Econometric Theory, **38**, 2022, 986–1013.
doi:10.1017/S0266466621000268

TAIL BEHAVIOR OF STOPPED LÉVY PROCESSES WITH MARKOV MODULATION

BRENDAN K. BEARE
University of Sydney

WON-KI SEO
University of Sydney

ALEXIS AKIRA TODA
University of California San Diego

This article concerns the tail probabilities of a light-tailed Markov-modulated Lévy process stopped at a state-dependent Poisson rate. The tails are shown to decay exponentially at rates given by the unique positive and negative roots of the spectral abscissa of a certain matrix-valued function. We illustrate the use of our results with an application to the stationary distribution of wealth in a simple economic model in which agents with constant absolute risk aversion are subject to random mortality and income fluctuation.

1. INTRODUCTION

This article provides machinery for characterizing the tail probabilities $\mathbb{P}(W_T > w)$ of a Markov-modulated Lévy process $(W_t)_{t\geq 0}$ stopped at a random time T. The class of Markov-modulated Lévy processes is a generalization of the class of Lévy processes in which the distributions of increments are governed by a Markov process $(J_t)_{t\geq 0}$, assumed here to have finite state-space. The stopping time T is triggered at a Poisson rate dependent on the Markov state. We assume that our Markov-modulated Lévy process is light-tailed, in the sense that $\mathbb{P}(W_t > w)$ decays to zero exponentially or faster as $w \to \infty$ for any fixed t. Our main result is that $\mathbb{P}(W_T > w)$ decays to zero at an exponential rate given by the unique positive root of the spectral abscissa of a certain matrix-valued function. An analogous result holds for the lower tail.

Our article is intended as a companion piece to Beare and Toda (2021), where closely related results are obtained in a discrete-time setting. In both articles, the problem we solve is motivated by an influential contribution of Reed (2001). Reed observed that in an economy consisting of a large number of units (e.g.,

Address correspondence to Bredan K. Beare, School of Economics, University of Sydney, Sydney, Australia; e-mail: brendan.beare@sydney.edu.au.

© The Author(s), 2021. Published by Cambridge University Press.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

firms, households, and cities) with an exponential age distribution, each of which has been growing like a geometric Brownian motion since birth, the cross-sectional distribution of sizes is that of a geometric Brownian motion stopped at an exponentially distributed time. This turns out to be the double Pareto distribution. Reed's observation explains, at an intuitive level, why numerous heterogeneous-agent models of macroeconomic activity generate Pareto-tailed cross-sectional distributions of income, wealth, and consumption; see, e.g., Toda (2014), Toda and Walsh (2015), Albornoz, Fanelli, and Hallak (2016), Arkolakis (2016), Benhabib, Bisin, and Zhu (2016), Gabaix et al. (2016), Aoki and Nirei (2017), Cao and Luo (2017), Jones and Kim (2018), and Kasa and Lei (2018). Related insights can be gleaned from an earlier literature in the 1950s, including Champernowne (1953), Rutherford (1955), Simon (1955), Sargan (1957), Wold and Whittle (1957), and Simon and Bonini (1958), concerning the tails of cross-sectional distributions in dynamic economic systems.

If we take the logarithm of a random variable with a double Pareto distribution, then we obtain a random variable with a Laplace distribution. The tails of a Laplace distribution decay at exponential rates. Consequently, the aforementioned observation of Reed (2001) implies that the distribution of a Brownian motion stopped at an exponentially distributed time has tails that decay at exponential rates. The results of this article show that we continue to obtain exponential rates of tail decay in a more general setting where in place of the Brownian motion, we have any light-tailed Markov-modulated Lévy process, and where in place of the exponentially distributed stopping time, we have a stopping time occurring at a state-dependent Poisson rate. This greatly expands the scope of potential applications: light-tailed Lévy processes are used widely in financial modeling as an empirically relevant generalization of Brownian motion, while Markov-modulation is a ubiquitous feature of empirically calibrated models in macroeconomics.

The remainder of this article is structured as follows. In Section 2, we briefly introduce the essential features of Lévy and Markov-modulated Lévy processes. Our primary results are presented in Section 3. We illustrate the use of those results in Section 4 with an application to the distribution of wealth in a simple economic model with agents subject to random mortality and income fluctuation. Appendices A.1–A.3 contain a presentation of mathematical results that play a key role in our proofs and may not be widely known among specialists in economic and econometric theory. Appendices A.4 and A.5 contain proofs of results stated in Sections 3 and 4.

2. PRELIMINARIES

In this section, we briefly recall the definitions of Lévy and Markov-modulated Lévy processes, and associated concepts. Our discussion is based mostly on Sato (1999) and Asmussen (2003).

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

2.1. Lévy Processes

A real-valued stochastic process $W = (W_t)_{t\geq 0}$ is called a Lévy process if it satisfies the following conditions.

(i) $\mathbb{P}(W_0 = 0) = 1$.
(ii) W has càdlàg (right continuous with left limits) paths with probability one.
(iii) For any $s \geq 0$, $(W_{s+t} - W_s)_{t\geq 0}$ is independent of $(W_t)_{0\leq t\leq s}$.
(iv) For any $s \geq 0$, the law of $(W_{s+t} - W_s)_{t\geq 0}$ is the same as the law of $(W_t)_{t\geq 0}$.

The law of a Lévy process W is uniquely determined by its Lévy exponent $\psi(z) = \log \mathbb{E}(e^{zW_1})$, where log denotes the complex logarithm, defined here in terms of the principal value. For imaginary z, the Lévy exponent $\psi(z)$ is the complex logarithm of the characteristic function of W_1; for complex z at which the expectation is well-defined, it is the complex logarithm of the Laplace transform or moment generating function (MGF) of W_1. Throughout this article, we take the domain of $M_X(z) = \mathbb{E}(e^{zX})$, the MGF of an arbitrary random variable X, to be the set

$$\mathcal{S}_X = \{z \in \mathbb{C} : \mathbb{E}(e^{(\mathrm{Re}\, z)X}) < \infty\},$$

a strip in the complex plane. Every MGF is holomorphic on the interior of its domain, and has log-convex restriction to the real axis. Consequently, every Lévy exponent is holomorphic on the interior of its domain, and has convex restriction to the real axis.

A Lévy process W is said to have a light upper (lower) tail if there is a positive (negative) real s such that $\mathbb{E}(e^{sW_1}) < \infty$, meaning that the domain of the Lévy exponent $\psi(z)$ has a positive (negative) real element. Note that the property of finiteness of the MGF of W_t is time-invariant, in the following sense: for any given real s, the three statements (a) $\mathbb{E}(e^{sW_1}) < \infty$, (b) $\mathbb{E}(e^{sW_t}) < \infty$ for some $t > 0$, and (c) $\mathbb{E}(e^{sW_t}) < \infty$ for all $t > 0$, are equivalent (Sato, 1999, Thm. 25.17). Moreover, for any complex z in the domain of $\psi(z)$, we have $\mathbb{E}(e^{zW_t}) = e^{t\psi(z)}$ for all $t \geq 0$. A Lévy process is said to be light-tailed if it has light upper and lower tails.

Prominent examples of Lévy processes which we will make use of in our discussion include Brownian motions with drift, Poisson processes, and Cauchy processes. A Brownian motion with drift has Lévy exponent $\psi(z) = \mu z + \frac{1}{2}\sigma^2 z^2$, with real parameters μ and $\sigma^2 \geq 0$ called the drift and diffusion. A Poisson process has Lévy exponent $\psi(z) = \phi(e^z - 1)$, with real parameter $\phi \geq 0$ called the intensity. A Cauchy process has Lévy exponent $\psi(z) = -|z|$, defined only for imaginary z.

2.2. Markov-Modulated Lévy Processes

Let $\mathcal{N} = \{1, \ldots, N\}$, a finite set. Consider a bivariate stochastic process $(W, J) = (W_t, J_t)_{t\geq 0}$ taking values in $\mathbb{R} \times \mathcal{N}$. We say that W is a Markov-modulated Lévy process, and that J is the Markov-modulator of W, if the following conditions are satisfied.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

(i) $\mathbb{P}(W_0 = 0) = 1$.
(ii) W and J have càdlàg paths with probability one.
(iii) For any $s \geq 0$, $(W_{s+t} - W_s, J_{s+t})_{t\geq 0}$ is conditionally independent of $(W_t, J_t)_{0\leq t\leq s}$ given J_s.
(iv) For any $n \in \mathcal{N}$ and any $s \geq 0$, the law of $(W_{s+t} - W_s, J_{s+t})_{t\geq 0}$ conditional on $J_s = n$ is the same as the law of $(W_t, J_t)_{t\geq 0}$ conditional on $J_0 = n$.

It is clear from conditions (iii) and (iv) that the Markov-modulator J is a time-homogeneous Markov process taking values in $\mathcal{N}$. When $\mathcal{N}$ is a singleton, conditions (i)–(iv) reduce to their analogues in Section 2.1, indicating that W is a Lévy process. When $\mathcal{N}$ has multiple elements, W is not in general a Lévy process but may be regarded, loosely speaking, as behaving like a Lévy process in a Markov environment.

The joint law of a Markov-modulated Lévy process W and its Markov-modulator J may be specified in terms of the following four objects.

