DETERMINATION OF PARETO EXPONENTS IN ECONOMIC MODELS DRIVEN BY MARKOV MULTIPLICATIVE PROCESSES

BRENDAN K. BEARE
School of Economics, University of Sydney

ALEXIS AKIRA TODA
Department of Economics, University of California San Diego

This article contains new tools for studying the shape of the stationary distribution of sizes in a dynamic economic system in which units experience random multiplicative shocks and are occasionally reset. Each unit has a Markov-switching type, which influences their growth rate and reset probability. We show that the size distribution has a Pareto upper tail, with exponent equal to the unique positive solution to an equation involving the spectral radius of a certain matrix-valued function. Under a nonlattice condition on growth rates, an eigenvector associated with the Pareto exponent provides the distribution of types in the upper tail of the size distribution.

KEYWORDS: Markov multiplicative process, Pareto tails, power law, random growth.

1. INTRODUCTION

THIS ARTICLE presents new tools for studying the shape of the stationary distribution of a dynamic economic system. We have in mind models in which a population of economic units experience random multiplicative shocks to their size over time, occasionally perishing at random and being replaced with a new unit. An economic unit could be, for instance, a household with size measured by wealth, or a firm with size measured by market capitalization. Economic units have a type, which may vary over time, such as a worker/entrepreneur type for households, or a productivity type for firms. The type affects the multiplicative growth rate, and perhaps also the survival rate, of an economic unit. Many heterogeneous-agent models of the Bewley–Aiyagari kind fit this general description if agents are subject to random mortality rather than having a certain finite or infinite lifespan.

It has long been known that random multiplicative growth is a generative mechanism for power laws. Random multiplicative growth plays a central role in economics and is known as Gibrat's law, while power laws are often referred to as Pareto tails. Early contributions to economics drawing a connection between random multiplicative growth and power laws include Champernowne (1953), Wold and Whittle (1957), and Simon and Bonini (1958). The topic has attracted a resurgence of interest in economics following the publication of Gabaix (1999) and Reed (2001); see Gabaix (2009, 2016) and Benhabib and Bisin (2018) for partial surveys. In mathematics, the central contribution is Kesten (1973). See Mitzenmacher (2004) for a discussion of many other relevant contributions across a range of disciplines.

Brendan K. Beare: brendan.beare@sydney.edu.au
Alexis Akira Toda: atoda@ucsd.edu

The primary contribution of this article is an equation whose unique positive solution is the Pareto (power law) exponent for the upper tail of the stationary distribution of sizes in a dynamic economic system of the kind described above. The equation is

$$\rho\big(\Pi \odot \Upsilon \odot \Psi(z)\big) = 1, \tag{1}$$

where $\rho(\cdot)$ is the spectral radius (maximum modulus of eigenvalues) of a square matrix, $\odot$ is the Hadamard (entrywise) product of matrices of the same size, Π is a matrix of transition probabilities for types, Υ is a matrix of survival rates, and $\Psi(z)$ is a matrix of moment generating functions of random growth rates. If equation (1) admits a unique positive solution (which is the case under mild conditions), say $z = \alpha$, then our main result, Theorem 1, establishes that α is the Pareto exponent for the upper tail of the stationary distribution of sizes.

In the simple case where there is only a single type of economic unit (which, in the class of models we consider, implies serial independence of the multiplicative shocks to size and is therefore quite restrictive), equation (1) simplifies to

$$\upsilon\psi(z) = 1, \tag{2}$$

where υ is the survival rate and $\psi(z)$ is the moment generating function of the random growth rate, both constant over time. Equation (2) appears as equation (10) in Manrubia and Zanette (1999), and now seems to be fairly widely known. Applications include Nirei and Aoki (2016) and Mukoyama and Osotimehin (2019) in macroeconomics, Montero and Villarroel (2013), Yamamoto (2014), and Meylahn, Sabhapandit, and Touchette (2015) in statistical physics, and Beare and Toda (2020) in epidemiology. Equation (2) is also closely related to the Cramér–Lundberg estimate of ruin probabilities used in actuarial science (see, e.g., Embrechts, Klüppelberg, and Mikosch (1997, Chapter 1)).

Our equation (1) extends (2) to a Markov setting in which economic units switch between types over time. There is a growing recognition in the literature that persistent heterogeneity in types plays a critical role in matching salient empirical regularities. Benhabib, Bisin, and Zhu (2011) study an overlapping generations model in which agents are subject to random labor (additive) and capital (multiplicative) income shocks. They allow these shocks to be persistent through their dependence on a latent Markov state (i.e., type), writing (p. 127) "the i.i.d. condition is very restrictive. Positive autocorrelations in [capital and labor income shocks] capture variations in social mobility in the economy, for example, economies in which returns on wealth and labor earning abilities are transmitted across generations." Cao and Luo (2017) study a similar model in which investors switch at random between high and low productivity types, while subject to random mortality. Regarding the importance of persistent heterogeneity in types, they write (p. 302) "the model with homogeneous returns, when calibrated to match the salient aggregate statistics, produces a tail index that is an order of magnitude too high compared to the one in the data." Based on an analysis of Norwegian administrative tax records, Fagereng, Guiso, Malacrino, and Pistaferri (2020) confirm these observations on the importance of type dependence for explaining the upper tail of the wealth distribution, writing that (pp. 118–119) "persistent traits of individual investors (such as financial sophistication, the ability to process and use financial information, the ability to overcome inertia, and for entrepreneurs, the talent to manage and organize their businesses) are capable of generating persistent differences in returns to wealth that may be as relevant as those conventionally attributed in household finance to differences in risk exposure or scale." Taking a somewhat different angle, Gabaix, Lasry, Lions, and Moll (2016) emphasize the

importance of persistent heterogeneity in productivity types not only for generating an income distribution with a tail sufficiently heavy to match U.S. data, but also for generating transition dynamics that are sufficiently fast to match the rate of increase of top income inequality. Additional articles discussing the importance of persistent heterogeneity in types for constructing plausibly calibrated models are discussed and cited therein (p. 2095).

Mathematically, the new results introduced in this article concern a class of Markov chains called *Markov multiplicative processes with reset*, which we define in Section 2. Loosely, these are processes driven by positive multiplicative shocks dependent on a Markov state variable, and subject to occasional reset to one at a state-dependent rate. After a heuristic discussion in Section 3 aimed at building intuition, we state our main results in Section 4. Under mild conditions, these results establish (i) that the left-hand side of equation (1) is convex in z and admits at most one positive solution; (ii) a sufficient condition for the existence of a positive solution; (iii) the existence of a unique stationary distribution for a Markov multiplicative process with reset; (iv) a formula for the moment generating function of the stationary distribution (after taking logs); (v) that a positive solution to equation (1) is the upper Pareto exponent for the stationary distribution. In Section 5, we explore refinements obtaining under a nonlattice condition on growth rates, establishing that (i) the upper tail of the stationary distribution satisfies a stronger form of Pareto decay; (ii) an eigenvector associated with the Pareto exponent provides the distribution of types in the upper tail. We lay out some directions for future research in Section 6. A mathematical Appendix to Sections 4–5 (Appendix A) contains the proofs of all numbered results, and a second Appendix includes a discussion of comparative statics for the Pareto exponent (Appendix B).

Continuous-time versions of the results in this article have been established in a companion article, Beare, Seo, and Toda (2022). There we study the tail probabilities of a Markov-modulated Lévy process stopped at a state-dependent Poisson rate. The tail exponents in continuous time are determined by an equation similar to (1), but with a spectral abscissa taking the role of the spectral radius.

While we do not provide a substantive economic application of our results in this article, several recent articles and circulated manuscripts provide applications.[1] Toda (2019, Theorem 4) uses equation (1) to determine the Pareto exponent for the wealth distribution in a Huggett economy with stochastic discounting. Ma, Stachurski, and Toda (2020, Theorem 3.3) do the same in a model with stochastic discounting, returns on wealth, and labor income. Gomez and Gouin-Bonenfant (2020, Proposition 3) use our results to determine the Pareto exponent of wealth in an economy populated with entrepreneurs and rentiers. Gouin-Bonenfant (2020, Proposition A.10) uses our results to calculate the Pareto exponent for the distribution of firm size in a model of the labor market, finding that Zipf's law is approximately satisfied. Gouin-Bonenfant and Toda (2022) show how to improve grid-based calculation of aggregate quantities in dynamic economic models by using equation (1) to extrapolate beyond the grid. Beare, Seo, and Toda (2022) use a continuous-time version of equation (1) to study the tails of wealth in a Huggett economy inhabited by agents with constant absolute risk aversion.

---

[1]Articles in which our results are discussed sometimes cite an earlier version of this article titled "Geometrically stopped Markovian random growth processes and Pareto tails," available on the arXiv e-print repository since December 2017: https://arxiv.org/abs/1712.01431.

2. MARKOV MULTIPLICATIVE PROCESSES WITH RESET

Let $\mathbb{Z}_+$ be the set of nonnegative integers and $\mathcal{N} = \{1, \ldots, N\}$, a finite set. The elements of $\mathcal{N}$ index the types of agents in a dynamic economic system. We will study the behavior of a homogeneous[2] Markov chain $(S_t, J_t)_{t \in \mathbb{Z}_+}$ on $(0, \infty) \times \mathcal{N}$. In each time period t, S_t indicates the size of an agent (e.g., their wealth) and J_t their type. The transition probabilities of our Markov chain are parametrized in terms of four objects:

(a) *Transition probabilities for J_t without reset*: An $N \times N$ matrix Π with nonnegative entries $\pi_{nn'}$, and all rows summing to one. The entries of Π are the probabilities with which agents transition between types each period.

(b) *Type-dependent survival probabilities*: An $N \times N$ matrix Υ whose entries $\upsilon_{nn'}$ are between zero and one inclusive. The entries of Υ are the survival probabilities for agents transitioning from type n to type n'.

(c) *Type probabilities for J_t upon reset*: An $N \times 1$ vector ϖ with nonnegative entries ϖ_n summing to one. The entries of ϖ provide the distribution of types for a new agent replacing an agent who has perished.

(d) *Type-dependent growth rate distributions for S_t*: A collection of N^2 cumulative distribution functions (CDFs) on $(0, \infty)$, denoted $\Phi_{nn'}$ with $n, n' \in \mathcal{N}$. The CDF $\Phi_{nn'}$ characterizes the distribution of growth rates for an agent transitioning from type n to type n'.

Conditional on having $S_t = s$ and $J_t = n$, the subsequent pair (S_{t+1}, J_{t+1}) may be generated in the following way:

(i) Choose J_{t+1}^* at random from $\mathcal{N}$, with $\mathrm{P}(J_{t+1}^* = n') = \pi_{nn'}$. This will be the value taken by J_{t+1} contingent on reset not occurring.

(ii) Conditional on choosing $J_{t+1}^* = n'$ in Step (i), reset occurs with probability $1 - \upsilon_{nn'}$.

