Journal of Economic Theory*, 79: 1–45.

Simon, H.A., and C.P. Bonini. 1958. "The Size Distribution of Business Firms." *American Economic Review*, 48(1): 607–617.

Simon, Herbert A. 1955. "On a Class of Skew Distribution Functions." *Biometrika*, 42(3-4): 425–440.

Stachurski, J. 2009. *Economic Dynamics: Theory and Computation. Computational economics*, MIT Press.

Appendix

A1. Proof of Proposition 5

From proposition 4, we rewrite the expression (19) as

$$\mu_{t+1}(s) = \sum_{s'} P(\varphi^s, \varphi^{s'})\mu_t(s')\mathbb{1}\{\varphi^s \geq \varphi^{s^*}\} + MG(\varphi^{s'}) \tag{A1}$$

From my setup, we have

$$M = \sum_{q=1}^{s^*-1} \mu_t(q) \tag{A2}$$

By the above (A2), (A1) becomes

$$\mu_{t+1}(s) = \sum_{s'} P(\varphi^s, \varphi^{s'})\mu_t(s')\mathbb{1}\{\varphi^s \geq \varphi^{s^*}\} + \sum_{q=1}^{s^*-1} \mu_t(q)G(\varphi^{s'}) \tag{A3}$$

Since $\mu_t = (\mu_{t,1}, \mu_{t,2}, \cdots, \mu_{t,S})$, we obtain for all s,

$$\mu_{t+1}(s) = \mu_t P(\varphi^s, \cdot)\mathbb{1}\{\varphi^s \geq \varphi^{s^*}\} + \mu_t I \cdot \mathbb{1}\{\varphi^s < \varphi^{s^*}\}G(\varphi^{s'}) \tag{A4}$$

where $I = (1, 1, ..., 1)^T$ is S-dimension vector with all elements of 1.

Or equivalently, we can write (A4) in the form of matrix

$$\mu_{t+1} = \mu_t \tilde{P} + \mu_t \tilde{G} = \mu_t(\tilde{P} + \tilde{G}) \tag{A5}$$

where

- $\tilde{P}$ is the transition matrix P with first $s^*(\mu_t) - 1$ rows are replaced with $0s$.

- $\tilde{G}$ is the S-dimensional square zero matrix 0 with first $s^* - 1$ rows are replaced with Gs

Let $\mathbb{Q} = \tilde{P} + \tilde{G}$, that is, $\mathbb{Q}$ is the transition matrix P whose each of the first $s^* - 1$ rows are replaced with Gs. Then (A5) becomes

$$\mu_{t+1} = \mu_t \mathbb{Q} \tag{A6}$$

Divided both sides of the (A6) by N, we get

$$\frac{\mu_{t+1}}{N} = \frac{\mu_t}{N}\mathbb{Q} \tag{A7}$$

Let $\hat{\mu}_t = \frac{\mu_t}{N}$, we have

$$\mu_{t+1} = \mu_t \mathbb{Q} \tag{A8}$$

which is what we want. ■

A2. Proof of Conjecture 1's Uniqueness Statement

A2.1 Proof of Existence

Since P is a stochastic matrix, by definition of Markov matrix, we have

(i) $P \geq 0$; (ii) $\mathbb{1} \cdot P = \mathbb{1}$.

And for $G = (G(1), G(2), G(3), \cdots, G(S))$, by its definition, we have

(i) $G \geq 0$; (ii) $\mathbb{1} \cdot G = 1$.

By definition of the matrix $\mathbb{Q}$, since $\mathbb{Q}$ is the transition matrix P, where the first $(s^* - 1)$ rows are replaced by Gs, we have

(i) $\mathbb{Q} \geq 0$; (ii) $\mathbb{1} \cdot Q = \mathbb{1}$.

By definition, the matrix $\mathbb{Q}$ is also a Markov matrix.

By Proposition 1, the Markov matrix $\mathbb{Q}$ has at least one stationary distribution.

A2.2 Proof of Uniqueness

In order to prove the uniqueness of the stationary distribution of the matrix $\mathbb{Q}$, we need to prove its irreducibility and aperiodicity first.

A2.2.a Proof of Irreducibililty

To prove $\mathbb{Q}$ is irreducible, by definition of irreducibility we need to show that for all $(x, y) \in \{1, \cdots, S\} \times \{1, \cdots, S\}$, there exist positive integers n such that $\mathbb{Q}^n(x, y) > 0$.