(i) *Initial state probabilities*: An $N \times 1$ vector ϖ with non-negative entries ϖ_n summing to one.
(ii) *Infinitesimal generator*: An $N \times N$ matrix Π with non-negative off-diagonal entries $\pi_{nn'}$, and all rows summing to zero.
(iii) *State-dependent Lévy exponents*: An $N \times N$ diagonal matrix-valued function $\Psi(z)$ with Lévy exponents $\psi_n(z)$ along the diagonal. (We take the domain of $\Psi(z)$ to be the intersection of the domains of the diagonal components.)
(iv) *MGFs for additional jumps*: An $N \times N$ matrix-valued function $\Upsilon(z)$ with diagonal entries equal to one and MGFs $\upsilon_{nn'}(z)$ in the off-diagonal entries. (We set $\upsilon_{nn'}(z) = 1$ if the corresponding transition rate $\pi_{nn'}$ is zero, and take the domain of $\Upsilon(z)$ to be the intersection of the domains of the off-diagonal components.)

These four objects uniquely determine the law of (W, J) in the following way. First, the law of the time-homogeneous Markov process J is uniquely determined by the law of J_0, which is given by ϖ, and the infinitesimal generator matrix Π. Specifically, we may draw J_0 according to the probabilities $\mathbb{P}(J_0 = n) = \varpi_n$, and then generate $(J_t)_{t>0}$ according to the transition rates

$$\lim_{t\downarrow 0} \frac{1}{t} \mathbb{P}(J_{s+t} = n' \mid J_s = n) = \pi_{nn'}, \quad n \neq n'.$$

Next, the law of the Markov-modulated Lévy process W conditional on J is uniquely determined by $\Psi(z)$ and $\Upsilon(z)$ as follows. Let $t_0 = 0$, and let $t_1, t_2, \ldots$ denote the successive times at which J transitions between states. Given J, the Markov-modulated Lévy process W is initialized at $W_0 = 0$ and evolves as a Lévy process with Lévy exponent $\psi_n(z)$ over each interval (t_k, t_{k+1}) with $J_{t_k} = n$. The transition points t_k may trigger additional jumps in W: if J transitions from state n to state n' at time t_k, then W experiences a jump at time t_k with size determined by the MGF $\upsilon_{nn'}(z)$. Note that the MGF $\upsilon_{nn'}(z)$ assigns probability one to a jump of size zero if $\upsilon_{nn'}(z) = 1$.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

If a state n is never reached with probability one, then the Lévy exponent $\psi_n(z)$ and the transition densities $\pi_{nn'}$ and jump MGFs $\upsilon_{nn'}(z)$ will be undetermined for each $n' \neq n$. In this case, we should simply reformulate the state space to exclude the redundant state n.

Let $\odot$ denote the Hadamard (entry-wise) product between two matrices of the same size. The following characterization of the conditional MGFs of a Markov-modulated Lévy process is a key input to our results.

PROPOSITION 2.1. *Let W be a Markov-modulated Lévy process with Markov-modulator J, parametrized by ϖ, Π, $\Psi(z)$ and $\Upsilon(z)$ as above. For any $t > 0$, any z in the domains of $\Psi(z)$ and $\Upsilon(z)$, and any $n, n' \in \mathcal{N}$, the conditional expectation $\mathbb{E}(e^{zW_t} 1(J_t = n') \mid J_0 = n)$ is equal to the (n,n')-entry of the matrix $\exp(t(\Psi(z) + \Pi \odot \Upsilon(z)))$.*

Proof. See Proposition 2.2 of Asmussen (2003, p. 311). ■

For readers seeking additional details on Markov-modulated Lévy processes, a useful textbook treatment may be found in Asmussen (2003, ch. XI), where such processes are called Markov additive processes.

3. RESULTS

Let $W = (W_t)_{t \geq 0}$ be a Markov-modulated Lévy process, with Markov-modulator $J = (J_t)_{t \geq 0}$ taking values in a finite set $\mathcal{N} = \{1, \ldots, N\}$. Let ϖ, Π, $\Psi(z)$ and $\Upsilon(z)$ be, respectively, the $N \times 1$ vector of initial probabilities, $N \times N$ infinitesimal generator matrix, $N \times N$ diagonal matrix of Lévy exponents, and $N \times N$ matrix of jump MGFs for (W, J), as described in Section 2.2. These four objects uniquely determine the law of (W, J). We seek to characterize the tails of W_T, where T is a random time.

We would like to interpret T as a time-of-death when mortality occurs at a state-dependent Poisson rate. To this end, we define a second Markov-modulated Lévy process $V = (V_t)_{t \geq 0}$, which shares the Markov-modulator J of W. As a counterpart to $\Psi(z)$, we introduce an $N \times N$ diagonal matrix Φ with non-negative real entries ϕ_n along the diagonal. As in Section 2.2, let $t_0 = 0$, and let $t_1, t_2, \ldots$ denote the successive times at which J transitions between states. We suppose that V is initialized at $V_0 = 0$ and evolves as a Poisson process of intensity ϕ_n over each interval (t_k, t_{k+1}) with $J_{t_k} = n$. We suppose that V is continuous at each transition time t_k with probability one; therefore, the implicit counterpart to $\Upsilon(z)$ is an $N \times N$ matrix of ones. This uniquely specifies the law of V given J. To complete our specification of the law of (V, W, J), we assume that V is conditionally independent of W given J. We set $T = \inf\{t \geq 0 : V_t > 0\}$, the time at which V experiences its first jump. A condition on Π and Φ that is necessary and sufficient to have $\mathbb{P}(T < \infty) = 1$ is given in Proposition 3.2.

Let $\mathcal{S}$ be the intersection of the domains of $\Psi(z)$ and $\Upsilon(z)$; that is, the intersection of the domains of all of the Lévy exponents $\psi_n(z)$ and MGFs $\upsilon_{nn'}(z)$.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Recalling our discussion of such domains in Section 2.1, we observe that $\mathcal{S}$ is a strip in the complex plane, and that its real part $\mathcal{I} = \{\mathrm{Re}\, z : z \in \mathcal{S}\}$ is an interval, possibly a singleton or unbounded, containing zero. In order to apply Theorems 3.1 and 3.2 to the upper (lower) tail of W_T, we will require that $\mathcal{I}$ contains a positive (negative) element. In this sense, W is required to be light-tailed; this extends the notion of a light-tailed Lévy process given in Section 2.1 to a Markov-modulated setting.

For $z \in \mathcal{S}$, define the complex matrix-valued function

$$\mathrm{A}(z) = \Psi(z) + \Pi \odot \Upsilon(z) - \Phi. \tag{3.1}$$

Our characterization of the tails of W_T will depend on the behavior of the spectral abscissa of $\mathrm{A}(s)$ as a function of the real variable $s \in \mathcal{I}$. Recall that the spectral abscissa of a complex square matrix A, which we denote $\zeta(A)$, is the maximum of the real parts of the complex eigenvalues of A; that is,

$$\zeta(A) = \max\{\mathrm{Re}\,\lambda : \lambda \text{ is an eigenvalue of } A\}.$$

Remark 3.1. A real square matrix is said to be Metzler if its off-diagonal entries are non-negative. As discussed in Appendix A.3, the spectral abscissa of a Metzler matrix is an eigenvalue of that matrix, and has associated left and right eigenvectors with non-negative entries. This result resembles the Perron–Frobenius theorem for non-negative matrices. The matrix $\mathrm{A}(s)$ defined by (3.1) is Metzler for every $s \in \mathcal{I}$.

Our first result concerns the shape of the spectral abscissa of $\mathrm{A}(s)$ as a function of the real variable $s \in \mathcal{I}$. The proof, which relies on a result of Nussbaum (1986) on the convexity of spectral abscissae, may be found in Appendix A.4, along with the proofs of all other numbered propositions and theorems stated in this section.

PROPOSITION 3.1. *The spectral abscissa $\zeta(\mathrm{A}(s))$ is a convex function of $s \in \mathcal{I}$ and satisfies $\zeta(\mathrm{A}(0)) \leq 0$. If $\zeta(\mathrm{A}(0)) < 0$, then the equation $\zeta(\mathrm{A}(s)) = 0$ admits at most one positive solution $s = \alpha \in \mathcal{I}$ and at most one negative solution $s = -\beta \in \mathcal{I}$.*

We illustrate the content of Proposition 3.1 in Figure 1. In the case depicted, the endpoints of $\mathcal{I}$ are finite, but these may also be infinite. The condition $\zeta(\mathrm{A}(0)) < 0$ is satisfied; note that $\mathrm{A}(0) = \Pi - \Phi$. The following result shows that this condition has a natural interpretation.

PROPOSITION 3.2. *We have $\mathbb{P}(T < \infty) = 1$ if and only if $\zeta(\mathrm{A}(0)) < 0$.*

Our main results, Theorems 3.1 and 3.2, concern the case where there exists a positive and/or negative s in the interior of $\mathcal{I}$ solving $\zeta(\mathrm{A}(s)) = 0$. Figure 1 depicts a case where both positive and negative solutions exist, but this need not always occur. We now give three examples of ways in which there might not exist a positive or negative solution. In each example, we have $N = 1$, so W is a Lévy process with no Markov-modulation, and the equation $\zeta(\mathrm{A}(s)) = 0$ reduces to $\psi(s) = \phi$.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

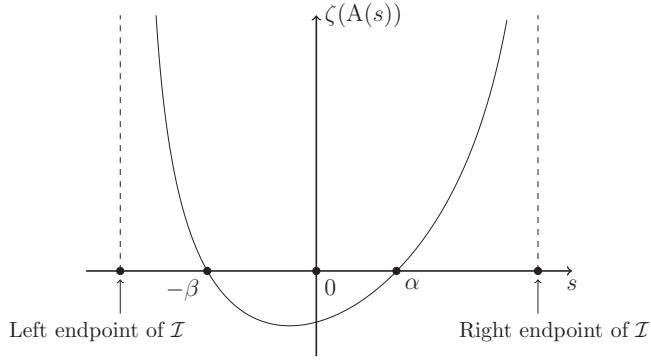


FIGURE 1. Determination of α and β from the spectral abscissa of $\mathrm{A}(s)$.