(iii) If reset occurs, then we set $S_{t+1} = 1$ and choose J_{t+1} at random from $\mathcal{N}$, with $\mathrm{P}(J_{t+1} = n') = \varpi_{n'}$.

(iv) If reset does not occur, then we set $J_{t+1} = J_{t+1}^*$ and, conditional on choosing $J_{t+1}^* = n'$ in Step (i), we draw a positive random variable $G_{nn'}$ from the CDF $\Phi_{nn'}$ and set $S_{t+1} = G_{nn'}s$.

Steps (i)–(iv) determine the Markov kernel for $(S_t, J_t)_{t \in \mathbb{Z}}$. Specifically, given $n, n' \in \mathcal{N}$ and $s, s' > 0$, the Markov kernel is

$$
P_{sn}(s', n') := \mathrm{P}\big(S_{t+1} \le s', J_{t+1} = n' \mid S_t = s, J_t = n\big)
$$
$$
= \underbrace{\pi_{nn'} \upsilon_{nn'} \Phi_{nn'}(s'/s)}_{J_{t+1}^* = n' \to \text{no reset} \to S_{t+1} \le s'} + \mathbb{1}(s' \ge 1) \sum_{n''} \underbrace{\pi_{nn''}(1 - \upsilon_{nn''})\varpi_{n'}}_{J_{t+1}^* = n'' \to \text{reset} \to J_{t+1} = n'} . \tag{3}
$$

We refer to a homogeneous Markov chain on $(0, \infty) \times \mathcal{N}$ with Markov kernel given by (3) as a *Markov multiplicative process with reset*. We call the multiplicative factors $G_{nn'}$ gross growth rates, and their logarithms growth rates.

Under further regularity conditions, a Markov multiplicative process with reset can be made stationary with a suitable choice of its distribution at time $t = 0$. This will be made clear in Proposition 3 below.

3. HEURISTIC DERIVATIONS

To build intuition, in this section we provide a brief heuristic derivation of our equation determining the Pareto exponent of a stationary Markov multiplicative process with reset.

---

[2]Homogeneity means that $\mathrm{P}(S_{t+1} \le s', J_{t+1} = n' \mid S_t = s, J_t = n)$ does not depend on t.

Let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be such a process, and conjecture that for all sufficiently large $s > 1$ we have

$$P(S_t > s, J_t = n) = y_n s^{-\alpha} \tag{4}$$

for some constant $y_n > 0$ and Pareto exponent $\alpha > 0$. Then

$$y_{n'} s^{-\alpha} = P(S_{t+1} > s, J_{t+1} = n') = \sum_{n=1}^{N} \pi_{nn'} \upsilon_{nn'} P(G_{nn'} S_t > s, J_t = n), \tag{5}$$

where $G_{nn'} > 0$ is the gross growth rate drawn from the CDF $\Phi_{nn'}$ in Step (iv) in Section 2. Applying the law of iterated expectations, we obtain

$$
\begin{aligned}
P(G_{nn'} S_t > s, J_t = n) &= E\big(P(S_t > s/G_{nn'}, J_t = n \mid G_{nn'})\big) \\
&= E\big(y_n (s/G_{nn'})^{-\alpha}\big) = y_n E\big(G_{nn'}^{\alpha}\big) s^{-\alpha}. \tag{6}
\end{aligned}
$$

For real z, set

$$\psi_{nn'}(z) := \int_0^{\infty} s^z \, d\Phi_{nn'}(s), \tag{7}$$

the moment generating function (MGF) of $\log G_{nn'}$. Let $\Psi(z)$ be the $N \times N$ matrix-valued function with entries $\psi_{nn'}(z)$. Noting that $\psi_{nn'}(\alpha) = E(G_{nn'}^{\alpha})$, and combining (5), (6), and (7), we obtain

$$y_{n'} = \sum_{n=1}^{N} y_n \pi_{nn'} \upsilon_{nn'} \psi_{nn'}(\alpha). \tag{8}$$

Letting $y = (y_1, \ldots, y_N)^{\top}$ and collecting (8) into a vector, we obtain

$$y^{\top} = y^{\top} \big(\Pi \odot \Upsilon \odot \Psi(\alpha)\big).$$

Thus $\Pi \odot \Upsilon \odot \Psi(\alpha)$ has a unit eigenvalue, with associated left eigenvector y.

The matrix $\Pi \odot \Upsilon \odot \Psi(\alpha)$ is nonnegative; suppose that it is also irreducible. The Perron–Frobenius theorem (see, e.g., Horn and Johnson (2013, Theorem 8.4.4)), a fundamental result in the theory of Markov chains, asserts that the spectral radius of a square, nonnegative, and irreducible matrix is the maximum real eigenvalue of that matrix, and that there are unique (up to positive scalar multiplication) left and right eigenvectors with positive entries corresponding to that eigenvalue. Since y has positive entries and is a left eigenvector of $\Pi \odot \Upsilon \odot \Psi(\alpha)$ associated with a unit eigenvalue, we are thus led to suspect that $\Pi \odot \Upsilon \odot \Psi(\alpha)$ may have spectral radius equal to one. This suspicion, if valid, suggests the possibility that we may be able to determine the Pareto exponent α by finding a positive number z that solves equation (1). The fact that this is indeed often possible is the main result of our article, Theorem 1.

The left eigenvector y has a natural interpretation. It follows from (4) that

$$P(J_t = n \mid S_t > s) = \frac{P(S_t > s, J_t = n)}{P(S_t > s)} = \frac{y_n s^{-\alpha}}{\displaystyle\sum_{n'=1}^{N} y_{n'} s^{-\alpha}} = \frac{y_n}{\displaystyle\sum_{n'=1}^{N} y_{n'}}.$$

Therefore, if the left eigenvector y is normalized such that its entries sum to one, then we have $\mathrm{P}(J_t = n \mid S_t > s) = y_n$. We may thus interpret y to be, loosely, the distribution of types in the upper tail of the size distribution. We formalize this claim in Theorem 2.

The preceding discussion is largely heuristic. The conjectured equality (4) cannot be expected to hold exactly except in very special cases. We will see that (4) holds approximately for large s in fairly general settings. The above derivations based on (4) are therefore also approximate. We have skirted technical issues such as the existence of a unique stationary distribution for (S_t, J_t), the finiteness of $\psi_{nn'}(\alpha)$, and the existence of a unique positive solution to (1). Derivations similar to those above appear in Gouin-Bonenfant and Toda (2022). We now turn to a rigorous statement of our main results.

4. MAIN RESULTS

In this section, we provide a rigorous statement of our primary result, which is that under regularity conditions a Markov multiplicative process with reset has a Pareto upper tail with exponent α given by the unique positive solution to equation (1). We rely on the following regularity conditions.

ASSUMPTION 1: (*i*) *The matrix* $\Pi \odot \Upsilon$ *is irreducible.* (*ii*) *There exist states* $n, n' \in \mathcal{N}$ *such that* $\pi_{nn'} > 0$ *and* $v_{nn'} < 1$. (*iii*) $\psi_{nn'} \equiv 1$ *whenever* $\pi_{nn'} v_{nn'} = 0$.

Assumption 1(i) means that, for any pair of states (n, n'), if J_t is in state n then there is a positive probability of it eventually reaching state n' before reset occurs. Assumption 1(ii) means that there is some state n such that, if J_t is in state n, then reset occurs next period with positive probability. These two conditions together ensure that J_t visits all states infinitely often and that reset occurs infinitely often. Assumption 1(iii), imposed without loss of generality, is merely a normalization since the growth rate when J_t transitions from state n to state n' is unidentified if this transition never occurs.

To clarify the heuristic discussion in Section 3, the first thing we will do is restrict the domain of the matrix-valued function $\Psi(z)$ such that each of its entries $\psi_{nn'}(z)$, defined in (7), is finite-valued. Define the set

$$\mathcal{I} = \left\{ z \in \mathbb{R} : \psi_{nn'}(z) < \infty \text{ for all } n, n' \in \mathcal{N} \right\}.$$

Since MGFs are always convex and are equal to one at zero, $\mathcal{I}$ is a convex set containing zero. We restrict the domain of $\Psi(z)$ to $\mathcal{I}$ and, for $z \in \mathcal{I}$, define the $N \times N$ matrix-valued function

$$\mathrm{A}(z) = \Pi \odot \Upsilon \odot \Psi(z). \tag{9}$$

Using (9), equation (1) may now be rewritten more simply as $\rho(\mathrm{A}(z)) = 1$.

We claimed in Section 3 that it is often the case that there is a unique positive value of z that solves the equation $\rho(\mathrm{A}(z)) = 1$. The following result concerning the shape of the function $\rho(\mathrm{A}(z))$ helps to explain why this is true.

PROPOSITION 1: *The spectral radius* $\rho(\mathrm{A}(z))$ *is a convex function of* $z \in \mathcal{I}$ *and satisfies* $\rho(\mathrm{A}(0)) \leq 1$. *If Assumption 1 holds, then* $\rho(\mathrm{A}(0)) < 1$ *and the equation* $\rho(\mathrm{A}(z)) = 1$ *has at most one positive solution* $z = \alpha \in \mathcal{I}$ *and at most one negative solution* $z = -\beta \in \mathcal{I}$.
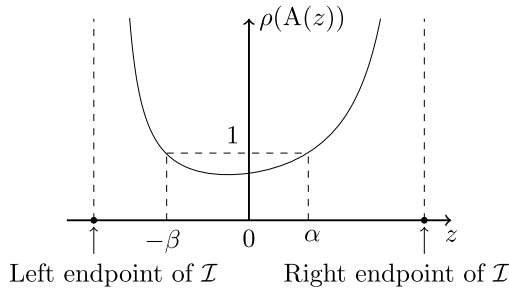
FIGURE 1.—Determination of α and β from the spectral radius of $A(z)$.

Figure 1 depicts a typical shape for $\rho(A(z))$ as a function of $z \in \mathcal{I}$. The graph of $\rho(A(z))$ is convex, less than one at zero (note that $A(0) = \Pi \odot Y$), and diverging to infinity at the left and right endpoints of $\mathcal{I}$ (which may in general be finite or infinite). There is a unique positive number α at which $\rho(A(z))$ crosses one. We will see that α is the Pareto exponent for the upper tail of the stationary distribution of S_t. There is also a unique negative number $-\beta$ at which $\rho(A(z))$ crosses one. We will see that β is the Pareto exponent for the lower tail of the stationary distribution of S_t.