Without loss of generosity, let's assume that $x \leq y$, then the definition of irreducibility says for all $(x, y) \in \{1, \cdots, S\} \times \{1, \cdots, S\}$, there exists an positive integer n such that

(i) $\mathbb{Q}^n(x, y) > 0$ and; (ii) $\mathbb{Q}^n(y, x) > 0$.

Since $\mathbb{Q}^{(y-x)}(x,y)$ can be represented as

$$\text{(A9)} \qquad \mathbb{Q}^{(y-x)}(x,y) = \begin{cases} \sum_{s=1}^{S} \mathbb{Q}^{(y-x-1)}(x,s)\mathbb{Q}(s,y) & if\ x < y \\ \mathbb{Q}(x,y) & if\ x = y \end{cases}$$

We notice that there exists one term $\mathbb{Q}(x,x+1)\mathbb{Q}(x+1,x+2)\cdots\mathbb{Q}(y-1,y)$ in the expansional expression of $\mathbb{Q}^{(y-x)}(x,y)$ when $x < y$; and $\mathbb{Q}^{(y-x)}(x,y) = \mathbb{Q}(x,y)$ when $x = y$.

Similarly, there exists one term $\mathbb{Q}(y,y-1)\mathbb{Q}(y-1,y-2)\cdots\mathbb{Q}(x+1,x)$ involved in the expansional expression of $\mathbb{Q}^{(y-x)}(y,x)$ when $y > x$; and $\mathbb{Q}^{(y-x)}(x,y) = \mathbb{Q}(x,y)$ when $x = y$.

Let $n = \{(y-x)|(x,y) \in \{1,\cdots,S\} \times \{1,\cdots,S\}\ and\ x \le y\} = S-1$.

Since other possible terms in the above expansional expressions are all non-negative, if we can prove that

(i) $\mathbb{Q}(s,s+1) > 0, \forall s \in \{1,\cdots,S-1\}$; (ii) $\mathbb{Q}(s,s-1) > 0, \forall s \in \{2,\cdots,S\}$. (iii) $\mathbb{Q}(s,s) > 0, \forall s \in \{1,\cdots,S\}$.

Then we can prove that for all $(x,y) \in \{(x,y) \in \{1,\cdots,S\}\times\{1,\cdots,S\}|x \le y\}$, and $n = S-1$, we have

(i) $\mathbb{Q}^n(x,y) > 0$ and; (ii) $\mathbb{Q}^n(y,x) > 0$.

By the definition of irreducibility, we know that the Markov matrix $\mathbb{Q}$ is irreducible.

Let's check whether the following inqualities are true for all $s^* \in \{1,\cdots,S\}$ by consider 3 cases.

(i) $\mathbb{Q}(s,s+1) > 0, \forall s \in \{1,\cdots,S-1\}$; (ii) $\mathbb{Q}(s,s-1) > 0, \forall s \in \{2,\cdots,S\}$. (iii) $\mathbb{Q}(s,s) > 0, \forall s \in \{1,\cdots,S\}$.

Case 1: $s^* = 1$**.**

In this case, we have $\mathbb{Q} = P$. By definition, we have

(i) $\mathbb{Q}(s,s+1) = P(s,s+1) = c, \forall s \in \{1,\cdots,S-1\}$; (ii) $\mathbb{Q}(s,s-1) = P(s,s-1) = a, \forall s \in \{2,\cdots,S\}$. (iii) $\mathbb{Q}(s,s) = P(s,s) = b, \forall z \in \{2,\cdots,S-1\}$; (iv) $\mathbb{Q}(1,1) = P(1,1) = a+b$; (v) $\mathbb{Q}(S,S) = P(S,S) = b+c$.

By the definition of the matrix P, we have $a > 0$, $b > 0$ and $c > 0$, so we obtain

(i) $\mathbb{Q}(s,s+1) = P(s,s+1) = c > 0, \forall s \in \{1,\cdots,S-1\}$; (ii) $\mathbb{Q}(s,s-1) = P(s,s-1) = a > 0, \forall s \in \{2,\cdots,S\}$. (iii) $\mathbb{Q}(s,s) = P(s,s) = b > 0, \forall s \in \{2,\cdots,S-1\}$; (iv) $\mathbb{Q}(1,1) = P(1,1) = a+b > 0$; (v) $\mathbb{Q}(S,S) = P(S,S) = b+c > 0$.