Example 3.1. If W is a Cauchy process then $\mathcal{I} = \{0\}$ and thus there cannot exist any positive or negative $s \in \mathcal{I}$. The problem is that a Cauchy process is not light-tailed.

Example 3.2. If W is a Poisson process with intensity $\gamma > 0$, then its Lévy exponent satisfies $\psi(s) = \gamma(\mathrm{e}^s - 1) < 0$ for all $s < 0$, so there cannot exist a negative solution to $\psi(s) = \phi$. The problem is that a Poisson process is non-negative. If $\phi > 0$, then the unique positive solution is $s = \log(1 + \phi/\gamma)$.

Example 3.3. This example is more subtle. Fix $c > 0$ and define a measure ν on $\mathbb{R}$ by $\nu(\mathrm{d}x) = 1(x \geq 1)cx^{-2}\mathrm{e}^{-x}\,\mathrm{d}x$. Let W be a pure jump process with Lévy measure ν, meaning that W has Lévy exponent satisfying

$$\psi(s) = \int_{-\infty}^{\infty} (\mathrm{e}^{sx} - 1)\nu(\mathrm{d}x) = c\int_{1}^{\infty} (\mathrm{e}^{(s-1)x} - \mathrm{e}^{-x})x^{-2}\,\mathrm{d}x$$

for all real s in its domain. The last integral is finite for $s \leq 1$ and infinite for $s > 1$, so $\mathcal{I} = (-\infty, 1]$. If $\phi > 0$, then we may choose c small enough that $\psi(1) < \phi$, in which case there will not exist a positive $s \in \mathcal{I}$ solving $\psi(s) = \phi$. The problem is that $\int (\mathrm{e}^{sx} - 1)\nu(\mathrm{d}x)$ jumps from below ϕ to infinity at the right endpoint of $\mathcal{I}$.

Remark 3.2. A simple sufficient condition for the existence of a positive s in the interior of $\mathcal{I}$ solving $\zeta(\mathrm{A}(s)) = 0$ is that (i) $\mathbb{P}(T < \infty) = 1$, and (ii) at least one of the Lévy exponents $\psi_n(s)$, viewed as a function of real $s \in \mathcal{I}$, diverges to infinity as s increases to the (positive, possibly infinite) right endpoint of $\mathcal{I}$. In this case, $\zeta(\mathrm{A}(s))$ is negative for $s = 0$ (Proposition 3.2) and increases to infinity as s approaches the right endpoint of $\mathcal{I}$, since

$$\zeta(\mathrm{A}(s)) \geq \zeta(\mathrm{diag}(\mathrm{A}(s))) = \max\{\psi_n(s) + \pi_{nn} - \phi_n : n \in \mathcal{N}\} \to \infty,$$

using a monotonicity property of the spectral abscissa (Theorem A.5) to obtain the inequality. Since $\zeta(\mathrm{A}(s))$ is convex (Proposition 3.1), hence continuous, on

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

the interior of $\mathcal{I}$, it follows from the intermediate value theorem that there exists a positive s solving $\zeta(\mathrm{A}(s)) = 0$. A similar remark applies to the existence of a negative solution.

When $\mathbb{P}(T < \infty) = 1$, the random variable W_T is well-defined on a set of probability one. The remainder of our results concern its distribution. We begin by identifying the MGF of W_T. Define

$$\mathcal{I}_- = \{s \in \mathcal{I} : \zeta(\mathrm{A}(s)) < 0\}, \quad \mathcal{S}_- = \{z \in \mathbb{C} : \mathrm{Re}\, z \in \mathcal{I}_-\}.$$

We will obtain an expression for the MGF of W_T on $\mathcal{S}_-$. The following property of $\mathcal{S}_-$ is important because it ensures that $\mathrm{A}(z)$ is invertible on $\mathcal{S}_-$.

PROPOSITION 3.3. *We have* $\zeta(\mathrm{A}(z)) < 0$ *for every* $z \in \mathcal{S}_-$.

Let 1_N denote an $N \times 1$ vector with each entry equal to one. We have the following representation of the MGF of W_T on $\mathcal{S}_-$.

PROPOSITION 3.4. *If* $\mathbb{P}(T < \infty) = 1$ *then, for each* $z \in \mathcal{S}_-$, *the MGF of* W_T *satisfies*

$$M_{W_T}(z) = \mathbb{E}(\mathrm{e}^{zW_T}) = \varpi^\top \mathrm{A}(z)^{-1}\mathrm{A}(0)1_N. \tag{3.2}$$

Example 3.4. (Brownian motion with drift.) Consider the simple case where there is no Markov-modulation ($N = 1$), and where W is a Brownian motion with drift, with Lévy exponent $\psi(z) = \mu z + \frac{1}{2}\sigma^2 z^2$. We then have $\zeta(\mathrm{A}(z)) = \mu z + \frac{1}{2}\sigma^2 z^2 - \phi$, a quadratic polynomial in z with positive and negative roots

$$\alpha = \frac{-\mu + \sqrt{\mu^2 + 2\sigma^2\phi}}{\sigma^2}, \quad -\beta = \frac{-\mu - \sqrt{\mu^2 + 2\sigma^2\phi}}{\sigma^2}.$$

Applying Proposition 3.4, we find that the MGF of W_T is given by

$$M_{W_T}(z) = \frac{\phi}{\phi - \mu z - \frac{1}{2}\sigma^2 z^2} = \frac{\alpha\beta}{(\alpha - z)(\beta + z)}.$$

This is the MGF of the asymmetric Laplace (or double exponential) distribution, which has probability density function

$$f(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta}\mathrm{e}^{-\alpha x} & \text{for } x \geq 0, \\ \frac{\alpha\beta}{\alpha+\beta}\mathrm{e}^{\beta x} & \text{for } x < 0. \end{cases}$$

The fact that a Brownian motion with drift stopped at an exponentially distributed time has the asymmetric Laplace distribution was shown by Reed (2001).

In Example 3.4, the distribution of W_T has exponentially decaying tail probabilities, with the rates of exponential decay in each tail given by the positive and negative roots α and $-\beta$ of $\zeta(\mathrm{A}(s))$, as in Figure 1. Our main result, Theorem 3.1, shows that this is a general feature of light-tailed Markov-modulated

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Lévy processes stopped at a state-dependent Poisson rate. To establish it, we rely on an assumption that Π is irreducible. A complex $N \times N$ matrix $A = (a_{nn'})$ is said to be irreducible if, for any nonempty proper subset $\mathcal{M}$ of $\mathcal{N}$, there is an $n \in \mathcal{M}$ and an $n' \in \mathcal{N} \setminus \mathcal{M}$ such that $a_{nn'} \neq 0$. For the infinitesimal generator matrix Π, irreducibility means that the state variable J_t cannot become permanently confined to any strict subset of the state space.

THEOREM 3.1. *Suppose that* $\mathbb{P}(T < \infty) = 1$ *and that* Π *is irreducible. If the equation* $\zeta(\mathrm{A}(s)) = 0$ *admits a positive solution* $s = \alpha$ *in the interior of* $\mathcal{I}$ *then*

$$0 < \liminf_{w\to\infty} \mathrm{e}^{\alpha w}\,\mathbb{P}(W_T > w) \leq \limsup_{w\to\infty} \mathrm{e}^{\alpha w}\,\mathbb{P}(W_T > w) < \infty. \tag{3.3}$$

Similarly, if the equation $\zeta(\mathrm{A}(s)) = 0$ *admits a negative solution* $s = -\beta$ *in the interior of* $\mathcal{I}$ *then*

$$0 < \liminf_{w\to\infty} \mathrm{e}^{\beta w}\,\mathbb{P}(W_T < -w) \leq \limsup_{w\to\infty} \mathrm{e}^{\beta w}\,\mathbb{P}(W_T < -w) < \infty. \tag{3.4}$$

Remark 3.3. Our proof of Theorem 3.1 involves showing that α and $-\beta$ are poles of $\mathrm{A}(z)^{-1}$. Under the irreducibility condition on Π, we show that these poles of $\mathrm{A}(z)^{-1}$ are simple, and are also simple poles of $M_{W_T}(z)$. This allows us to apply a theorem of Nakagawa (2007), discussed in Appendix A.1, to obtain positive and finite lower and upper bounds for the limits inferior and superior in (3.3) and (3.4).

Remark 3.4. If (3.3) is satisfied, then there exist $w_0, c, c' \in (0, \infty)$ such that

$$c \leq \mathrm{e}^{\alpha w}\,\mathbb{P}(W_T > w) \leq c'$$

for all $w \geq w_0$. Taking logarithms, dividing by w, and letting $w \to \infty$, we obtain

$$\lim_{w\to\infty} \frac{1}{w} \log \mathbb{P}(W_T > w) = -\alpha. \tag{3.5}$$

Thus (3.3) implies (3.5). A similar statement applies to the lower tail.

We illustrate the use of Theorem 3.1 with applications to a two-state Brownian motion with drift, and a two-state linear trend.