The graph of $\rho(A(z))$ does not always have the typical shape depicted in Figure 1. Three atypical cases in which no positive z solves $\rho(A(z)) = 1$ are depicted in Figure 2. If all growth rates are nonpositive, so that $P(G_{nn'} \leq 1) = 1$ for all $n, n' \in \mathcal{N}$, then $\rho(A(z))$ is nonincreasing in z and we have the case depicted in Figure 2(a). If any growth rate does not have a light upper tail, so that the corresponding MGF $\psi_{nn'}(z)$ is infinite for all positive z, then the right endpoint of $\mathcal{I}$ is zero and we have the case depicted in Figure 2(b). Problems may also arise if the tails of growth rates are insufficiently light. For instance, suppose for simplicity that we have no Markov modulation ($N = 1$) and that $\log G$ has PDF:

$$f(x) = \begin{cases} c(x+2)^{-2}e^{-ax} & \text{for } x \geq -1, \\ 0 & \text{for } x < -1, \end{cases}$$

where $a > 0$ and c is a positive constant such that the PDF integrates to one. In this case, $E(G^z)$ is finite for $z \leq a$ and infinite for $z > a$, so that $\mathcal{I} = (-\infty, a]$. The existence of a positive solution to $\rho(A(z)) = 1$ thus depends on whether $\rho(A(a))$ is greater than one. The case where $a = 1$ and $v = 0.7$ is depicted in Figure 2(c); we see that $\rho(A(a)) < 1$, so that there is no positive solution to $\rho(A(z)) = 1$.

The following result establishes that, to have a unique positive solution to $\rho(A(z)) = 1$, it suffices that all growth rates have exponential moments of all orders, with at least one growth rate $\log G_{nn}$ positive with positive probability. An analogous result applies symmetrically to the existence of a unique negative solution to $\rho(A(z)) = 1$.

PROPOSITION 2: *Suppose that Assumption 1 is satisfied. If $E(G_{nn'}^z) < \infty$ for all $z > 0$ and all $n, n' \in \mathcal{N}$, and if $P(G_{nn} > 1) > 0$ for some $n \in \mathcal{N}$, then there is a unique positive value of z in the interior of $\mathcal{I}$ such that $\rho(A(z)) = 1$.*

We mentioned at the end of Section 2 that, under suitable regularity conditions, a Markov multiplicative process with reset can be made stationary with a suitable choice of its distribution at time $t = 0$. The following result indicates that Assumption 1 provides
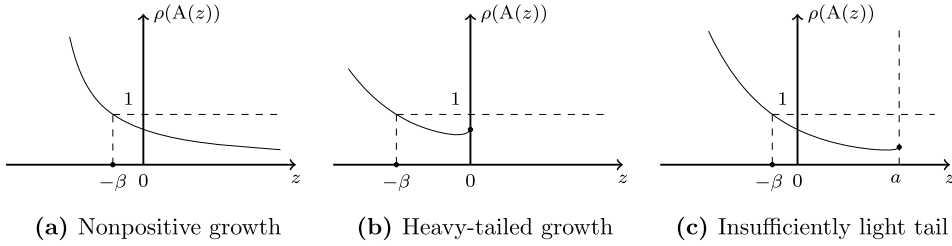
FIGURE 2.—Examples where no positive z solves $\rho(\mathrm{A}(z)) = 1$.

sufficient regularity. It is proved by observing that reset generates a positive recurrent accessible atom of $(S_t, J_t)_{t \in \mathbb{Z}_+}$, which implies the existence of a unique stationary distribution.

PROPOSITION 3: *Let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be a Markov multiplicative process with reset. If Assumption 1 is satisfied, then there exists a unique probability distribution for (S_0, J_0) such that $(S_t, J_t)_{t \in \mathbb{Z}_+}$ is stationary.*

We denote by p the $N \times 1$ vector of stationary probabilities for J_t, and by q the $N \times 1$ vector whose nth entry is $q_n = \sum_{n'} \pi_{nn'}(1 - v_{nn'})$, the conditional probability of reset in period $t + 1$ given $J_t = n$. The unconditional probability of reset is then $r := \sum_n p_n q_n$. It will also be useful to introduce notation for the subset of $\mathcal{I}$ on which $\mathrm{A}(z)$ has spectral radius less than one. We therefore set $\mathcal{I}_- = \{z \in \mathcal{I} : \rho(\mathrm{A}(z)) < 1\}$, which is a convex subset of $\mathcal{I}$ by Proposition 1, nonempty and containing zero under Assumption 1. In typical cases such as the one depicted in Figure 1, we have $\mathcal{I}_- = (-\beta, \alpha)$.

The following result, Proposition 4, is established in the course of proving our main result, Theorem 1 below, but is also of independent interest. It reveals the form of the conditional MGF of $\log S_t$ given J_t, which has domain $\mathcal{I}_-$. The notation I refers to an $N \times N$ identity matrix.

PROPOSITION 4: *Let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be a stationary Markov multiplicative process with reset. Then, for each $z \in \mathcal{I}_-$, the matrix $\mathrm{I} - \mathrm{A}(z)$ is invertible and*

$$\left[\mathrm{E}\big(S_t^z \mathbb{1}(J_t = 1)\big) \quad \cdots \quad \mathrm{E}\big(S_t^z \mathbb{1}(J_t = N)\big) \right] = r \boldsymbol{\varpi}^\top \big(\mathrm{I} - \mathrm{A}(z)\big)^{-1}.$$

An immediate consequence of Proposition 4 is that the MGF of $\log S_t$ is given by $\mathrm{E}(S_t^z) = r \boldsymbol{\varpi}^\top (\mathrm{I} - \mathrm{A}(z))^{-1} 1_N$ for $z \in \mathcal{I}_-$, where 1_N is an $N \times 1$ vector of ones. Proposition 4 thus completely characterizes the stationary distribution of sizes whenever the interior of $\mathcal{I}_-$ contains zero.

We now state our main result.

THEOREM 1: *Let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be a stationary Markov multiplicative process with reset satisfying Assumption 1. If the equation $\rho(\mathrm{A}(z)) = 1$ admits a unique positive solution $z = \alpha$ in the interior of $\mathcal{I}$, then the limits inferior and superior of $s^\alpha \mathrm{P}(S_t > s)$ as $s \to \infty$ are positive and finite, and*

$$\lim_{s \to \infty} \frac{\log \mathrm{P}(S_t > s)}{\log s} = -\alpha. \tag{10}$$

Equation (10) indicates that there is an approximately linear relationship in log-log scale between the tail probability $P(S_t > s)$ and the threshold s for large s, with slope $-\alpha$. In this sense, the upper tail of the distribution of S_t is Pareto with exponent α. The fact that this approximately linear relationship in log-log scale arises in distributions with a Pareto upper tail is the basis for the popular log-log rank-size regression method of estimating the Pareto exponent; see, for example, Gabaix and Ibragimov (2011). Note that Theorem 1 does not assert the convergence of $s^\alpha P(S_t > s)$ to a positive and finite limit, which would be a stronger notion of Pareto tail decay. This stronger property is not satisfied in general, but may be guaranteed by imposing a nonlattice condition on the distribution of growth rates, as discussed in Section 5. To obtain the convergence (10), it suffices that $s^\alpha P(S_t > s)$ has positive and finite limits inferior and superior. To see why, note that the latter condition implies the existence of positive and finite constants c_- and c_+ such that $c_- \leq s^\alpha P(S_t > s) \leq c_+$ for all sufficiently large s. Taking logarithms, dividing by $\log s$, and letting $s \to \infty$, we obtain (10).

We close this section with four remarks on Theorem 1 and a brief example.

REMARK 1: By replacing S_t with $1/S_t$, it is easy to deduce from Theorem 1 that the lower tail probability $s^{-\beta} P(S_t < s) = (1/s)^\beta P(1/S_t > 1/s)$ has a positive and finite limit as $s \downarrow 0$, where $z = -\beta$ is the unique *negative* solution to the equation $\rho(A(z)) = 1$ in the interior of $\mathcal{I}$ (if it exists). Thus the stationary distribution of S_t also exhibits a power law in the *lower* tail.[3]

REMARK 2: If the survival probabilities $v_{nn'}$ are constant across states, so that we may write $v_{nn'} = v \in (0, 1)$, then the equation $\rho(A(z)) = 1$ reduces to $v\rho(\Pi \odot \Psi(z)) = 1$. If in addition the growth rate MGFs $\psi_{nn'}$ depend only on the current state n', so that we may write $\psi_{nn'} = \psi_{n'}$, then letting $D(z) = \text{diag}(\psi_1(z), \ldots, \psi_N(z))$, the equation $\rho(A(z)) = 1$ reduces to

$$v\rho\big(\Pi D(z)\big) = 1. \tag{11}$$

Similarly, if the growth rate MGFs $\psi_{nn'}$ depend only on the *previous* state n, then the equation $\rho(A(z)) = 1$ reduces to $v\rho(D(z)\Pi) = 1$, which is identical to (11) noting that $\rho(AB) = \rho(BA)$ in general.

REMARK 3: Our characterization of the upper Pareto exponent α in Theorem 1 is implicit, in the sense that it is given by the positive solution to $\rho(A(z)) = 1$. It may be useful for economic applications to provide comparative statics for α; that is, results indicating how α varies as we vary the parameters that enter into the matrix $A(z)$. We provide a rigorous statement of such results in Appendix B. The main findings are as follows: (i) A marginal increase in any survival rate $v_{nn'}$ reduces α. (ii) A marginal increase in the mean of any growth rate $\log G_{nn'}$ reduces α. (iii) A marginal increase in the scale of any growth

---

[3]Although not widely known, lower-tail power-law behavior in the distribution of economic aggregates has been documented in a number of empirical studies. For city size, see the bottom-right panel of Figure 1 of Reed (2001) (who writes "lower-tail power-law behaviour…is not apparently widely recognized"), top panels of Figures 1 and 2 of Reed (2002) (who writes "the lower-tail plots exhibit linearity"), footnote 8 of Giesen, Zimmermann, and Suedekum (2010), who write "Among the 100 smallest cities we also find a distinctive power law pattern," and Figure 2 of Devadoss, Luckstead, Danforth, and Akhundjanov (2016). Aside from city size, power law behavior in the lower tail has been documented in income (bottom-left panel of Figure 1 of Reed (2001); Figure 3 of Toda (2011); Figure 1 of Toda (2012)) and consumption (Figure 2 of Toda and Walsh (2015); and Toda (2017)).

rate $\log G_{nn'}$ reduces α. (iv) If the survival probabilities $\upsilon_{nn'}$ and growth rates $\log G_{nn'}$ depend only on the current state n', or only on the previous state n, then a marginal increase in the persistence of the Markov modulator J_t reduces α.

REMARK 4: Proposition 4 and Theorem 1 are related to Theorem 15 in Toda (2014) on geometric sums. The latter result establishes that the Laplace distribution provides a small p asymptotic approximation to the distribution of a sum of T_p weakly dependent random variables, where T_p is a geometric random variable with success probability p. It implies a double Pareto approximation to the stationary distribution of S_t when the survival probability does not vary with types and is close to one. Proposition 4 provides the exact MGF of $\log S_t$, while Theorem 1 shows that the tail probabilities of S_t decay at Pareto rates determined by the positive and negative solutions to $\rho(\mathrm{A}(z)) = 1$, with neither result relying on asymptotic approximation or type-invariance of survival rates.