That is, when $s^* = 1$, the following inequalities hold:

(i) $\mathbb{Q}(s,s+1) > 0, \forall s \in \{1,\cdots,S-1\}$; (ii) $\mathbb{Q}(s,s-1) > 0, \forall s \in \{2,\cdots,S\}$. (iii) $\mathbb{Q}(s,s) > 0, \forall s \in \{1,\cdots,S\}$.

Case 2: $s^* \in \{2,...,S-1\}$**.**

In this case, by definition of the matrix $\mathbb{Q}$, for $s \in \{s^*, s^*+1,\cdots,S\}$, $\mathbb{Q}$'s row s are the same as in P.

By definition, we still have

(i) $\mathbb{Q}(s,s+1) = P(s,s+1) = c > 0, \forall s \in \{s^*,\cdots,S-1\}$; (ii) $\mathbb{Q}(s,s-1) =$

$P(s, s-1) = a > 0, \forall s \in \{s^*, \cdots, S\}$; (iii) $\mathbb{Q}(s, s) = P(s, s) = b > 0, \forall s \in \{s^*, \cdots, S-1\}$; (iv) $\mathbb{Q}(S, S) = P(S, S) = b + c > 0$.

Since for $s \in \{1, 2, \cdots, s^* - 1\}$, $\mathbb{Q}$'s row s are replaced by G, by definition of G, we have

(i) $\mathbb{Q}(s, s+1) = G(s+1) > 0, \forall s \in \{1, 2, \cdots, s^*-1\}$; (ii) $\mathbb{Q}(s, s-1) = G(s-1) > 0, \forall s \in \{2, 3, \cdots, s^* - 1\}$. (iii) $\mathbb{Q}(s, s) = G(s) > 0, \forall s \in \{1, \cdots, s^* - 1\}$.

Therefore, when $s^* \in \{2, ..., S - 1\}$, we still have

(i) $\mathbb{Q}(s, s+1) > 0, \forall s \in \{1, 2, \cdots, S-1\}$; (ii) $\mathbb{Q}(s, s-1) > 0, \forall s \in \{2, 3, \cdots, S\}$; (iii) $\mathbb{Q}(s, s) > 0, \forall s \in \{1, \cdots, S\}$.

Case 3: $s^* = S$

In this case, by definition of the matrix $\mathbb{Q}$, for $s = S$, $\mathbb{Q}$'s row s are the same as in P, that is,

(i) $\mathbb{Q}(S, S-1) = P(S, S-1) = a > 0$; (ii) $\mathbb{Q}(S, S) = P(S, S) = b + c > 0$.

Since for $s \in \{1, 2, \cdots, S - 1\}$, $\mathbb{Q}$'s row s are replaced by G, by definition of G, we have

(i) $\mathbb{Q}(s, s+1) = G(s+1) > 0, \forall s \in \{1, 2, \cdots, S-1\}$; (ii) $\mathbb{Q}(s, s-1) = G(s-1) > 0, \forall s \in \{2, 3, \cdots, S - 1\}$. (iii) $\mathbb{Q}(s, s) = G(s) > 0, \forall s \in \{1, \cdots, S - 1\}$.

Therefore, when $s^* = S$, we still have

(i) $\mathbb{Q}(s, s+1) > 0, \forall s \in \{1, 2, \cdots, S-1\}$; (ii) $\mathbb{Q}(s, s-1) > 0, \forall s \in \{2, 3, \cdots, S\}$; (iii) $\mathbb{Q}(s, s) > 0, \forall s \in \{1, \cdots, S\}$.

Overall, for all $s^* \in \{1, 2, ..., S\}$, we have

(i) $\mathbb{Q}(s, s + 1) > 0, \forall s \in \{1, \cdots, S - 1\}$; (ii) $\mathbb{Q}(s, s - 1) > 0, \forall s \in \{2, \cdots, S\}$. (iii) $\mathbb{Q}(s, s) > 0, \forall s \in \{1, \cdots, S\}$.

which is what we want to prove.

By the expansional expression of $\mathbb{Q}^{(y-x)}(x, y)$ and $\mathbb{Q}^{(y-x)}(y, x)$, we can conclude that for all $(x, y) \in \{(x, y) \in \{1, \cdots, S\} \times \{1, \cdots, S\} | x \leq y\}$, and $, n = \{(y - x) | (x, y) \in \{1, \cdots, S\} \times \{1, \cdots, S\} \; and \; x \leq y\} = S - 1$, we have

(i) $\mathbb{Q}^n(x, y) > 0$ and; (ii) $\mathbb{Q}^n(y, x) > 0$.