Example 3.5 (Two-state Brownian motion with drift)**.** Suppose next that there are two Markov states ($N = 2$), and that in state n, W evolves like a Brownian motion with drift μ_n and diffusion $\sigma_n^2 > 0$, and V evolves like a Poisson process with intensity ϕ_n. Suppose that transitions do not trigger jumps, so that all entries of $\Upsilon(z)$ are equal to one. The rows of the infinitesimal generator Π must sum to zero, so we write it as

$$\Pi = \begin{bmatrix} -\pi_1 & \pi_1 \\ \pi_2 & -\pi_2 \end{bmatrix}.$$

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Suppose that $\pi_1, \pi_2 > 0$, so that Π is irreducible, and suppose that $\phi_1 > 0$ or $\phi_2 > 0$, so that $\mathbb{P}(T < \infty) = 1$. The matrix $\mathrm{A}(s)$ in (3.1) is given by

$$\mathrm{A}(s) = \begin{bmatrix} \frac{1}{2}\sigma_1^2 s^2 + \mu_1 s - \phi_1 - \pi_1 & \pi_1 \\ \pi_2 & \frac{1}{2}\sigma_2^2 s^2 + \mu_2 s - \phi_2 - \pi_2 \end{bmatrix} =: \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

It has two real eigenvalues given by the solutions to the quadratic equation

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0,$$

the larger of which is the spectral abscissa

$$\zeta(\mathrm{A}(s)) = \frac{a+d+\sqrt{(a+d)^2 - 4(ad-bc)}}{2}. \tag{3.6}$$

Note that the discriminant satisfies

$$(a+d)^2 - 4(ad-bc) = (a-d)^2 + 4bc > 0$$

due to the fact that $bc = \pi_1\pi_2 > 0$.

To apply Theorem 3.1, we seek positive and negative values of s such that the right-hand side of (3.6) is equal to zero. This occurs when $a+d \le 0$ and $ad - bc = 0$; that is, when

$$\frac{1}{2}(\sigma_1^2 + \sigma_2^2)s^2 + (\mu_1 + \mu_2)s - (\phi_1 + \phi_2) - (\pi_1 + \pi_2) \le 0 \tag{3.7}$$

and

$$\left(\frac{1}{2}\sigma_1^2 s^2 + \mu_1 s - \phi_1 - \pi_1\right)\left(\frac{1}{2}\sigma_2^2 s^2 + \mu_2 s - \phi_2 - \pi_2\right) - \pi_1\pi_2 = 0. \tag{3.8}$$

Denote by $f(s)$ the left-hand side of (3.8), a quartic polynomial in s, and denote by $g_1(s)$ and $g_2(s)$, the two quadratic polynomials in parentheses on the left-hand side of (3.8). Observe that

$$f(0) = \pi_1\phi_2 + \phi_1\pi_2 + \phi_1\phi_2 > 0 \tag{3.9}$$

since by assumption $\pi_1, \pi_2 > 0$ and $\phi_1 > 0$ or $\phi_2 > 0$. Each quadratic polynomial $g_n(s)$ has positive and negative roots $\alpha_n, -\beta_n$, at which our quartic polynomial $f(s)$ satisfies

$$f(-\beta_1) = f(-\beta_2) = f(\alpha_1) = f(\alpha_2) = -\pi_1\pi_2 < 0. \tag{3.10}$$

From (3.9), (3.10), $f(\pm\infty) = \infty$, and the intermediate value theorem, we deduce that each of the four disjoint intervals

$$(-\infty, -\max\{\beta_1, \beta_2\}),\ (-\min\{\beta_1, \beta_2\}, 0),\ (0, \min\{\alpha_1, \alpha_2\}),\ (\max\{\alpha_1, \alpha_2\}, \infty) \tag{3.11}$$

contains at least one root of $f(s)$. Since f is a quartic polynomial, by the fundamental theorem of algebra, f has exactly four complex roots. Therefore, f has four distinct real roots, one in each interval in (3.11). The two quadratic polynomials $g_1(s)$ and $g_2(s)$ are negative over the interval $(0, \min\{\alpha_1, \alpha_2\})$, implying that their

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

sum is negative, and so the quadratic inequality (3.7) is satisfied. On the other hand, $g_1(s)$ and $g_2(s)$ are both positive over the interval $(\max\{\alpha_1, \alpha_2\}, \infty)$, implying that their sum is positive, and so (3.7) is not satisfied. Thus the unique positive root of the quartic polynomial $f(s)$ belonging to the interval $(0, \min\{\alpha_1, \alpha_2\})$ is the unique positive solution $s = \alpha$ to the equation $\zeta(\mathrm{A}(s)) = 0$ and, by Theorem 3.1, is equal to the rate of exponential decay in the upper tail of W_T. Similarly, the unique negative root of $f(s)$ belonging to the interval $(-\min\{\beta_1, \beta_2\}, 0)$ is the unique negative solution $s = -\beta$ to the equation $\zeta(\mathrm{A}(s)) = 0$, and provides the rate of exponential decay in the lower tail of W_T.

Example 3.6 (Two-state linear trend)**.** Suppose we eliminate the diffusive component in the two-state Brownian motion considered in Example 3.5 by setting $\sigma_1^2 = \sigma_2^2 = 0$, thereby obtaining a two-state linear trend. In this case, the quadratic inequality (3.7) reduces to

$$(\mu_1 + \mu_2)s \leq \phi_1 + \phi_2 + \pi_1 + \pi_2, \tag{3.12}$$

and the quartic equation (3.8) reduces to

$$(\mu_1 s - \phi_1 - \pi_1)(\mu_2 s - \phi_2 - \pi_2) - \pi_1\pi_2 = 0. \tag{3.13}$$

Denote by $g(s)$ the left-hand side of (3.13), a quadratic polynomial in s if $\mu_1\mu_2 \neq 0$, or linear otherwise. Note that $g(0) > 0$ by the same argument as in Example 3.5. Excluding the case $\mu_1\mu_2 = 0$ and assuming without loss of generality that $\mu_1 \leq \mu_2$, there are three cases to consider.

Case 1: $\mu_1 < 0 < \mu_2$. *Since* $g(0) > 0$ *and* $g(\pm\infty) = -\infty$*, the quadratic polynomial* $g(s)$ *has unique positive and negative roots* α *and* $-\beta$*. Thus, by Theorem 3.1,* W_T *has exponential upper and lower tails with decay rates* α *and* β *respectively.*

Case 2: $0 < \mu_1 \leq \mu_2$. *Let* $\alpha_n = (\pi_n + \phi_n)/\mu_n > 0$*, so that* $g(\alpha_1) = g(\alpha_2) = -\pi_1\pi_2 < 0$*. Since* $g(0) > 0$ *and* $g(\infty) = \infty$*, the disjoint intervals* $(0, \min\{\alpha_1, \alpha_2\})$ *and* $(\max\{\alpha_1, \alpha_2\}, \infty)$ *must each contain a root of the quadratic polynomial* $g(s)$*. The larger root violates inequality (3.12). Letting* α *denote the smaller root, we deduce from Theorem 3.1 that* W_T *has an exponential upper tail with decay rate* α*.*

Case 3: $\mu_1 \leq \mu_2 < 0$. *Symmetrically to Case 2,* W_T *has an exponential lower tail with decay rate given by the maximum of the two negative roots of* $g(s)$*.*

Remark 3.5. Cao and Luo (2017) study a general equilibrium neoclassical growth model in continuous-time with persistent heterogeneous returns to investment. Productivity in their model is a two-state Markov process. This leads to a stationary distribution of log-wealth whose upper tail is determined in a manner similar to that of W_T in Example 3.6 above. Thm. 3 of Cao and Luo (2017)

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

establishes that the upper tail exponent of their stationary distribution of wealth is given by the solution to a quadratic equation similar to (3.13).

Theorem 3.1 establishes conditions under which W_T has upper tail decaying exponentially at rate α in the sense of (3.3), and lower tail decaying exponentially at rate β in the sense of (3.4). It may be reasonable to ask whether (3.3) and (3.4) may be strengthened to

$$0 < \liminf_{w\to\infty} e^{\alpha w}\,\mathbb{P}(W_T > w) = \limsup_{w\to\infty} e^{\alpha w}\,\mathbb{P}(W_T > w) < \infty \quad \text{and} \tag{3.14}$$

$$0 < \liminf_{w\to\infty} e^{\beta w}\,\mathbb{P}(W_T < -w) = \limsup_{w\to\infty} e^{\beta w}\,\mathbb{P}(W_T < -w) < \infty, \tag{3.15}$$

respectively, so that $e^{\alpha w}\,\mathbb{P}(W_T > w)$ and $e^{\beta w}\,\mathbb{P}(W_T < -w)$ converge to positive and finite limits as $w \to \infty$. In fact, this is not possible in general: Example 3.7 below provides a case where W_T satisfies (3.3) but not (3.14).

If we would like to strengthen Theorem 3.1 such that the tails of W_T decay exponentially in the sense of (3.14) and (3.15), then it suffices to impose a nonlattice condition. We say that an MGF is lattice if it corresponds to a distribution that is supported on a lattice; that is, a distribution that assigns probability one to a set of the form $\{a + bm : m \in \mathbb{Z}\}$ for some real a and $b > 0$. Otherwise, we say that an MGF is nonlattice.

THEOREM 3.2. *Suppose that either (i) for some $n \in \mathcal{N}$, the MGF $\exp\psi_n(z)$ is nonlattice, or (ii) for some $n, n' \in \mathcal{N}$ with $\pi_{nn'} > 0$, the MGF $\upsilon_{nn'}(z)$ is nonlattice. Then Theorem 3.1 remains valid with (3.14) in place of (3.3), and (3.15) in place of (3.4).*

Remark 3.6. Our proof of Theorem 3.2 involves showing that our nonlattice condition ensures that the simple poles α and $-\beta$ of the MGF $M_{W_T}(z)$ in (3.2) are the unique singularities on the lines $\mathrm{Re}\, z = \alpha, -\beta$, called the axes of convergence of $M_{W_T}(z)$. This strengthens the lower and upper bounds on the limits inferior and superior in (3.3) and (3.4) obtained using the theorem of Nakagawa (2007), such that the lower and upper bounds are equal.

In situations where Theorem 3.1 applies but the nonlattice condition in Theorem 3.2 is not satisfied, (3.14) or (3.15) may fail to hold, as the following example shows.