EXAMPLE 1: Suppose that there are two states ($N = 2$). By the Perron–Frobenius theorem, $\rho(\mathrm{A}(z))$ is the maximum real eigenvalue of $\mathrm{A}(z)$, and so by applying the usual formula for the eigenvalues of a 2×2 matrix (see, e.g., Horn and Johnson (2013, p. 39)) we compute

$$\rho\big(\mathrm{A}(z)\big) = \frac{1}{2}\Big(a_{11}(z) + a_{22}(z) + \sqrt{\big(a_{11}(z) - a_{22}(z)\big)^2 + 4a_{12}(z)a_{21}(z)}\,\Big), \qquad (12)$$

where $a_{nn'}(z) := \pi_{nn'}\upsilon_{nn'}\psi_{nn'}(z)$. Setting $\rho(\mathrm{A}(z))$ equal to one and solving for z gives a unique positive solution $z = \alpha$ if the conditions in Proposition 2 are satisfied. To illustrate concretely, suppose that

$$\Pi = \begin{bmatrix} 1 - \pi_{12} & \pi_{12} \\ \pi_{12} & 1 - \pi_{12} \end{bmatrix}$$

for some $\pi_{12} \in (0, 1)$, that $\upsilon_{nn'} = \upsilon = 0.95$, and that each $\psi_{nn'}$ is Gaussian, with locations $\mu_{11} = \mu_{21} = 0.03$ and $\mu_{12} = \mu_{22} = 0.01$ and scales $\sigma_{nn'} = \sigma = 0.01$. We graph the unique positive solution $z = \alpha$ to $\rho(\mathrm{A}(z)) = 1$ as a function of π_{12} in Figure 3(a). The graph shows that the Pareto exponent α increases smoothly from around 1.7 to around 2.56 as the transition probability π_{12} increases from zero to one. The fact that α is increasing in π_{12} is consistent with the comparative statics presented in Appendix B and previewed in Remark 3. The lower limit of approximately 1.7 is the value of z solving the equation $\upsilon\psi_{11}(z) = 1$; that is, it is the Pareto exponent for an agent always of the higher growth type. The upper limit of approximately 2.56 is the value of z solving the equation $\upsilon\psi(z) = 1$, where ψ is an equally weighted mixture of ψ_{12} and ψ_{21}. We will discuss Figure 3(b) when we return to this example in Section 5.

Examples 3.5 and 3.6 in Beare, Seo, and Toda (2022) concern a continuous-time reformulation of Example 1 in which the log-size of agents evolves as a two-state Brownian motion with drift. There the determination of the Pareto exponent boils down to examining the roots of a quartic polynomial, or of a quadratic polynomial in the absence of a diffusive component. The latter case corresponds closely to the economic model in Cao and Luo (2017), in which the upper Pareto exponent for the stationary distribution of wealth is given by the unique positive root of a quadratic polynomial.
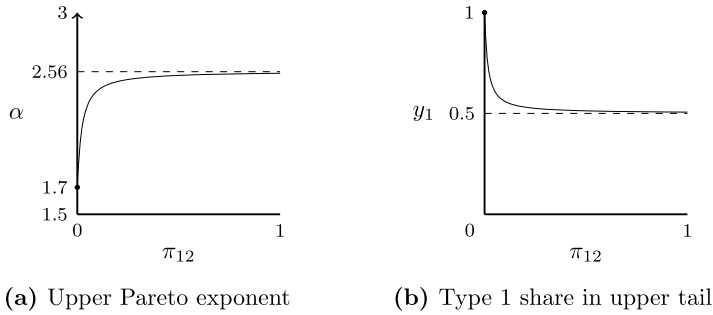
(a) Upper Pareto exponent   (b) Type 1 share in upper tail

FIGURE 3.—Upper Pareto exponent and type 1 share in Example 1.

5. REFINEMENTS UNDER A NONLATTICE CONDITION

While Theorem 1 establishes Pareto decay of the upper tail of the stationary distribution of S_t in the sense of there being an asymptotically linear relationship in log-log scale between the tail probability $P(S_t > s)$ and threshold s with slope $-\alpha$, it is notable that Theorem 1 does not assert the convergence of $s^\alpha P(S_t > s)$ to a positive and finite limit. Such convergence does not hold in general. We show in this section that it obtains under a nonlattice condition on growth rates. We also establish a characterization of upper-tail type shares that obtains under the nonlattice condition.

ASSUMPTION 2: *There do not exist constants $d > 0$ and $c_{nn'}$, $n \neq n'$, such that (i)* $\mathrm{supp}(\log G_{nn}) \subset d\mathbb{Z}$ *for each $n \in \mathcal{N}$, and (ii)* $\mathrm{supp}(\log G_{nn'}) \subset c_{nn'} + d\mathbb{Z}$ *for each $n, n' \in \mathcal{N}$ with $n \neq n'$.*

A simple sufficient (but not necessary) condition for Assumption 2 is that at least one growth rate is not a discrete random variable. The following result strengthens the conclusion of Theorem 1 when Assumption 2 is satisfied, and also shows that in this case a left eigenvector of $A(\alpha)$ characterizes type shares in the upper tail of S_t.

THEOREM 2: *Let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be a stationary Markov multiplicative process with reset satisfying Assumptions 1 and 2. If the equation $\rho(A(z)) = 1$ admits a unique positive solution $z = \alpha$ in the interior of $\mathcal{I}$, then $s^\alpha P(S_t > s)$ converges to a positive and finite limit as $s \to \infty$, so that S_t has a Pareto upper tail with exponent α. Moreover, $A(\alpha)$ has a unique left eigenvector $y = (y_1, \ldots, y_N)^\top$ with strictly positive entries summing to one, and for each $n \in \mathcal{N}$ we have*

$$\lim_{s \to \infty} P(J_t = n \mid S_t > s) = y_n. \tag{13}$$

We illustrate the statement about limiting type shares in Theorem 2 by revisiting Example 1.

EXAMPLE 1—continued: At $z = \alpha$, a left eigenvector with strictly positive entries associated with the maximum real eigenvalue computed in (12) is

$$\tilde{y} = \begin{bmatrix} \dfrac{\frac{1}{2}\left(a_{11}(\alpha) - a_{22}(\alpha) + \sqrt{\left(a_{11}(\alpha) - a_{22}(\alpha)\right)^2 + 4a_{12}(\alpha)a_{21}(\alpha)}\right)}{a_{12}(\alpha)} \end{bmatrix}.$$

Scaling the entries of $\tilde{y}$ such that they sum to one yields the left eigenvector y referred to in Theorem 2. In Figure 3(b), adopting the conditionally Gaussian parametrization used in our earlier discussion of this example, we graph y_1, the share of type 1 agents in the upper tail of the size distribution, as a function of the transition probability π_{12}. The graph shows that the share of type 1 agents in the upper tail decreases smoothly from one to one half as the transition probability π_{12} increases from zero to one.

The following simple example shows that the conclusions of Theorem 2 need not be satisfied if Assumption 2 is dropped.

EXAMPLE 1: Suppose again that $N = 2$, and let (S_t, J_t) be a Markov multiplicative process with reset on $(0, \infty) \times \{1, 2\}$ with $G_{11} = G_{22} = 1$, $G_{12} = 4$ and $G_{21} = 2$, and with survival rate $v \in (0, 1)$ constant across states. Suppose that we set $S_t = 1$ and $J_t = 1$ upon reset, and that in the absence of reset J_t has transition probability matrix

$$\Pi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

These assumptions fully specify the transition probabilities of (S_t, J_t). Noting that Assumption 1 is satisfied, we deduce from Proposition 3 that (S_t, J_t) has a unique stationary distribution, and assume it to be initialized at this distribution.

Assumption 2 is not satisfied in this example, as may be seen by taking $d = 1$ and $c_{12} = 2c_{21} = 2\log 2$. We show in Appendix A that the equation $\rho(A(z)) = 1$ admits a unique positive solution $z = \alpha = -(2/3)\log_2 v$, and that

$$v = \liminf_{s \to \infty} s^\alpha P(S_t > s) < \limsup_{s \to \infty} s^\alpha P(S_t > s) = v^{-1/3} \quad \text{and}$$

$$\frac{v}{1+v} = \liminf_{s \to \infty} P(J_t = 1 \mid S_t > s) < \limsup_{s \to \infty} P(J_t = 1 \mid S_t > s) = \frac{1}{1+v}.$$

The conclusions of Theorem 2 therefore do not hold. Nevertheless, Theorem 1 implies that

$$\lim_{s \to \infty} \frac{\log P(S_t > s)}{\log s} = -\alpha = \frac{2}{3}\log_2 v.$$

Example 1 illustrates the fragility of Theorem 2. Since any collection of growth rate distributions satisfying Assumption 2 can be well approximated by a collection of growth rate distributions not satisfying it, and vice versa, it is natural to be skeptical of results that depend on this condition for their validity. Nonlattice conditions have been used elsewhere to establish Pareto tail behavior of stationary solutions to stochastic difference equations.[4] An advantage of our approach is that our primary result, Theorem 1, does not rely on any nonlattice condition for its validity. The form of Pareto tail decay given in (10) is in this sense robust.

---

[4]See, for instance, (1.11) in Theorem A of Kesten (1973), (2) in Theorems 1 and 2 of de Saporta (2005), (A7) in Assumption 1.2 of Roitershtein (2007), and the "spread out" condition on pages 1408–1409 of Collamore (2009); and in an economic application, see footnote 56 in Benhabib, Bisin, and Zhu (2011). Note also that results of this kind typically exclude the possibility of reset; see (1.9) in Theorem A of Kesten (1973), the requirement that zero be excluded from the state space in Theorems 1 and 2 of de Saporta (2005), (A5) in Assumption 1.2 of Roitershtein (2007), and the requirement that $\log A_n$ be finite-valued on page 1408 of Collamore (2009).

6. FINAL REMARKS

Given the importance of persistent heterogeneity in productivity types for constructing plausibly calibrated economic models, we expect to see modelers adopting this feature more widely in future, and hope that the results we have presented here may prove useful to them. We conclude by outlining three potential generalizations of our results, which may broaden the scope of applications.

First, it may be useful to pursue a relaxation of our assumption of irreducibility. The evolution of types may follow a reducible Markov chain if, for instance, the health of agents randomly and irreversibly deteriorates over time, affecting their access to different productivity types. We rely on irreducibility at various points in our technical arguments. In particular, in the proof of Theorem 3 in Appendix A, which is used to prove Theorems 1 and 2, irreducibility is critical to establishing that poles in the MGF of log-size determining the Pareto exponents are simple. The reducible case may require a more general treatment of these poles.

Second, it may be useful to pursue a relaxation of our assumption that the number of types is finite. This would allow, for instance, productivity to be modelled as a general autoregressive process. Such a generalization would likely entail adapting our numerous arguments involving matrices such that they apply with linear operators on infinite dimensional spaces.