Therefore, by definition of Markov matrix's irreducibility, the Markov matrix Q is irreducible.

A2.2.b Proof of Aperiodicity

In order to prove $\mathbb{Q}$'s aperiodicity, we need to show that for all $s \in \{1, \cdots, S\}$, we have

$$gcd\{t : \mathbb{Q}^t(s, s) > 0\} = 1 \tag{A10}$$

where *gcd* is the greatest common divisor.

That is, for all $s \in \{1, \cdots, S\}$, we have $\mathbb{Q}(s, s) > 0$ We need to consider 2 cases by different values of s^*.

Case 1: $s^* = 1$**.**

By definition of $\mathbb{Q}$, $\mathbb{Q} = P$, so we have

(i) $\mathbb{Q}(s,s) = b, \forall s \in \{2, \cdots, S-1\}$ (ii) $\mathbb{Q}(s,s) = a+b, \forall s = 1$; (iii) $\mathbb{Q}(s,s) = b+c, \forall s = S$.

Since $a > 0$, $c > 0$ and $b > 0$, we have

(i) $\mathbb{Q}(s,s) = b > 0, \forall s \in \{2, \cdots, S-1\}$ (ii) $\mathbb{Q}(s,s) = a+b > 0, \forall s = 1$; (iii) $\mathbb{Q}(s,s) = b+c > 0, \forall s = S$.

When $s^* = 1$, we have $\mathbb{Q}(s,s) > 0, \forall s \in \{1, \cdots, S\}$.

That is, for all states $s \in \{1, 2, \cdots, S-1, S\}$, when $s^* = 1$, we have $gcd\{t : \mathbb{Q}^t(s,s) > 0\} = 1$ By definition of the aperiodicity of a Markov matrix, we know that when $s^* = 1$, the Markov matrix $\mathbb{Q}$ is aperiodic.

Case 2: $s^* \in \{2, ..., S-1\}$**.**

For $s \in \{s^*, \cdots, S\}$, rows s in Q are the same as in P, so we still have

(i) $\mathbb{Q}(s,s) = b > 0, \forall s \in \{s^*, \cdots, S-1\}$; (ii) $\mathbb{Q}(s,s) = b+c > 0, \forall s = S$.

For $s \in \{1, \cdots, s^*-1\}$, rows s in $\mathbb{Q}$ are replaced by G, so we have $\mathbb{Q}(s,s) = G(s) > 0, \forall s \in \{1, \cdots, s^*-1\}$;

When $s^* \in \{2, \cdots, S-1\}$, we still have $\mathbb{Q}(s,s) > 0, \forall s \in \{1, \cdots, S\}$;

That is, for states $s \in \{1, 2, \cdots, S-1, S\}$, when $s^* \in \{2, ..., S-1\}$, we also have $gcd\{t : \mathbb{Q}^t(s,s) > 0\} = 1$ By definition of the aperiodicity of a Markov matrix, we know that the Markov matrix $\mathbb{Q}$ is aperiodic, when $s^* \in \{2, ..., S-1\}$.

case 3: $s^* = S$**.**

For $s = S$, rows s in $\mathbb{Q}$ are the same as in P, so we still have

(i) $\mathbb{Q}(S,S) = P(S,S) = b+c > 0$; (ii) $\mathbb{Q}(S,S-1) = P(S,S-1) = a > 0$.

For $s \in \{1, \cdots, S-1\}$, rows s in $\mathbb{Q}$ are replaced by G, so we have $\mathbb{Q}(s,s) = G(s) > 0, \forall s \in \{1, \cdots, S-1\}$;

When $s^* = S$, we still have $\mathbb{Q}(s,s) > 0, \forall s \in \{1, \cdots, S\}$;

That is, for states $s \in \{1, 2, \cdots, S-1, S\}$, when $s^* = S$, we also have $gcd\{t : \mathbb{Q}^t(s,s) > 0\} = 1$ By definition of the aperiodicity of a Markov matrix, we know that the Markov matrix $\mathbb{Q}$ is aperiodic, when $s^* = S$.

By the above three cases, we can conclude that for $s^* \in \{1, \cdots, S\}$, we have $\mathbb{Q}(s,s) > 0, \forall s \in \{1, \cdots, S\}$;

That is, for states $s \in \{1, 2, \cdots, S-1, S\}$, when $s^* \in \{1, \cdots, S\}$, we also have $gcd\{t : \mathbb{Q}^t(s,s) > 0\} = 1$ By definition of the aperiodicity of a Markov matrix, we know that the Markov matrix $\mathbb{Q}$ is aperiodic, for all $s^* \in \{1, \cdots, S\}$.