Example 3.7. (Poisson process.) Consider the case where there is no Markov-modulation ($N = 1$), and where W is a Poisson process with intensity $\gamma > 0$, as in Example 3.2. We saw there that when $\phi > 0$, the unique positive solution to $\psi(s) = \phi$ is $s = \log(1 + \phi/\gamma)$. Therefore, applying Theorem 3.1, we find that the upper tail of W_T decays exponentially in the sense of (3.3), with decay rate $\alpha = \log(1 + \phi/\gamma)$.

Does the upper tail of W_T decay exponentially in the stronger sense of (3.14)? We cannot deduce this from Theorem 3.2, because the MGF $\exp\psi(z)$ corresponds

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

to a Poisson distribution, which is supported on the lattice of integers. In fact, the upper tail of W_T does not satisfy (3.14), as we now show. Using the formula for the probability mass function of the Poisson distribution, for any non-negative integer k we obtain

$$\mathbb{P}(W_T = k) = \mathbb{E}(\mathbb{P}(W_T = k \mid T)) = \int_0^\infty \frac{(\gamma x)^k}{k!} \mathrm{e}^{-\gamma x} \phi \mathrm{e}^{-\phi x} \, \mathrm{d}x = \frac{\phi \gamma^k}{(\gamma + \phi)^{k+1}}.$$

It follows that

$$\mathbb{P}(W_T > w) = \mathbb{P}(W_T \geq \lfloor w \rfloor + 1) = \frac{\phi}{\gamma + \phi} \sum_{k=\lfloor w \rfloor + 1}^{\infty} \left(\frac{\gamma}{\gamma + \phi} \right)^k = \left(\frac{\gamma}{\gamma + \phi} \right)^{\lfloor w \rfloor + 1},$$

where $\lfloor \cdot \rfloor$ rounds down to the nearest integer. Since $\mathrm{e}^\alpha = (\gamma + \phi)/\gamma$, we deduce that

$$\mathrm{e}^{\alpha w} \mathbb{P}(W_T > w) = \left(\frac{\gamma}{\gamma + \phi} \right)^{\lfloor w \rfloor - w + 1},$$

which oscillates between 1 and $\gamma/(\gamma + \phi)$ as $w \to \infty$, in contravention of (3.14).

4. APPLICATION TO THE DISTRIBUTION OF WEALTH

We illustrate our results with an application to the distribution of wealth in a simple economic model in which agents with constant absolute risk aversion (CARA) are subject to idiosyncratic income fluctuation. The model we consider is similar to those of Wang (2003) and Achdou et al. (2021). Throughout this section, to avoid repeated qualifications about sets of measure zero, we regard any two stochastic processes whose paths are equal with probability one to be the same object.

4.1. Single Agent Problem

Let $J = (J_t)_{t \geq 0}$ be a càdlàg Markov process taking values in the finite set $\mathcal{N} = \{1, \ldots, N\}$, with irreducible $N \times N$ infinitesimal generator matrix Π, and $N \times 1$ vector of initial probabilities ϖ, which we take to be given by the stationary distribution of J. Let $t_0 = 0$, and let $t_1, t_2, \ldots$ denote the successive times at which J transitions between states. We suppose that for t belonging to an interval $[t_k, t_{k+1})$ with $J_t = n$, an agent receives a constant flow of income $Y_t = y_n \in \mathbb{R}$. The agent chooses a consumption flow $C = (C_t)_{t \geq 0}$, which is a stochastic process satisfying two conditions:

(i) C is adapted to the filtration generated by J.
(ii) C is càdlàg, with discontinuities permitted only at $t_1, t_2, \ldots$.

We denote by $\mathcal{C}$ the collection of all stochastic processes C satisfying (i) and (ii). Condition (i) ensures that the agent's consumption at time t depends only on information available at time t. Condition (ii) allows us to define the agent's wealth over each interval $[t_k, t_{k+1})$ to be the unique solution to an ordinary differential

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

equation (ODE) satisfying a boundary condition at t_k. Specifically, we define the agent's wealth $W^C = (W_t^C)_{t\geq 0}$, whose dependence on C we make explicit in our notation, to be the unique continuous stochastic process initialized at $W_0^C = 0$ which, over each interval $[t_k, t_{k+1})$, satisfies the ODE

$$\frac{\mathrm{d}W_t^C}{\mathrm{d}t} = rW_t^C + Y_t - C_t, \tag{4.1}$$

where $r > 0$ is a fixed rate of interest. The existence of a unique solution to this ODE satisfying a boundary condition at t_k is ensured by the fact that $Y_t - C_t$ is continuous on $[t_k, t_{k+1})$ with finite limit at t_{k+1} (Hartman, 1982, Cor. 5.1, p. 31).

The agent's objective function is the expected discounted flow of future utility, defined for $C \in \mathcal{C}$ by

$$U(C) = \mathbb{E}_0 \int_0^\infty \mathrm{e}^{-\rho t} u(C_t)\,\mathrm{d}t. \tag{4.2}$$

Here, $\mathbb{E}_0$ denotes expectation conditional on J_0, the parameter $\rho > 0$ is the agent's rate of discounting, and $u(\cdot)$ is the CARA utility function $u(c) = -\mathrm{e}^{-\gamma c}/\gamma$, with coefficient of absolute risk aversion $\gamma > 0$. Since C is measurable under condition (ii) and $u(\cdot)$ is bounded from above, the integral in (4.2) is well-defined, though possibly equal to minus infinity.

The agent's consumption flow C is required to satisfy the no-Ponzi condition

$$\liminf_{t\to\infty} \mathrm{e}^{-\rho t}\,\mathbb{E}_0(\mathrm{e}^{-\gamma r W_t^C}) = 0. \tag{4.3}$$

The no-Ponzi condition serves as a borrowing constraint by limiting the agent's ability to accumulate debt. We will see in Appendix A.5 that it corresponds to the transversality condition for the agent's optimization problem. We denote by $\mathcal{C}_1$ the collection of all $C \in \mathcal{C}$ satisfying (4.3).

The optimization problem we have just described resembles a continuous-time reformulation of the discrete-time problem studied by Wang (2003). The one substantive difference is that Wang (2003) assumes an autoregressive law of motion for income, which we are unable to accommodate due to our assumption that the number of Markov states is finite. Our optimization problem is also very similar to the leading example discussed by Achdou et al. (2021), also in continuous-time, which assumes a two-state income process. Those authors use a fixed lower bound for wealth as a borrowing constraint; we instead follow Wang (2003) and use the transversality condition as a borrowing constraint. The optimal consumption flow is given by the following result, proved in Appendix A.5.

PROPOSITION 4.1. *$U(C)$ is uniquely maximized over $\mathcal{C}_1$ by a consumption flow C^*, with associated wealth process W^*, such that*

$$C_t^* = rW_t^* + b_n \quad \textit{while } J_t = n, \tag{4.4}$$

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

where $b_1, \ldots, b_N$ are the unique real numbers solving the system of equations

$$b_n = y_n - \frac{1}{\gamma} + \frac{\rho}{\gamma r} - \frac{1}{\gamma r} \sum_{n'=1}^{N} \pi_{nn'} \mathrm{e}^{\gamma(b_n - b_{n'})}, \quad n \in \mathcal{N}. \tag{4.5}$$

We see from Proposition 4.1 that the agent consumes the interest earned on wealth, plus a quantity determined by the current income level. Combining (4.1) and (4.4) reveals that over each interval (t_k, t_{k+1}) with $J_{t_k} = n$, the agent's wealth has slope $y_n - b_n$. Wealth is therefore a piecewise linear trend with slope determined by the current Markov state. This is a simple example of a Markov modulated Lévy process; we studied the two-state case in Example 3.6.

4.2. Partial and General Equilibrium

We now characterize the distribution of wealth in a partial or general equilibrium model with a unit mass of agents. We adopt the perpetual youth formulation of birth and death introduced by Yaari (1965), in which agents die at a constant Poisson rate $\phi > 0$, returning their wealth or debt to a central bank, and are reborn with zero initial wealth and initial income state drawn from the stationary distribution ϖ. (The central bank may alternatively be viewed as an insurance company buying and selling actuarial notes, as in Yaari, 1965, p. 140.) The stationary age distribution under this scheme is exponential with parameter ϕ, which is the same as the distribution of an individual agent's lifespan T. Agents maximize their expected discounted flow of future utility. Since the mortality rate is constant, an agent's lifespan T is independent of their income flow, so that with discount rate $\tilde{\rho} > 0$ we have

$$\mathbb{E}_0 \int_0^T \mathrm{e}^{-\tilde{\rho} t} u(C_t)\, \mathrm{d}t = \mathbb{E}_0 \int_0^\infty \mathrm{e}^{-(\tilde{\rho}+\phi)t} u(C_t)\, \mathrm{d}t.$$

The infinite horizon problem discussed in Section 4.1 is therefore the same as the one solved by agents subject to mortality if we set $\rho = \tilde{\rho} + \phi$.

If agents have wealth processes that are independent copies of W^*, then the stationary distribution of wealth in the economy is the distribution of W^*_T, where T is independent of W^*. This is the essential insight of Reed (2001). As discussed in Section 4.1, combining the wealth dynamics (4.1) with the optimal consumption rule (4.4) reveals that W^* is a continuous piecewise linear trend, with slope $y_n - b_n$ while the current Markov state is n. This means that W^* is a Markov-modulated Lévy process parametrized by ϖ, Π, $\Psi(z)$ and $\Upsilon(z)$, where $\Psi(z)$ is a diagonal matrix with nth diagonal entry $(y_n - b_n)z$, and $\Upsilon(z)$ has all entries equal to one. The corresponding matrix $\mathrm{A}(z)$ in (3.1) has diagonal entries $\mathrm{A}_{nn}(z) = (y_n - b_n)z + \pi_{nn} - \phi$ and off-diagonal entries $\mathrm{A}_{nn'}(z) = \pi_{nn'}$. In view of Propositions 3.1 and 3.2 and Remark 3.2, if $b_n < y_n$ for some n, then the equation $\zeta(\mathrm{A}(s)) = 0$ admits a unique positive solution $s = \alpha$. A similar statement applies to the existence of a unique

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

negative solution. The following result is therefore an immediate consequence of Theorem 3.1.