Third, it may be useful to generalize our results such that they apply when the random growth of agents is only asymptotically, rather than exactly, multiplicative. For instance, while capital income is naturally associated with multiplicative growth in wealth, the effect of labor income on wealth is additive. In an economy in which agents randomly accrue both capital and labor income, the growth in the wealth of the wealthiest agents is predominantly driven by capital income, and thus approximately multiplicative. Since it is the wealthiest agents which determine the upper Pareto exponent, we expect that our results on the upper Pareto exponent may be applied in a setting of this sort. This has been done in Gouin-Bonenfant and Toda (2022), with heuristic justification. A rigorous demonstration of the validity of our characterization of the upper Pareto exponent under asymptotically multiplicative growth would place such applications on firmer footing.

APPENDIX A: PROOFS OF CLAIMS IN SECTIONS 4 AND 5

PROOF OF PROPOSITION 1: Since each entry of $A(z)$ is a nonnegative multiple of a MGF finite-valued on $\mathcal{I}$, a result of Kingman (1961) guarantees that the spectral radius $\rho(A(z))$ is a convex function of $z \in \mathcal{I}$. And since $A(0)$ is nonnegative and bounded entrywise by Π, which is nonnegative with row sums of one, Theorems 8.1.22 and 8.4.5 in Horn and Johnson (2013) imply that $\rho(A(0)) \leq \rho(\Pi) = 1$. Under Assumption 1, $A(0)$ is nonnegative and irreducible with at least one entry strictly less than the corresponding entry of Π, so problem 15 on page 515 in Horn and Johnson (2013) implies that $\rho(A(0)) < \rho(\Pi)$; thus in this case $\rho(A(0)) < 1$. Since $\rho(A(z))$ is convex on $\mathcal{I}$, under Assumption 1 there can therefore be at most one positive and one negative solution to the equation $\rho(A(z)) = 1$ in $\mathcal{I}$.   *Q.E.D.*

PROOF OF PROPOSITION 2: The assumption that $E(G_{nn'}^z) < \infty$ for all $z > 0$ and all $n, n' \in \mathcal{N}$ guarantees that $[0, \infty) \subset \mathcal{I}$ and that $\rho(A(z)) < \infty$ for all $z > 0$, while the assumption that $P(G_{nn} > 1) > 0$ guarantees that $\psi_{nn}(z) \to \infty$ as $z \to \infty$. We therefore have

$$\rho\big(A(z)\big) \geq \rho\big(\mathrm{diag}(A(z))\big) = \max\big\{\pi_{nn} v_{nn} \psi_{nn}(z) : n \in \mathcal{N}\big\} \to \infty$$
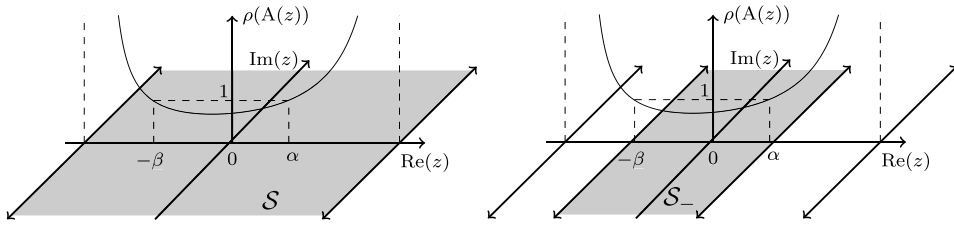
FIGURE A.1.—The strips $\mathcal{S}$ and $\mathcal{S}_-$ in the complex plane.

as $z \to \infty$, using a monotonicity property of the spectral radius (Horn and Johnson (2013, Theorem 8.1.18)) to obtain the inequality, and noting that $\pi_{nn} v_{nn} > 0$ under Assumption 1(iii). Under Assumption 1, Proposition 1 implies that $\rho(\mathrm{A}(0)) < 1$ and that $\rho(\mathrm{A}(z))$ is convex, hence continuous, on the interior of $\mathcal{I}$. It follows from the intermediate value theorem that $\rho(\mathrm{A}(z)) = 1$ for some positive z in the interior of $\mathcal{I}$.   Q.E.D.

PROOF OF PROPOSITION 3: It will be convenient to augment the state space $\mathcal{N}$ such that we are able to keep track of when reset occurs. To this end, let $\widetilde{\mathcal{N}} = \{\widetilde{n}_1, \ldots, \widetilde{n}_{2N}\}$, where

$$\widetilde{n}_i = \begin{cases} (i, 0) & \text{for } i = 1, \ldots, N, \\ (i - N, 1) & \text{for } i = N + 1, \ldots, 2N. \end{cases}$$

Let K_t be equal to one if there is reset at time t, or equal to zero otherwise. The sequence of pairs $(J_t, K_t)_{t \in \mathbb{Z}_+}$ is then a Markov chain on $\widetilde{\mathcal{N}}$. Fix $n \in \mathcal{N}$ with $\varpi_n > 0$. Under Assumption 1, since $\varpi_n > 0$, the state $(n, 1)$ is accessible from any other state; therefore, since there are finitely many states, the state $(n, 1)$ must be positive recurrent (Douc, Moulines, Priouret, and Soulier (2018, Corollary 7.2.3)).

Consider the singleton $\mathcal{R}_n := \{1\} \times \{n\}$. Since $(S_t, J_t) \in \mathcal{R}_n$ whenever $(J_t, K_t) = (n, 1)$, the fact that $(n, 1)$ is a positive recurrent accessible state of $(J_t, K_t)_{t \in \mathbb{Z}_+}$ implies that $\mathcal{R}_n$ is a positive recurrent accessible atom of $(S_t, J_t)_{t \in \mathbb{Z}_+}$. It therefore follows from Theorem 6.4.2 in Douc et al. (2018) that there exists a unique probability distribution for (S_0, J_0) such that $(S_t, J_t)_{t \in \mathbb{Z}_+}$ is stationary.   Q.E.D.

Until now, we have regarded $\mathrm{A}(z)$ to be a function of a real variable $z \in \mathcal{I}$, where $\mathcal{I}$ was defined to be the smallest convex subset of the real line on which all of the MGFs $\psi_{nn'}(z)$ are finite. In what follows, it will be more useful to regard $\mathrm{A}(z)$ and each of the MGFs $\psi_{nn'}(z)$ to be functions of a complex variable. We therefore extend $\mathcal{I}$ to a strip in the complex plane by setting $\mathcal{S} = \{z \in \mathbb{C} : \mathrm{Re}(z) \in \mathcal{I}\}$. Owing to the fact that

$$\int |s^z| \, d\Phi_{nn'}(s) = \int s^{\mathrm{Re}(z)} \, d\Phi_{nn'}(s) = \psi_{nn'}(\mathrm{Re}(z)), \tag{14}$$

it is assured that each $\psi_{nn'}(z) := \int s^z \, d\Phi_{nn'}(s)$ is well-defined as a complex-valued function of $z \in \mathcal{S}$, and that $\mathrm{A}(z)$ is well-defined as a complex matrix-valued function of $z \in \mathcal{S}$. The functions $\psi_{nn'}(z)$ are called Mellin–Stieltjes transforms of the corresponding CDFs $\Phi_{nn'}(s)$. We also extend $\mathcal{I}_-$ to a strip in the complex plane by setting $\mathcal{S}_- = \{z \in \mathbb{C} : \mathrm{Re}(z) \in \mathcal{I}_-\}$. Figure A.1 illustrates $\mathcal{S}$ and $\mathcal{S}_-$.

LEMMA 1: *For each $z \in \mathcal{S}_-$, we have $\rho(\mathrm{A}(z)) < 1$.*

PROOF OF LEMMA 1: In view of (14), the entries of $A(z)$ are bounded in modulus by the corresponding entries of $A(\mathrm{Re}(z))$ for all $z \in \mathcal{S}$. Using a monotonicity property of the spectral radius (Horn and Johnson (2013, Theorem 8.1.18)), we therefore have $\rho(A(z)) \leq \rho(A(\mathrm{Re}(z)))$ for all $z \in \mathcal{S}$. The desired result now follows from the definitions of $\mathcal{I}_-$ and $\mathcal{S}_-$.   Q.E.D.

The quantity $\mathrm{E}(S_t^z)$, viewed as a function of the complex variable z, is called the Mellin transform of the distribution of S_t. The following lemma provides a formula for $\mathrm{E}(S_t^z \mathbb{1}(J_t = n))$ valid for each $z \in \mathcal{S}_-$.

LEMMA 2: *Let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be a stationary Markov chain on $(0, \infty) \times \mathcal{N}$ with Markov kernel given by (3). Then, for each $z \in \mathcal{S}_-$, the matrix $I - A(z)$ is invertible and*

$$\big[\mathrm{E}\big(S_t^z \mathbb{1}(J_t = 1)\big) \quad \cdots \quad \mathrm{E}\big(S_t^z \mathbb{1}(J_t = N)\big)\big] = r\varpi^\top \big(I - A(z)\big)^{-1}. \tag{15}$$

PROOF OF LEMMA 2: Let $F_n(s) := \mathrm{P}(S_t \leq s, J_t = n)$. Then, using (3),

$$F_{n'}(s') = \sum_n \int \mathrm{P}\big(S_{t+1} \leq s', J_{t+1} = n' \mid S_t = s, J_t = n\big) \, \mathrm{d}F_n(s)$$

$$= \sum_n \pi_{nn'} v_{nn'} \int \Phi_{nn'}(s'/s) \, \mathrm{d}F_n(s)$$

$$+ \mathbb{1}(s' \geq 1)\varpi_{n'} \sum_{n,n''} p_n \pi_{nn''}(1 - v_{nn''}). \tag{16}$$

Noting that $\int (s')^z \Phi_{nn'}(\mathrm{d}s'/s) = s^z \psi_{nn'}(z)$ for each $z \in \mathcal{S}$, and using the definition of the unconditional reset probability r, we take the Mellin–Stieltjes transform of either side of (16) to obtain

$$\int s^z \, \mathrm{d}F_{n'}(s) = \sum_n \pi_{nn'} v_{nn'} \psi_{nn'}(z) \int s^z \, \mathrm{d}F_n(s) + r\varpi_{n'}. \tag{17}$$

Letting $L(z)$ denote the $N \times 1$ vector with nth entry

$$\mathrm{E}\big(S_t^z \mathbb{1}(J_t = n)\big) = \int s^z \, \mathrm{d}F_n(s),$$

we may rewrite (17) using matrix notation as $L(z) = A(z)^\top L(z) + r\varpi$ or, equivalently,

$$L(z)^\top \big(I - A(z)\big) = r\varpi^\top. \tag{18}$$

Equation (18) is valid for each $z \in \mathcal{S}$. Lemma 1 implies that the matrix $I - A(z)$ is invertible for each $z \in \mathcal{S}_-$. We may therefore post-multiply both sides of (18) by $(I - A(z))^{-1}$ to obtain (15), valid for each $z \in \mathcal{S}_-$.   Q.E.D.

PROOF OF PROPOSITION 4: The result follows immediately from Lemma 2.   Q.E.D.