A.2.2.c Proof of Uniqueness

Since the Markov matrix $\mathbb{Q}$ is irreducible and aperiodic, by Proposition 2, we can conclude that

- $\mathbb{Q}$ has exactly one stationary distribution μ^*;
- For any initial distribution μ_0, we have

$$\|\mu_0 \mathbb{Q}^t - \mu^*\| \to 0 \ as \ t \to \infty$$

■

A3. Simulation Results

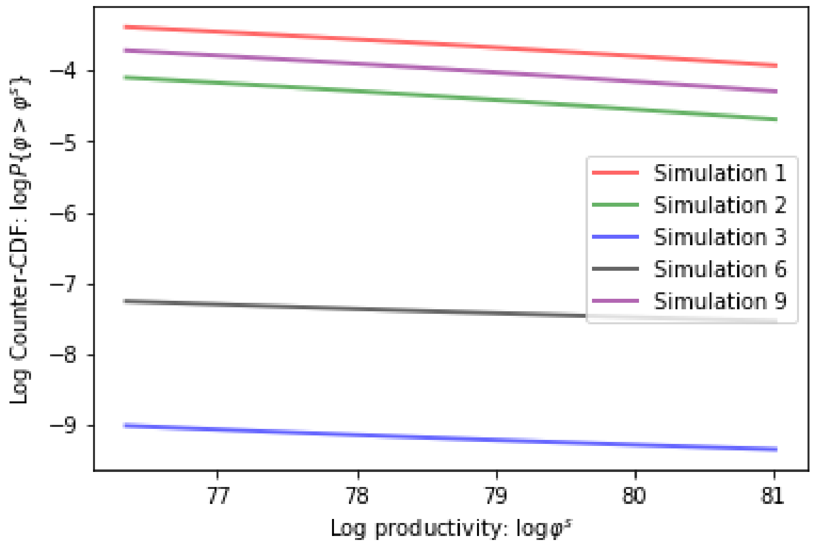

Figure A1. Upper Tails of Resulting Stationary Distributions in Simulation 1, 2, 3, 6, 9