PROPOSITION 4.2. *If $b_n < y_n$ and $b_{n'} > y_{n'}$ for some n, n', then the equation $\zeta(\mathrm{A}(s)) = 0$ has unique positive and negative solutions $s = \alpha, -\beta$, and the upper and lower tails of the stationary distribution of wealth decay exponentially at rates α and β in the sense of (3.3) and (3.4).*

The preceding analysis may be regarded as one of partial equilibrium because the interest rate r was treated as exogenous. In a general equilibrium model, the interest rate needs to adjust so that the asset market clears. Changes in r affect the system of equations (4.5) determining $b_1, \ldots, b_N$ in Proposition 4.1. We therefore write $b_n = b_n(r)$, and set $b(r) = (b_1(r), \ldots, b_N(r))^\top$ and $y = (y_1, \ldots, y_N)^\top$. Since an individual agent's rate of saving or borrowing in state n is $y_n - b_n$, and the proportion of agents in state n is ϖ_n, the condition for general equilibrium is

$$\varpi^\top (y - b(r)) = 0. \tag{4.6}$$

Note that under condition (4.6), the expected slope of W^* at any given point in time is zero, implying that aggregate wealth—that is, $\mathbb{E}(W_T^*)$, the expected value of a random draw from the stationary distribution of wealth—is zero. Consequently, the net flow of assets returned by departing agents is zero.

Our final result, proved in Appendix A.5, shows that general equilibrium exists, and that in general equilibrium the condition in Proposition 4.2 for exponential tail decay is automatically satisfied when income fluctuates.

PROPOSITION 4.3. *There exists an interest rate $r > 0$ such that (4.6) is satisfied. For any such r, if $y_1, \ldots, y_N$ take at least two distinct values, then $b_n(r) < y_n$ and $b_{n'}(r) > y_{n'}$ for some n, n'. Consequently, the tails of the stationary distribution of wealth decay exponentially as in Proposition 4.2.*

Note that Proposition 4.3 guarantees the existence of general equilibrium, but says nothing about uniqueness. Toda (2017) shows that, in a discrete-time version of our model, there is a unique equilibrium if income follows a first-order autoregressive law of motion as in Wang (2003), but multiple equilibria may obtain under a number of other common time series models for income.

Achdou et al. (2021) analyze the stationary distribution of wealth in a continuous-time heterogeneous-agent model in which, as in our model, the sole source of heterogeneity is bounded fluctuation in income. They establish that the support of the stationary distribution of wealth in their model is bounded from above. This is an instance of a more general principle established by Stachurski and Toda (2019) in a discrete-time context: for a wide class of heterogeneous-agent models in which income fluctuation is the sole source of heterogeneity, the stationary distribution of wealth inherits the upper tail behavior (in this instance, boundedness) of income shocks. Contrary to this principle, we have established

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

that the stationary distribution of wealth in our model has exponential (hence unbounded) upper tail even though income has finite (hence bounded) support. The critical feature of our model leading to this discrepancy is that agents have CARA utility, violating a bounded relative risk aversion condition imposed by Stachurski and Toda (2019) and Achdou et al. (2021). Under CARA, the optimal consumption path obtained here in continuous-time and by Wang (2003) in discrete-time leads to a nonstationary wealth path. We are thus forced to introduce random birth/death as in Yaari (1965) so as to obtain a stationary distribution of wealth. Under bounded relative risk aversion, Stachurski and Toda (2019) and Achdou et al. (2021) obtain a stationary distribution of wealth even with infinitely lived agents.

The technical machinery developed in Section 3 of this article may be applied more generally to models in which the wealth path of an agent subject to random birth/death may be written as an increasing function of a Markov-modulated Lévy process. In a companion article, Beare and Toda (2021) study a discrete-time model in which agents are exposed to fluctuations in capital income. Owing to the multiplicative nature of capital income, it turns out to be the logarithm of an agent's wealth path which is (the discrete-time counterpart to) a Markov-modulated Lévy process. The effect of the logarithm is to generate a stationary distribution of wealth with a Pareto upper tail. In the simpler economic model considered in this article, the absence of investment risk leads to a stationary distribution of wealth with an exponential upper tail, which may be less desirable from an empirical perspective.

5. CONCLUDING REMARKS

The primary results established in this article provide a closed-form expression for the Laplace transform of the distribution of a light-tailed Markov-modulated Lévy process stopped at a state-dependent Poisson rate, and an equation whose unique positive and negative solutions are the rates of exponential decay in the tails of that distribution. We conclude with some brief remarks on further extensions and applications.

Our main result on exponential tail decay rates, Theorem 3.1, relies on the assumption that the infinitesimal generator matrix Π is irreducible. As discussed in Remark 3.3, irreducibility of Π is used to show that the poles α and $-\beta$ of the Laplace transform $M_{W_T}(z)$ are simple. When Π is reducible, these poles may no longer be simple, and so the version of the theorem of Nakagawa (2007) presented in Appendix A.1 is no longer applicable. We must instead use a more general version of Nakagawa's theorem applicable to poles of arbitrary order, which leads to a subtly different characterization of tail probabilities. We plan to address the reducible case thoroughly in a subsequent article.

Throughout this article, we have assumed that the state-space $\mathcal{N}$ for the Markov state J_t is finite. For many applications, it may be more useful to consider a continuum of Markov states. In the continuous case, the matrices Π, $\Psi(z)$ and $\Upsilon(z)$ used to parametrize a Markov-modulated Lévy process, and the diagonal matrix Φ of stopping rates, are properly viewed as linear operators on a suitable

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

infinite-dimensional vector space. Our analysis of tail probabilities then hinges on spectral properties of the operator-valued function $\mathrm{A}(z)$; in particular, on the behavior of $\mathrm{A}(z)^{-1}$ at its poles α and $-\beta$. Consequently, to handle the continuous case, an adaptation of Theorem A.2 in Appendix A.2 to an infinite-dimensional setting may be required. We expect that results obtained recently by Beare and Seo (2020) and Franchi and Paruolo (2020) in the context of developing representation theory for cointegrated functional time series may be useful in this regard. Adaptation of the results on Metzler matrices presented in Appendix A.3 to an infinite-dimensional setting ought to be possible by applying the Krein–Rutman theorem in place of the Perron–Frobenius theorem.

The application of our results to problems of immediate practical interest is a priority. In an unpublished manuscript, Gouin-Bonenfant and Toda (2020) have used a discrete-time version of our results (as in Beare and Toda, 2021) to develop numerical procedures for computing aggregate quantities of interest from empirically calibrated heterogeneous-agent macroeconomic models. Accuracy superior to traditional grid-based procedures is achieved by using the knowledge of tail decay rates derived from our results to extrapolate into the upper tail of the wealth distribution. In other recent work, Beare and Toda (2020) have used our results (again, in discrete-time as in Beare and Toda, 2021) to study the spread of coronavirus disease across U.S. counties in the early months of 2020. It is found that the decay rate of the upper-tail of the cross-sectional distribution of coronavirus cases across counties matches the rate implied by our results in a simple model in which random multiplicative growth within counties is paired with the distribution of ages (i.e., times since first infection) across counties. We anticipate that further applications of our results, or generalizations thereof, may be possible in a diverse range of areas of interest.

APPENDIX

A. Mathematical Appendix

A.1. *Exponential Tails via the Laplace Transform.* A theorem of Nakagawa (2007) is the key tool we use to prove Theorems 3.1 and 3.2. Nakagawa's theorem places bounds on tail probabilities depending on singularities of the associated Laplace transform. The relevant singularities for our problem are simple poles. In this appendix, we present a simplified version of Nakagawa's theorem that is sufficient to handle simple poles.

First, we briefly review the Laplace transform. For more details see Widder (1941) and Lukacs (1970, ch. 7). Given a real random variable X with cumulative distribution function (CDF) $F(x)$, let $M(s) = \int_{-\infty}^{\infty} \mathrm{e}^{sx}\,\mathrm{d}F(x)$ be the corresponding MGF, defined initially for all real s such that the integral is finite. The set of all such s forms an interval containing zero, possibly a singleton, which we denote $\mathcal{I}$. The right and left boundary points of this interval, which may be real or $\pm\infty$, are denoted α and $-\beta$ and called the right and left abscissae of convergence. We extend the domain of M into the complex plane by setting $M(z) = \int_{-\infty}^{\infty} \mathrm{e}^{zx}\,\mathrm{d}F(x)$ for all complex z such that the integral is well-defined; this is precisely the set $\mathcal{S} = \{z \in \mathbb{C} : \mathrm{Re}\, z \in \mathcal{I}\}$. Using the dominated convergence theorem, it is easy to see

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

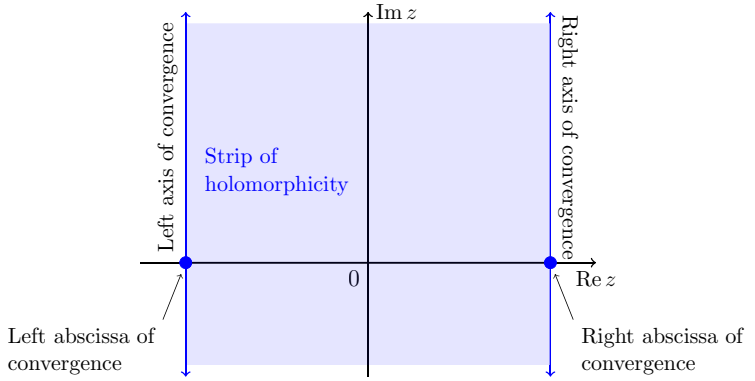


FIGURE A.1. Strip of holomorphicity of the Laplace transform.

that $M(z)$ is holomorphic on the interior of $\mathcal{S}$, which is called the strip of holomorphicity (Figure A.1). The lines $\mathrm{Re}\, z = \alpha, -\beta$ comprising the boundary of $\mathcal{S}$ are called the right and left axes of convergence. Viewed as a function of a complex variable, the MGF $M(z)$ is called the Laplace transform of the CDF $F(x)$, or of the random variable X.