THEOREM 3: *Suppose that Assumption 1 holds, and let $(S_t, J_t)_{t \in \mathbb{Z}_+}$ be a stationary homogeneous Markov chain on $(0, \infty) \times \mathcal{N}$ with Markov kernel given by (3). Suppose further that*

the equation $\rho(A(z)) = 1$ admits a unique positive solution $z = \alpha$ in the interior of $\mathcal{I}$. Then $A(\alpha)$ has an algebraically simple unit eigenvalue, and the associated right and left eigenspaces are spanned by unique right and left eigenvectors x and y with positive entries summing to one. Let

$$C = \frac{r\varpi^\top x}{y^\top A'(\alpha)x},$$

where $A'(\alpha)$ is the matrix of complex derivatives of $A(z)$ at $z = \alpha$. Let B be the supremum of all $b > 0$ such that $I - A(\alpha + i\tau)$ is invertible for $\tau \in (-b, b)$ except at $\tau = 0$. Then $B \in (0, \infty]$, $C \in (0, \infty)$, and for each $n \in \mathcal{N}$,

$$\frac{2\pi C/B}{e^{2\pi\alpha/B} - 1} y_n \le \liminf_{s\to\infty} s^\alpha P(S_t > s, J_t = n)$$

$$\le \limsup_{s\to\infty} s^\alpha P(S_t > s, J_t = n) \le \frac{2\pi C/B}{1 - e^{-2\pi\alpha/B}} y_n \tag{19}$$

if $B < \infty$, or otherwise

$$\lim_{s\to\infty} s^\alpha P(S_t > s, J_t = n) = \frac{Cy_n}{\alpha}. \tag{20}$$

PROOF OF THEOREM 3: Let Ω denote the interior of $\mathcal{S}$, an open and connected subset of the complex plane, nonempty since it contains α. The matrix-valued function $A(z)$ is holomorphic on Ω. Since $\rho(A(z))$ is a convex function of $z \in \mathcal{I}$ with $\rho(A(0)) < 1$ (Proposition 1), and $\rho(A(\alpha)) = 1$, it must be the case that $I - A(z)$ is invertible at all real numbers between 0 and α, which are elements of Ω. It therefore follows from Theorem A.2 in Beare, Seo, and Toda (2022) that $I - A(z)$ has a meromorphic inverse on Ω, with poles at the points of noninvertibility of $I - A(z)$. Since $A(\alpha)$ is nonnegative, we know from the Perron–Frobenius theorem that its spectral radius of one is an eigenvalue. Consequently, zero is an eigenvalue of $I - A(\alpha)$. We deduce that $I - A(z)$ is not invertible at $\alpha \in \Omega$, and that α is a pole of $(I - A(z))^{-1}$. Since every pole is an isolated singularity, it follows immediately that $B > 0$.

We now show that α is a simple pole of $(I - A(z))^{-1}$, and determine the associated residue. Since $A(\alpha)$ is nonnegative and irreducible under Assumption 1, it follows from the Perron–Frobenius theorem that the unit eigenvalue of $A(\alpha)$ is algebraically simple, hence geometrically simple, and is associated with unique right and left eigenvectors x, y with positive entries summing to one. From Theorem A.2 in Beare, Seo, and Toda (2022), we thus deduce that α is a simple pole of $(I - A(z))^{-1}$, with residue

$$R := -x(y^\top A'(\alpha)x)^{-1}y^\top = -cxy^\top,$$

where $c := (y^\top A'(\alpha)x)^{-1}$ is a positive real number.

Let $e^{(n)}$ denote an $N \times 1$ vector with nth entry equal to one and other entries equal to zero. By Lemma 2, we have

$$E(S_t^z \mid J_t = n) = (r/p_n)\varpi^\top(I - A(z))^{-1}e^{(n)} \tag{21}$$

for $z \in \mathcal{S}_-$, where r is the unconditional probability of reset and $p_n = P(J_t = n)$. Moreover, since $(I - A(z))^{-1}$ is meromorphic on Ω, (21) defines a meromorphic extension of $E(S_t^z \mid$

$J_t = n$) to Ω. As $z \to \alpha$, we obtain

$$
(z - \alpha) \mathrm{E}\big(S_t^z \mid J_t = n\big) = (r/p_n)\boldsymbol{\varpi}^\top (z - \alpha)\big(\mathrm{I} - \mathrm{A}(z)\big)^{-1} e^{(n)}
$$
$$
\to (r/p_n)\boldsymbol{\varpi}^\top R e^{(n)} = -rc\big(\boldsymbol{\varpi}^\top x\big) y_n / p_n = -C y_n / p_n.
$$

The eigenvectors x and y have positive entries, and $\boldsymbol{\varpi}$ has nonnegative entries with at least one entry positive, so $\boldsymbol{\varpi}^\top x > 0$ and $y_n > 0$. The unconditional reset probability r is positive under Assumption 1, and as noted above, $c > 0$, so also $C > 0$. Therefore, $\lim_{z \to \alpha}(z - \alpha)\mathrm{E}(S_t^z \mid J_t = n) \neq 0$. This shows that α is a simple pole of $\mathrm{E}(S_t^z \mid J_t = n)$ with residue $-C y_n / p_n$. Letting B' denote the supremum of all $b > 0$ such that α is the unique singularity of $\mathrm{E}(S_t^z \mid J_t = n)$ on $\{\alpha + i\tau : \tau \in (-b, b)\}$, it now follows from Theorem A.1 in Beare, Seo, and Toda (2022) that

$$
\frac{2\pi C/B'}{\mathrm{e}^{2\pi\alpha/B'} - 1} y_n / p_n \leq \liminf_{s \to \infty} s^\alpha \mathrm{P}(S_t > s \mid J_t = n)
$$
$$
\leq \limsup_{s \to \infty} s^\alpha \mathrm{P}(S_t > s \mid J_t = n) \leq \frac{2\pi C/B'}{1 - \mathrm{e}^{-2\pi\alpha/B'}} y_n / p_n \tag{22}
$$

if $B' < \infty$, or otherwise

$$
\lim_{s \to \infty} s^\alpha \mathrm{P}(S_t > s \mid J_t = n) = \frac{C y_n}{\alpha p_n}. \tag{23}
$$

It is apparent from (21) that if $\alpha + i\tau$ is a singularity of $\mathrm{E}(S_t^z \mid J_t = n)$ then it is also a singularity of $(\mathrm{I} - \mathrm{A}(z))^{-1}$. Thus $B \leq B'$. If $B = \infty$ then also $B' = \infty$, and so (20) follows from (23). If $B < \infty$ and $B' < \infty$, then (19) follows from (22) by noting that the lower (upper) bound in (22) is increasing (decreasing) in B'. If $B < \infty$ and $B' = \infty$, then (19) follows from (23) by noting that $C y_n / \alpha$ falls between the lower and upper bounds in (19). $\qquad$ *Q.E.D.*

REMARK 5: Theorems A.1 and A.2 in Beare, Seo, and Toda (2022), which play a critical role in the proof of Theorem 3, are not novel to that paper but are stated there in a way which is convenient for our purposes. The role played by Theorem A.1 is particularly important. This result, due to Nakagawa (2007), is the source of the tail probability bounds in (19).

PROOF OF THEOREM 1: Theorem 3 establishes that $s^\alpha \mathrm{P}(S_t > s)$ has positive and finite limits inferior and superior. There thus exist positive and finite constants c_- and c_+ such that $c_- \leq s^\alpha \mathrm{P}(S_t > s) \leq c_+$ for all sufficiently large s. Taking logarithms, dividing by $\log s$, and letting $s \to \infty$, we obtain (10). $\qquad$ *Q.E.D.*

LEMMA 3: *Let X be a real random variable and α a positive real number such that $\mathrm{E}(\mathrm{e}^{\alpha X}) < \infty$, and let τ be a nonzero real number. Then $|\mathrm{E}(\mathrm{e}^{(\alpha + i\tau)X})| = \mathrm{E}(\mathrm{e}^{\alpha X})$ if and only if $\mathrm{supp}(X) \subset c + (2\pi/\tau)\mathbb{Z}$ for some $c \in \mathbb{R}$. Moreover, $\mathrm{E}(\mathrm{e}^{(\alpha + i\tau)X}) = \mathrm{E}(\mathrm{e}^{\alpha X})$ if and only if $\mathrm{supp}(X) \subset (2\pi/\tau)\mathbb{Z}$.*

PROOF OF LEMMA 3: The result follows from an obvious modification to the proofs of Lemmas 3 and 4 in Feller (1971, pp. 500–501). $\qquad$ *Q.E.D.*

PROOF OF THEOREM 2: The result follows from Theorem 3 if we can show that $B = \infty$; that is, that α is the unique point of noninvertibility of $I - A(z)$ on the axis $\mathrm{Re}(z) = \alpha$. Suppose to the contrary that $I - A(\alpha + i\tau)$ is not invertible for some nonzero $\tau \in \mathbb{R}$. In view of (14), the entries of $A(\alpha + i\tau)$ are bounded in modulus by the corresponding entries of $A(\alpha)$. Using a monotonicity property of the spectral radius (Horn and Johnson (2013, Theorem 8.1.18)), we deduce that $\rho(A(\alpha + i\tau)) \le \rho(A(\alpha)) = 1$. But noninvertibility of $I - A(\alpha + i\tau)$ implies that $A(\alpha + i\tau)$ has a unit eigenvalue, so we must have $\rho(A(\alpha + i\tau)) = \rho(A(\alpha)) = 1$. It now follows from Theorem 8.4.5 in Horn and Johnson (2013) that $A(\alpha + i\tau) = DA(\alpha)D^{-1}$, where $D = \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_N})$ for some $\theta_1, \ldots, \theta_N \in \mathbb{R}$. We may therefore write

$$\pi_{nn'} v_{nn'} \psi_{nn'}(\alpha + i\tau) = \pi_{nn'} v_{nn'} e^{i(\theta_n - \theta_{n'})} \psi_{nn'}(\alpha)$$

for all $n, n' \in \mathcal{N}$. Consequently, $\psi_{nn}(\alpha + i\tau) = \psi_{nn}(\alpha)$ for all $n \in \mathcal{N}$ such that $\pi_{nn} v_{nn} > 0$, and $|\psi_{nn'}(\alpha + i\tau)| = \psi_{nn'}(\alpha)$ for all $n, n' \in \mathcal{N}$ such that $\pi_{nn'} v_{nn'} > 0$. Moreover, for any $n, n' \in \mathcal{N}$ such that $\pi_{nn'} v_{nn'} = 0$, we trivially have $\psi_{nn'}(\alpha + i\tau) = \psi_{nn'}(\alpha) = 1$ due to the normalization in Assumption 1(iii). It therefore follows from Lemma 3 that, for all $n, n' \in \mathcal{N}$, we have $\mathrm{supp}(\log G_{nn}) \subset (2\pi/\tau)\mathbb{Z}$ and $\mathrm{supp}(\log G_{nn'}) \subset c_{nn'} + (2\pi/\tau)\mathbb{Z}$ for some $c_{nn'} \in \mathbb{R}$. But this violates Assumption 2, so we conclude that $I - A(\alpha + i\tau)$ must be invertible for all nonzero $\tau \in \mathbb{R}$.   Q.E.D.