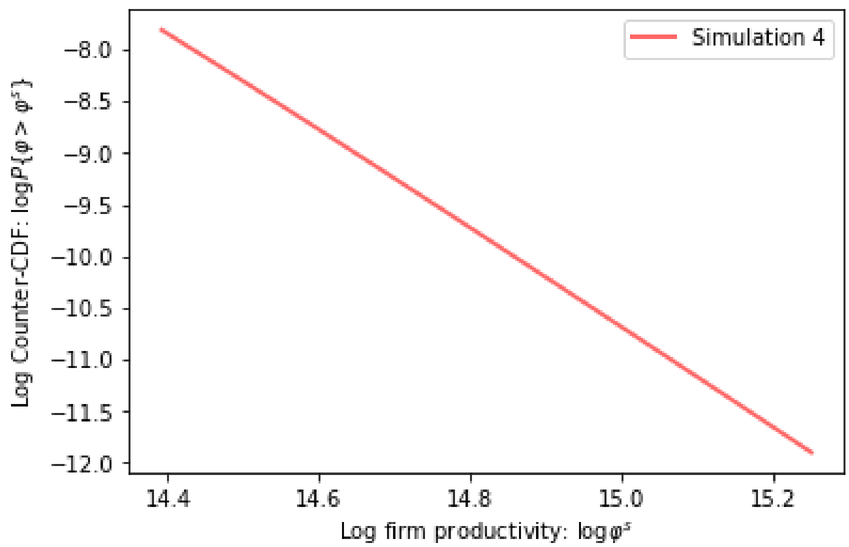

Figure A2. The Upper Tail of Resulting Distributions in Simulation 4

Note: $(S, p) = (S1, 0.8)$

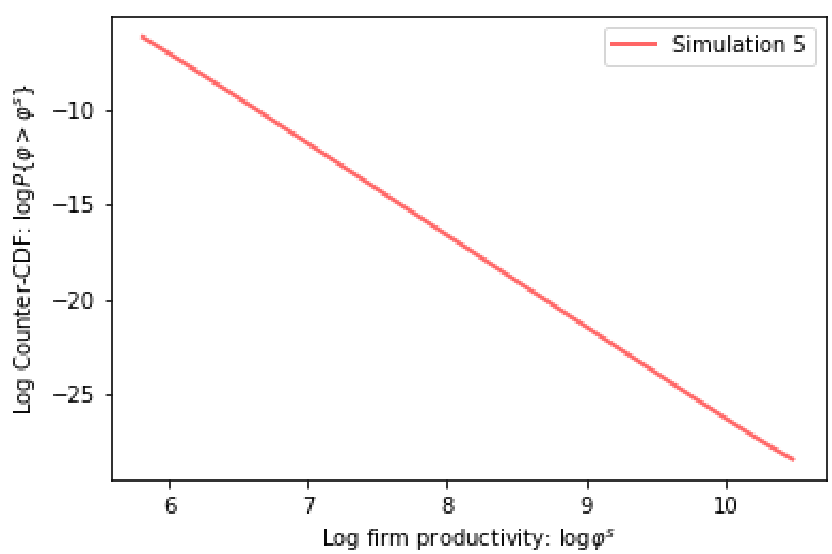

Figure A3. The Upper Tail of Resulting Stationary Distribution in Simulation 5

Note: $S = S1$

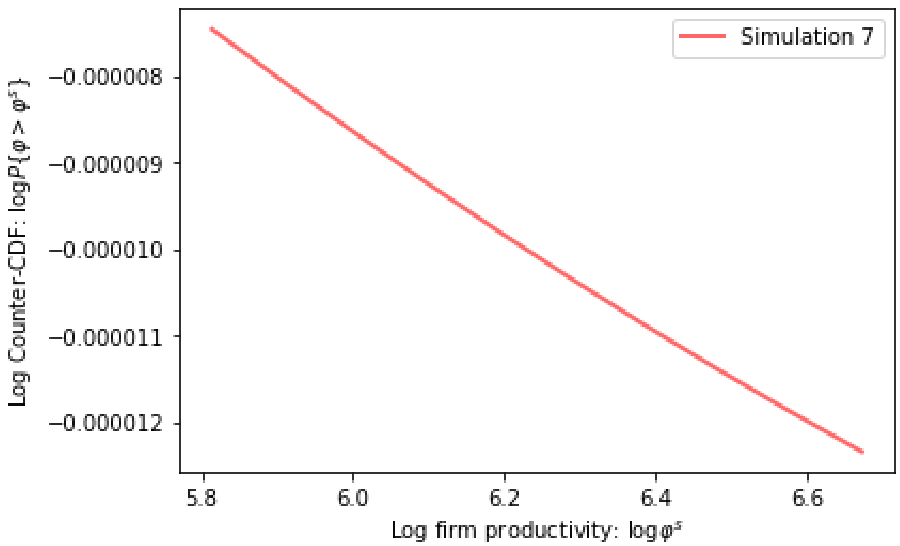

Figure A4. The Upper Tails of Resulting Stationary Distribution in Simulation 7

Note: $(S, p) = (S2, 0.5)$

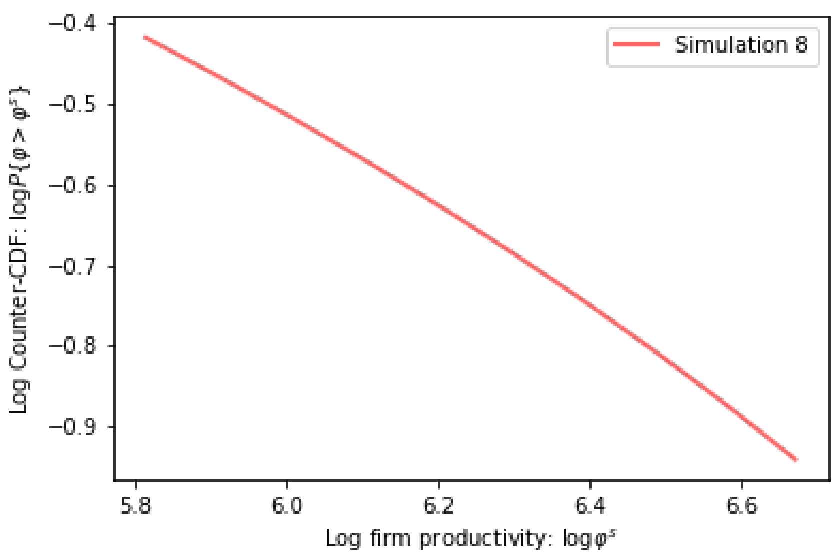

Figure A5. The Upper Tail of Resulting Stationary Distribution in Simulation 8

Note: $S = S2$