There is a close relationship between the tail probabilities of a CDF and the abscissae of convergence of its Laplace transform. In general, we have

$$\limsup_{x\to\infty} \frac{1}{x} \log \mathbb{P}(X > x) = -\alpha, \quad \limsup_{x\to\infty} \frac{1}{x} \log \mathbb{P}(X < -x) = -\beta$$

whenever the relevant abscissa is nonzero (Widder, 1941, pp. 42–43, 241; Lukacs, 1970, p. 194). Moreover, each abscissa is a singularity of the Laplace transform (Widder, 1941, p. 58). Nakagawa's theorem establishes a tighter relationship between the abscissae of convergence and the tail probabilities that depends on the nature of the singularities at the abscissae, and the location of other singularities along the axes of convergence. For our purposes, it will be enough to consider the case where the singularities at the abscissae are simple poles. We state a result providing bounds for the right tail probabilities; an analogous result holds for the left tail.

THEOREM A.1. *Let X be a real random variable and $M(z)$ its Laplace transform, with right abscissa of convergence $\alpha \in (0, \infty)$. Suppose that $M(z)$ can be meromorphically extended to an open set containing its right axis of convergence, with a simple pole at α. Let $-C < 0$ denote the residue of $M(z)$ at α, and let $B > 0$ denote the supremum of all $b > 0$ such that α is the unique singularity of $M(z)$ on $\{\alpha + it : t \in (-b, b)\}$. Then if $B < \infty$, we have*

$$\frac{2\pi C/B}{\mathrm{e}^{2\pi\alpha/B} - 1} \le \liminf_{x\to\infty} \mathrm{e}^{\alpha x}\, \mathbb{P}(X > x) \le \limsup_{x\to\infty} \mathrm{e}^{\alpha x}\, \mathbb{P}(X > x) \le \frac{2\pi C/B}{1 - \mathrm{e}^{-2\pi\alpha/B}},$$

and if $B = \infty$, we have

$$\lim_{x\to\infty} \mathrm{e}^{\alpha x}\, \mathbb{P}(X > x) = \frac{C}{\alpha}.$$

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Proof. The result follows from Theorem 5* of Nakagawa (2007), which provides more complicated bounds that obtain when the singularity at the right abscissa of convergence is a pole of arbitrary order. ■

A.2. *Simple Poles of Matrix-Valued Functions.* In this appendix, we state a result providing conditions under which a singularity of a matrix-valued function of a complex variable is a simple pole, and characterizing the residue of that pole. It is used to prove Theorems 3.1 and 3.2.

THEOREM A.2. *Let $A(z)$ be an $N \times N$ matrix-valued holomorphic function of $z \in \Omega$, where Ω is some open and connected subset of the complex plane. Suppose that $A(z)$ is invertible for some $z \in \Omega$. Then $A(z)$ has a meromorphic inverse $A(z)^{-1}$ on Ω, with poles at the points of noninvertibility of $A(z)$. If $A(z_0)$ has rank $r < N$ for some $z_0 \in \Omega$, so that z_0 is a pole of $A(z)^{-1}$, then the following conditions are equivalent.*

(i) *z_0 is a simple pole of $A(z)^{-1}$.*
(ii) *The zero eigenvalue of $A(z_0)$ has equal geometric and algebraic multiplicities.*
(iii) *The $r \times r$ matrix $y^\top A'(z_0)x$ is invertible, where x and y can be any $N \times r$ matrices of full column rank such that $A(z_0)x = 0$ and $y^\top A(z_0) = 0$.*

Under any of these equivalent conditions, the residue of $A(z)^{-1}$ at the simple pole z_0 is equal to $x(y^\top A'(z_0)x)^{-1}y^\top$.

Proof. Meromorphicity of $A(z)^{-1}$ when $A(z)$ is somewhere invertible was proved by Steinberg (1968). The equivalence of (i) and (ii) was proved by Howland (1971). The equivalence of (i) and (iii) and the residue formula were proved by Schumacher (1986); see also Schumacher (1991, pp. 562–563). ■

A.3. *Spectral Properties of Metzler Matrices.* A real square matrix is said to be Metzler if its off-diagonal entries are non-negative. In this appendix, we present three results on Metzler matrices that are used repeatedly in Appendices A.4 and A.5. The first result shows that the well-known Perron–Frobenius theorem for non-negative square matrices extends to Metzler matrices if we shift our focus from the spectral radius to the spectral abscissa.

THEOREM A.3. *Let A be a Metzler matrix. Then $\zeta(A)$ is an eigenvalue of A, and there are corresponding left and right eigenvectors with non-negative entries. If, in addition, A is irreducible, then $\zeta(A)$ is an algebraically simple eigenvalue of A, and there are corresponding left and right eigenvectors with strictly positive entries.*

Proof. The result may be proved by applying the Perron–Frobenius theorem to the matrix $A + dI$, where I is an identity matrix and $d \geq 0$ is chosen large enough to make $A + dI$ nonnegative. See, for example, Corollary 3.2 of Smith (1995, Chap. 4). ■

The next result concerns Metzler matrices with constant row sums.

THEOREM A.4. *If all row sums of a Metzler matrix A are equal, then $\zeta(A)$ is equal to the common row sum.*

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Proof. The result may be proved by applying Thm. 8.1.22 of Horn and Johnson (2013) to $A + dI$, where $d \geq 0$ is chosen large enough to make $A + dI$ non-negative. ■

We also require a monotonicity property of the spectral abscissa.

THEOREM A.5. *Let $A = (a_{nn'})$ be a Metzler matrix, and let $B = (b_{nn'})$ be a complex matrix of the same size. Suppose that the entries of A and B satisfy*

$$\operatorname{Re} b_{nn} \leq a_{nn} \text{ for all } n, \text{ and } |b_{nn'}| \leq a_{nn'} \text{ for all } n \neq n'. \tag{A.1}$$

Then $\zeta(B) \leq \zeta(A)$. If, in addition, A is irreducible and at least one of the inequalities in (A.1) is strict, then $\zeta(B) < \zeta(A)$.

Proof. Let $C = (c_{nn'})$ be the Metzler matrix with diagonal entries $c_{nn} = \operatorname{Re} b_{nn}$ and off-diagonal entries $c_{nn'} = |b_{nn'}|$. The weak inequality $\zeta(B) \leq \zeta(C)$ was established in Cor. 1 of Deutsch (1975). Since $C \leq A$, we may obtain $\zeta(C) \leq \zeta(A)$ by applying Cor. 8.1.9 of Horn and Johnson (2013) to the matrices $A + dI$ and $C + dI$, where $d \geq 0$ is chosen large enough to make both matrices non-negative. When A is irreducible and at least one of the inequalities in (A.1) holds strictly, we may obtain $\zeta(C) < \zeta(A)$ by applying Thm. 8.4.5 of Horn and Johnson (2013) in the same fashion to show that if $\zeta(C) = \zeta(A)$ then $C = A$. ■

A.4. *Proofs of Results in Section 3.*

Proof of Proposition 3.1. The off-diagonal entries of $\mathrm{A}(s)$ are log-convex functions of s (since each is a positively scaled MGF), while the diagonal entries of $\mathrm{A}(s)$ are convex functions of s (since each is a log-MGF shifted by a constant). Thm. 1.1 of Nussbaum (1986) therefore implies that $\zeta(\mathrm{A}(s))$ is a convex function of s.

Since $\mathrm{A}(0) \leq \Pi$, Theorem A.5 implies that $\zeta(\mathrm{A}(0)) \leq \zeta(\Pi)$. Theorem A.4 implies that $\zeta(\Pi) = 0$. Thus, $\zeta(\mathrm{A}(0)) \leq 0$. The last sentence of the proposition is an obvious consequence of the convexity of $\zeta(\mathrm{A}(s))$. ■

Proof of Proposition 3.2. It will be convenient to augment our state space with an absorbing state into which we transition at time T. To this end, let $\widetilde{\mathcal{N}} = \{1, \ldots, N+1\}$, and for $t \geq 0$ let

$$\widetilde{J}_t = \begin{cases} J_t & \text{if } t < T \\ N+1 & \text{if } t \geq T. \end{cases}$$

The process $\widetilde{J} = (\widetilde{J}_t)_{t \geq 0}$ is a time-homogeneous Markov process with state space $\widetilde{\mathcal{N}}$. The $(N+1) \times 1$ vector of initial state probabilities $\widetilde{\varpi}$ and $(N+1) \times (N+1)$ infinitesimal generator matrix $\widetilde{\Pi}$ for $\widetilde{J}$ are given by

$$\widetilde{\varpi} = \begin{bmatrix} \varpi \\ 0 \end{bmatrix}, \quad \widetilde{\Pi} = \begin{bmatrix} \Pi - \Phi & -(\Pi - \Phi)1_N \\ 0_N^\top & 0 \end{bmatrix} = \begin{bmatrix} \mathrm{A}(0) & -\mathrm{A}(0)1_N \\ 0_N^\top & 0 \end{bmatrix},$$

where 0_N (respectively, 1_N) denotes an $N \times 1$ vector with each entry equal to zero (respectively, one). For $n \in \mathcal{N}$, let $p_n = \mathbb{P}(T < \infty \mid J_0 = n)$, the nth absorption probability for state $N+1$. (We may set $p_n = 1$ for any state n such that $P(J_0 = n) = 0$.) Define the vector $p = (p_1, \ldots, p_N)^\top$. Clearly $\mathbb{P}(T < \infty) = 1$ if and only if $p = 1_N$. We will show that the conditions $p = 1_N$ and $\zeta(\mathrm{A}(0)) < 0$ are both equivalent to $\mathrm{A}(0)$ being invertible.