DETAILS OF EXAMPLE 1: Observe that

$$\rho(A(z)) = v\rho\left(\begin{bmatrix} 0 & 4^z \\ 2^z & 0 \end{bmatrix}\right) = 2^{3z/2} v.$$

It follows that the unique positive solution to the equation $\rho(A(z)) = 1$ is $z = -(2/3)\log_2 v$, and so we deduce from Theorem 1 that S_t has a Pareto upper tail with exponent $\alpha = -(2/3)\log_2 v$ in the sense of (10).

Assumption 2 is not satisfied in this example. We now show that the conclusions of Theorem 2 do not hold. If reset occurs in period τ, then in the absence of further reset, the values taken by $S_{\tau+k}$ for $k \ge 0$ are given by the (deterministic) sequence $1, 4, 8, 32, 64, 256, 512, \ldots$. Therefore, if in period t the most recent reset occurred in period $t - k$, then S_t is equal to the $(k+1)$th entry in that sequence, which is $2^{\lfloor (3k+1)/2 \rfloor}$. Since the probability that there have been exactly k periods since the most recent reset[5] is $v^k(1 - v)$, we find that

$$P(S_t > s) = \sum_{k=0}^{\infty} v^k(1 - v)\mathbb{1}(2^{\lfloor (3k+1)/2 \rfloor} > s). \tag{24}$$

Therefore, if $n - 1 \le \log_2 s < n$, then

$$P(S_t > s) = \sum_{k=\lceil (2n-1)/3 \rceil}^{\infty} v^k(1 - v) = v^{\lceil (2n-1)/3 \rceil} = v^{\lceil (2\lfloor \log_2 s \rfloor + 1)/3 \rceil}. \tag{25}$$

---

[5]Here, we implicitly extend (S_t, J_t) to a stationary Markov process indexed by $t \in \mathbb{Z}$ rather than $t \in \mathbb{Z}_+$, so that the number of periods since the most recent reset is unbounded.

Multiplying by s^α with $\alpha = -(2/3)\log_2 v$, we obtain

$$s^\alpha P(S_t > s) = s^{-(2/3)\log_2 v} v^{\lceil (2\lfloor \log_2 s\rfloor +1)/3\rceil} = v^{-(2/3)\log_2 s + \lceil (2\lfloor \log_2 s\rfloor +1)/3\rceil},$$

which oscillates between v and $v^{-1/3}$ as $s \to \infty$. Thus, while (10) is satisfied, $s^\alpha P(S_t > s)$ does not converge to any limit as $s \to \infty$.

The failure of Assumption 2 also causes the convergence (13) in Theorem 2 to fail in this example. To see why, observe first that, since the Markov state switches every period, summing the even-indexed terms in the series on the right-hand side of (24) gives

$$P(S_t > s, J_t = 1) = \sum_{k=0}^{\infty} v^{2k}(1 - v)\mathbb{1}\big(2^{3k} > s\big).$$

Therefore, if $n - 1 \le \log_2 s < n$, then

$$P(S_t > s, J_t = 1) = \sum_{k=\lceil n/3\rceil}^{\infty} v^{2k}(1 - v) = \frac{v^{2\lceil n/3\rceil}}{1+v} = \frac{v^{2\lceil (\lfloor \log_2 s\rfloor +1)/3\rceil}}{1+v}. \tag{26}$$

Dividing (26) by (25), we obtain

$$P(J_t = 1 \mid S_t > s) = \frac{v^{2\lceil (\lfloor \log_2 s\rfloor +1)/3\rceil - \lceil (2\lfloor \log_2 s\rfloor +1)/3\rceil}}{1+v},$$

which oscillates between $v/(1 + v)$ and $1/(1 + v)$ as $s \to \infty$. Thus $P(J_t = 1 \mid S_t > s)$ does not converge to any limit as $s \to \infty$. Similarly, $P(J_t = 2 \mid S_t > s)$ does not converge to any limit as $s \to \infty$. *Q.E.D.*

APPENDIX B: COMPARATIVE STATICS

We consider linear perturbations to the survival probabilities, to the parameters of location-scale transformations of the growth rates, and to a measure of the degree of persistence in the Markov modulator. Given $\mu_{nn'} \in \mathbb{R}$ and $\sigma_{nn'} > 0$, we denote the MGF of $\mu_{nn'} + \sigma_{nn'}\log G_{nn'}$ by $\psi_{nn'}(z; \mu_{nn'}, \sigma_{nn'}) = e^{z\mu_{nn'}}\psi_{nn'}(\sigma_{nn'}z)$, and normalize $\log G_{nn'}$ to have mean zero. Let $\Psi(z; M, \Sigma)$ be the $N \times N$ matrix of MGFs $\psi_{nn'}(z; \mu_{nn'}, \sigma_{nn'})$ parametrized by $M = (\mu_{nn'})$ and $\Sigma = (\sigma_{nn'})$. We parametrize persistence by setting $\Pi(\tau) = \tau I + (1 - \tau)\Pi$ for $\tau \in [0, 1]$. Increases (decreases) in τ are interpreted as increases (decreases) in persistence.

Collect the parameters to be perturbed into a single vector of parameters,

$$\theta = (Y, M, \Sigma, \tau) \in (0, 1)^{N \times N} \times \mathbb{R}^{N \times N} \times (0, \infty)^{N \times N} \times (0, 1) =: \Theta,$$

and let $A(z; \theta) = \Pi(\tau) \odot Y \odot \Psi(z; M, \Sigma)$. The following result, an application of the implicit function theorem, shows how perturbations to θ affect α.

PROPOSITION 5: *Suppose the vector of parameters $\theta_0 = (Y_0, M_0, \Sigma_0, \tau_0) \in \Theta$ is such that $\Pi(\tau_0)$, Y_0 and $\Psi(z; M_0, \Sigma_0)$ satisfy Assumption 1, and the equation $\rho(A(z; \theta_0)) = 1$ admits a unique positive solution $z = \alpha_0$ in the interior of*

$$\mathcal{I}_0 := \big\{ z \in \mathbb{R} : \text{all entries of } \Psi(z; M_0, \Sigma_0) \text{ are finite} \big\}.$$

Then there exists a neighborhood $U \subset \Theta$ of θ_0 and a unique continuously differentiable function $\alpha : U \to (0, \infty)$ such that $\rho(A(\alpha(\theta); \theta)) = 1$ on U. The partial derivatives of α satisfy the following inequalities at θ_0:

(i) *$\partial \alpha / \partial \upsilon_{nn'} \leq 0$: Longer lifespan implies a smaller Pareto exponent.*
(ii) *$\partial \alpha / \partial \mu_{nn'} \leq 0$: Higher growth implies a smaller Pareto exponent.*
(iii) *$\partial \alpha / \partial \sigma_{nn'} \leq 0$: Higher volatility implies a smaller Pareto exponent.*
(iv) *If all columns or all rows of $\Upsilon_0 \odot \Psi(z; M_0, \Sigma_0)$ are the same, then $\partial \alpha / \partial \tau \leq 0$: Higher persistence implies a smaller Pareto exponent.*

REMARK 6: In Proposition 5(iv), the requirement that all columns or all rows of $\Upsilon_0 \odot \Psi(z; M_0, \Sigma_0)$ are the same means that the survival probability and growth rate distribution depend only on the current state, or only on the previous state. Without this condition, higher persistence need not imply a smaller Pareto exponent. As a counterexample, set $\mathcal{N} = \{1, 2\}$ and suppose that $G_{nn'} = 2$ if $n \neq n'$ or $G_{nn'} = 1$ if $n = n'$. Suppose that the transition probability matrix for J_t is

$$\Pi(\tau) = \tau I + (1 - \tau) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \tau & 1 - \tau \\ 1 - \tau & \tau \end{bmatrix},$$

and the survival probability $\upsilon \in (0, 1)$ is constant. Then

$$\Pi(\tau) \odot \Upsilon \odot \Psi(z) = \upsilon \begin{bmatrix} \tau & (1 - \tau)2^z \\ (1 - \tau)2^z & \tau \end{bmatrix},$$

whose spectral radius is $\upsilon(\tau + (1 - \tau)2^z)$. Setting the spectral radius equal to one and solving for $z > 0$, we find that the Pareto exponent for the upper tail of the stationary distribution of S_t is

$$\alpha(\tau) = \log_2 \left(\frac{1/\upsilon - \tau}{1 - \tau} \right) = \log_2 \left(1 + \frac{1/\upsilon - 1}{1 - \tau} \right),$$

which is *increasing* in $\tau \in (0, 1)$. Therefore, in this example, increasing the persistence makes the upper tail lighter.

PROOF OF PROPOSITION 5: Define $F : (0, \infty) \times \Theta \to [-1, \infty]$ by

$$F(z; \theta) = \rho(A(z; \theta)) - 1,$$

where the spectral radius is understood to be infinite whenever one or more entries of $\Psi(z; M, \Sigma)$ are infinite. Note that $F(\alpha_0; \theta_0) = 0$ and that, since α_0 lies in the interior of $\mathcal{I}_0$, $F(z; \theta)$ is finite on a neighborhood of (α_0, θ_0). To apply the implicit function theorem, we need F to be continuously differentiable on a neighborhood of (α_0, θ_0), with nonzero partial derivative $\partial F / \partial z$ at (α_0, θ_0). Continuous differentiability follows from the fact that the entries of $\Pi(\tau)$ and $\Psi(z; M, \Sigma)$ are continuously differentiable with respect to their parameters and z, and the fact that the spectral radius of a nonnegative irreducible matrix is continuously differentiable with respect to its entries (see, e.g., Vahrenkamp (1976)). Positivity of the partial derivative $\partial F / \partial z$ at (α_0, θ_0) follows from Proposition 1. The implicit function theorem thus guarantees the existence of a neighborhood $U \subset \Theta$ of θ_0 and a unique continuously differentiable function $\alpha : U \to (0, \infty)$ such that $\rho(A(\alpha(\theta); \theta)) = 1$

on U. The partial derivatives of α on U are then given by

$$\nabla_\theta \alpha = -\frac{1}{\partial F/\partial z} \nabla_\theta F.$$

It remains to show that the partial derivatives of α have the signs asserted in (i)–(iv). Since $\partial F/\partial z > 0$, and since the spectral radius of a nonnegative matrix is nondecreasing in its entries (Horn and Johnson (2013, Theorem 8.1.18), to show (i)–(iii) it suffices to show that the entries of $A(z; \theta)$ are nondecreasing in $v_{nn'}$, $\mu_{nn'}$, and $\sigma_{nn'}$. For $v_{nn'}$ this is obvious. For $\mu_{nn'}$, it follows from the fact that

$$\frac{\partial \psi_{nn'}(z; \mu_{nn'}, \sigma_{nn'})}{\partial \mu_{nn'}} = z e^{z\mu_{nn'}} \psi_{nn'}(\sigma_{nn'} z),$$

which is positive for $z > 0$. For $\sigma_{nn'}$, observe that since $\log G_{nn'}$ is normalized to have zero mean, its MGF $\psi_{nn'}(z)$ satisfies $\psi'_{nn'}(0) = 0$, and thus, by convexity, $\psi'_{nn'}(z) \geq 0$ for $z > 0$. The partial derivative

$$\frac{\partial \psi_{nn'}(z; \mu_{nn'}, \sigma_{nn'})}{\partial \sigma_{nn'}} = z e^{z\mu_{nn'}} \psi'_{nn'}(\sigma_{nn'} z)$$

is therefore nonnegative for $z > 0$, and so the entries of $A(z; \theta)$ are nondecreasing in $\sigma_{nn'}$. To show (iv), observe that if all rows of $Y_0 \odot \Psi(z; M_0, \Sigma_0)$ are the same, then we may write $A(z; \theta_0) = \Pi(\tau_0) D_0(z)$, where $D_0(z) = \mathrm{diag}(Y_0 \odot \Psi(z; M_0, \Sigma_0))$. On the other hand, if all columns of $Y_0 \odot \Psi(z; M_0, \Sigma_0)$ are the same, then we may write $A(z; \theta_0) = D_0(z) \Pi(\tau_0)$. Noting that $\rho(AB) = \rho(BA)$ in general (Horn and Johnson (2013, Theorem 1.3.22)), in either case we have $\rho(A(z; \theta_0)) = \rho((\tau_0 I + (1-\tau_0)\Pi) D_0(z))$. The fact that $\rho(A(z; \theta_0))$ is nondecreasing in τ_0 now follows from Theorem 5.2 in Karlin (1982); see also Theorem 1.1 in Altenberg (2013). Since $\partial F/\partial z > 0$, it follows that α is nonincreasing in τ at θ_0. This establishes (iv). *Q.E.D.*

REFERENCES

ALTENBERG, LEE (2013): "On the Ordering of Spectral Radius Product $r(\mathbf{A})r(\mathbf{AD})$ Versus $r(\mathbf{A}^2\mathbf{D})$ and Related Applications," *SIAM Journal on Matrix Analysis and Applications*, 34, 978–998. [1831]

BEARE, BRENDAN K., AND ALEXIS A. TODA (2020): "On the Emergence of a Power Law in the Distribution of COVID-19 Cases," *Physica D: Nonlinear Phenomena*, 412, 132649. [1812]

BEARE, BRENDAN K., WON-KI SEO, AND ALEXIS A. TODA (2022): "Tail Behavior of Stopped Lévy Processes With Markov Modulation," *Econometric Theory* (forthcoming). [1813,1820,1826,1827]

BENHABIB, JESS, AND ALBERTO BISIN (2018): "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56, 1261–1291. [1811]

BENHABIB, JESS, ALBERTO BISIN, AND SHENGHAO ZHU (2011): "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 79, 123–157. [1812,1822]

CAO, DAN, AND WENLAN LUO (2017): "Persistent Heterogeneous Returns and Top End Wealth Inequality," *Review of Economic Dynamics*, 26, 301–326. [1812,1820]

CHAMPERNOWNE, DAVID G. (1953): "A Model of Income Distribution," *Economic Journal*, 63, 318–351. [1811]

COLLAMORE, JEFFREY F. (2009): "Random Recurrence Equations and Ruin in a Markov-Dependent Stochastic Economic Environment," *Annals of Applied Probability*, 19, 1404–1458. [1822]

DE SAPORTA (2005): "Tail of the Stationary Solution of the Stochastic Equation $Y_{n+1} = a_n Y_n + b_n$ With Markovian Coefficients," *Stochastic Processes and their Applications*, 115, 519–578. [1822]

DEVADOSS, STEPHEN, JEFF LUCKSTEAD, DIANA DANFORTH, AND SHERZOD AKHUNDJANOV (2016): "The Power Law Distribution for Lower Tail Cities in India," *Physica A*, 442, 193–196. [1819]

DOUC, RANDAL, ERIC MOULINES, PIERRE PRIOURET, AND PHILIPPE SOULIER (2018): *Markov Chains*. Springer International Publishing. [1824]

EMBRECHTS, PAUL, CLAUDIA KLÜPPELBERG, AND THOMAS MIKOSCH (1997): *Modelling Extremal Events: for Insurance and Finance*. Heidelberg: Springer-Verlag. [1812]

FAGERENG, ANDREAS, LUIGI GUISO, DAVIDE MALACRINO, AND LUIGI PISTAFERRI (2020): "Heterogeneity and Persistence in Returns to Wealth," *Econometrica*, 88, 115–170. [1812]

FELLER, WILLIAM (1971): *An Introduction to Probability Theory and Its Applications, Vol. II*. New York: John Wiley & Sons. [1827]

GABAIX, XAVIER (1999): "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114, 739–767. [1811]

——— (2009): "Power Laws in Economics and Finance," *Annual Review of Economics*, 1, 255–293. [1811]

——— (2016): "Power Laws in Economics: An Introduction," *Journal of Economic Perspectives*, 30, 185–206. [1811]

GABAIX, XAVIER, AND RUSTAM IBRAGIMOV (2011): "Rank−1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," *Journal of Business and Economic Statistics*, 29, 24–39. [1819]

GABAIX, XAVIER, JEAN-MICHEL LASRY, PIERRE-LOUIS LIONS, AND BENJAMIN MOLL (2016): "The Dynamics of Inequality," *Econometrica*, 84, 2071–2111. [1812]

GIESEN, KRISTIAN, ARNDT ZIMMERMANN, AND JENS SUEDEKUM (2010): "The Size Distribution Across All Cities—Double Pareto Lognormal Strikes," *Journal of Urban Economics*, 68, 129–137. [1819]

GOMEZ, MATTHIEU, AND ÉMILIEN GOUIN-BONENFANT (2020): "A Q-Theory of Inequality," Unpublished Manuscript, Columbia University. [1813]

GOUIN-BONENFANT, ÉMILIEN (2020): "Productivity Dispersion, Between-Firm Competition, and the Labor Share," Unpublished Manuscript, Columbia University. [1813]

GOUIN-BONENFANT, ÉMILIEN, AND ALEXIS A. TODA (2022): "Pareto Extrapolation: An Analytical Framework for Studying Tail Inequality," *Quantitative Economics* (forthcoming). [1813,1816,1823]

HORN, ROGER A., AND CHARLES R. JOHNSON (2013): *Matrix Analysis* (Second Ed.). Cambridge, U.K.: Cambridge University Press. [1815,1820,1823-1825,1828,1831]

KARLIN, SAMUEL (1982): "Classifications of Selection-Migration Structures and Conditions for a Protected Polymorphism," in *Evolutionary Biology*, Vol. 14, ed. by Max K. Hecht, Bruce Wallace, and Ghillean T. Prance. New York: Plenum, 61–204. [1831]

KESTEN, HARRY (1973): "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Mathematica*, 131, 207–248. [1811,1822]

KINGMAN, JOHN F. C. (1961): "A Convexity Property of Positive Matrices," *Quarterly Journal of Mathematics*, 12, 283–284. [1823]

MA, QINGYIN, JOHN STACHURSKI, AND ALEXIS A. TODA (2020): "The Income Fluctuation Problem and the Evolution of Wealth," *Journal of Economic Theory*, 187, 105003. [1813]

MANRUBIA, SUSANNA C., AND DAMIÁN H. ZANETTE (1999): "Stochastic Multiplicative Processes With Reset Events," *Physical Review E*, 59, 4945–4948. [1812]

MEYLAHN, JANUSZ M., SANJIB SABHAPANDIT, AND HUGO TOUCHETTE (2015): "Large Deviations for Markov Processes With Resetting," *Physical Review E*, 92, 062148. [1812]

MITZENMACHER, MICHAEL (2004): "A Brief History of Generative Models for Power Law and Lognormal Distributions," *Internet Mathematics*, 1, 226–252. [1811]

MONTERO, MIQUEL, AND JAVIER VILLARROEL (2013): "Monotonic Continuous-Time Random Walks With Drift and Stochastic Reset Events," *Physical Review E*, 87, 012116. [1812]

MUKOYAMA, TOSHIHIKO, AND SOPHIE OSOTIMEHIN (2019): "Barriers to Reallocation and Economic Growth: The Effects of Firing Costs," *American Economic Journal: Macroeconomics*, 11, 235–270. [1812]

NAKAGAWA, KENJI (2007): "Application of Tauberian Theorem to the Exponential Decay of the Tail Probability of a Random Variable," *IEEE Transactions on Information Theory*, 53, 3239–3249. [1827]

NIREI, MAKOTO, AND SHUHEI AOKI (2016): "Pareto Distribution of Income in Neoclassical Growth Models," *Review of Economic Dynamics*, 20, 25–42. [1812]

REED, WILLIAM J. (2001): "The Pareto, Zipf and Other Power Laws," *Economics Letters*, 74, 15–19. [1811, 1819]

——— (2002): "On the Rank-Size Distribution for Human Settlements," *Journal of Regional Science*, 42, 1–17. [1819]

ROITERSHTEIN, ALEXANDER (2007): "One-Dimensional Linear Recursions With Markov-Dependent Coefficients," *Annals of Applied Probability*, 17, 572–608. [1822]

SIMON, HERBERT A., AND CHARLES P. BONINI (1958): "The Size Distribution of Business Firms," *American Economic Review*, 48, 607–617. [1811]

TODA, ALEXIS A. (2011): "Income Dynamics With a Stationary Double Pareto Distribution," *Physical Review E*, 83, 046122. [1819]

——— (2012): "The Double Power Law in Income Distribution: Explanations and Evidence," *Journal of Economic Behavior and Organization*, 84, 364–381. [1819]

——— (2014): "Incomplete Market Dynamics and Cross-Sectional Distributions," *Journal of Economic Theory*, 154, 310–348. [1820]

——— (2017): "A Note on the Size Distribution of Consumption: More Double Pareto Than Lognormal," *Macroeconomic Dynamics*, 21, 1508–1518. [1819]

——— (2019): "Wealth Distribution With Random Discount Factors," *Journal of Monetary Economics*, 104, 101–113. [1813]

TODA, ALEXIS A., AND KIERAN WALSH (2015): "The Double Power Law in Consumption and Implications for Testing Euler Equations," *Journal of Political Economy*, 123, 1177–1200. [1819]

VAHRENKAMP, RICHARD (1976): "Derivatives of the Dominant Root," *Applied Mathematics and Computation*, 2, 29–39. [1830]

WOLD, HERMAN O. A., AND PETER WHITTLE (1957): "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, 25, 591–595. [1811]

YAMAMOTO, KEN (2014): "Stochastic Model of Zipf's Law and the Universality of the Power-Law Exponent," *Physical Review E*, 89, 042115. [1812]