We first show that $p = 1_N$ if and only if $\mathrm{A}(0)$ is invertible. Thm. 3.3.1 of Norris (1997) provides the following characterization of absorption probabilities: the system of linear

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

equations

$$\mathrm{A}(0)x = \mathrm{A}(0)1_N \tag{A.2}$$

admits a minimal non-negative solution, which is $x = p$. If $\mathrm{A}(0)$ is invertible, then $x = 1_N$ is the unique solution to (A.2), and so $p = 1_N$. If $\mathrm{A}(0)$ is not invertible then the set of solutions to (A.2) forms a hyperplane, so that the minimal non-negative solution must have at least one entry equal to zero, implying that $p \neq 1_N$.

We next show that $\mathrm{A}(0)$ is invertible if and only if $\zeta(\mathrm{A}(0)) < 0$. Since $\mathrm{A}(0)$ is Metzler, we know that its spectral abscissa is an eigenvalue (Theorem A.3), and so it suffices to show that $\zeta(\mathrm{A}(0)) \leq 0$. This was established in Proposition 3.1. ■

Our proofs of Proposition 3.3 and Theorem 3.2 rely on the following lemma, which resembles a well-known result on characteristic functions.

LEMMA A.1. *Let X be a real random variable and $M(z) = \mathbb{E}(\mathrm{e}^{zX})$ its Laplace transform, with right abscissa of convergence $\alpha > 0$. Then for any $s \in [0, \alpha)$ we have $|M(s+it)| \leq M(s)$ for all real t, with strict inequality for all real $t \neq 0$ if and only if the distribution of X is not supported on a lattice.*

Proof. The inequality $|M(s+it)| \leq M(s)$ is part of Thm. 7.1.2 of Lukacs (1970), and the remainder of the lemma follows from an obvious modification to the proof of Thm. 2.1.4 of Lukacs (1970). ■

Proof of Proposition 3.3. Fix $z \in \mathcal{S}_-$, and let $s = \mathrm{Re}\, z \in \mathcal{I}_-$. The diagonal entries of $\mathrm{A}(z)$ and $\mathrm{A}(s)$ are $\mathrm{A}_{nn}(z) = \psi_n(z) + \pi_{nn} - \phi_n$ and $\mathrm{A}_{nn}(s) = \psi_n(s) + \pi_{nn} - \phi_n$, and thus by Lemma A.1 satisfy $|\exp \mathrm{A}_{nn}(z)| \leq \exp \mathrm{A}_{nn}(s)$, which we may rewrite as $\mathrm{Re}\, \mathrm{A}_{nn}(z) \leq \mathrm{A}_{nn}(s)$. The off-diagonal entries of $\mathrm{A}(z)$ and $\mathrm{A}(s)$ are $\mathrm{A}_{nn'}(z) = \pi_{nn'} \upsilon_{nn'}(z)$ and $\mathrm{A}_{nn'}(s) = \pi_{nn'} \upsilon_{nn'}(s)$, and thus by Lemma A.1 satisfy $|\mathrm{A}_{nn'}(z)| \leq \mathrm{A}_{nn'}(s)$. Since $\mathrm{A}(s)$ is Metzler, we deduce from Theorem A.5 that $\zeta(\mathrm{A}(z)) \leq \zeta(\mathrm{A}(s))$. We conclude by observing that $\zeta(\mathrm{A}(s)) < 0$ due to the definition of $\mathcal{I}_-$. ■

Proof of Proposition 3.4. As in the proof of Proposition 3.2, we employ the augmented state space $\widetilde{\mathcal{N}}$. For $t \geq 0$, let $\widetilde{W}_t = W_{T \wedge t}$, and note that

$$\mathrm{e}^{z\widetilde{W}_t} 1(\widetilde{J}_t = N+1) = \mathrm{e}^{zW_T} 1(T \leq t). \tag{A.3}$$

The process $\widetilde{W} = (\widetilde{W}_t)_{t \geq 0}$ is a Markov-modulated Lévy process with Markov-modulator $\widetilde{J}$. The law of $\widetilde{W}$ conditional on $\widetilde{J}$ is parametrized by the $(N+1) \times (N+1)$ matrices

$$\widetilde{\Psi}(z) = \begin{bmatrix} \Psi(z) & 0_N \\ 0_N^\top & 0 \end{bmatrix}, \quad \widetilde{\Upsilon}(z) = \begin{bmatrix} \Upsilon(z) & 1_N \\ 1_N^\top & 1 \end{bmatrix}.$$

Fix $n \in \mathcal{N}$ such that $\mathbb{P}(J_0 = n) > 0$. By (A.3) and Proposition 2.1, for any $t > 0$, the conditional expectation $\mathbb{E}(\mathrm{e}^{zW_T} 1(T \leq t) \mid \widetilde{J}_0 = n)$ is equal to the entry in row n, column $N+1$ of the matrix

$$\exp(t(\widetilde{\Psi}(z) + \widetilde{\Pi} \odot \widetilde{\Upsilon}(z))) = \exp\left(t \begin{bmatrix} \mathrm{A}(z) & -\mathrm{A}(0)1_N \\ 0_N^\top & 0 \end{bmatrix}\right).$$

Applying Lem. 10.5.1 of Chen and Francis (1995), which provides a formula for the exponential of a block upper-triangular matrix, and using the fact that $\mathrm{A}(z)$ is invertible

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

for $z \in \mathcal{S}_-$ (Proposition 3.3), we find that $\mathbb{E}(e^{zW_T} 1(T \le t) \mid \widetilde{J}_0 = n)$ is equal to the entry in row n of the vector

$$-\int_0^t \exp(s\mathrm{A}(z))\mathrm{A}(0)1_N \, \mathrm{d}s = -\mathrm{A}(z)^{-1}(\exp(t\mathrm{A}(z)) - I)\mathrm{A}(0)1_N.$$

Since $\zeta(\mathrm{A}(z)) < 0$ for all $z \in \mathcal{S}_-$ (Proposition 3.3), we have $\exp(t\mathrm{A}(z)) \to 0$ as $t \to \infty$, and so $\mathbb{E}(e^{zW_T} 1(T \le t) \mid \widetilde{J}_0 = n)$ converges to the entry in row n of the vector $\mathrm{A}(z)^{-1}\mathrm{A}(0)1_N$ as $t \to \infty$.

It remains only to show that

$$\mathbb{E}(e^{zW_T} 1(T \le t) \mid J_0 = n) \to \mathbb{E}(e^{zW_T} \mid J_0 = n) \text{ as } t \to \infty. \tag{A.4}$$

Since $\mathbb{P}(T < \infty) = 1$, we have $\mathbb{P}(e^{zW_T} 1(T \le t) \to e^{zW_T} \mid J_0 = n) = 1$. For real $z = s \in \mathcal{I}_-$, we may thus obtain (A.4) by applying the monotone convergence theorem. In view of what was shown in the previous paragraph, this means that $\mathbb{E}(e^{sW_T} \mid J_0 = n)$ is the entry in row n of the vector $\mathrm{A}(s)^{-1}\mathrm{A}(0)1_N$, which is finite. We may thus obtain (A.4) for any complex $z \in \mathcal{S}_-$ by applying the dominated convergence theorem, using $e^{(\mathrm{Re}\, z)W_T}$ as the dominating function. ∎

Proof of Theorem 3.1. We will consider the case where the equation $\zeta(\mathrm{A}(s)) = 0$ admits a positive solution $s = \alpha$ in the interior of $\mathcal{I}$. The case of a negative solution may be handled by a symmetric argument.

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 $\mathrm{A}(z)$ is holomorphic on Ω. Since $\zeta(\mathrm{A}(s))$ is a convex function of $s \in \mathcal{I}$ (Proposition 3.1), with $\zeta(\mathrm{A}(0)) < 0$ (Proposition 3.2) and $\zeta(\mathrm{A}(\alpha)) = 0$, it must be the case that $\mathrm{A}(z)$ is invertible at all real numbers between 0 and α, which are elements of Ω. It follows from Theorem A.2 that $\mathrm{A}(z)$ has a meromorphic inverse $\mathrm{A}(z)^{-1}$ on Ω, with poles at the points of noninvertibility of $\mathrm{A}(z)$. Since $\mathrm{A}(\alpha)$ is Metzler, we know that its spectral abscissa of zero is an eigenvalue (Theorem A.3). Thus, $\mathrm{A}(z)$ is not invertible at $\alpha \in \Omega$, and α is a pole of $\mathrm{A}(z)^{-1}$.

We now show that α is a simple pole of $\mathrm{A}(z)^{-1}$, and determine the associated residue. Since $\mathrm{A}(\alpha)$ is irreducible Metzler, it follows from Theorem A.3 that the zero eigenvalue of $\mathrm{A}(\alpha)$ is algebraically simple, hence geometrically simple, and is associated with right and left eigenvectors x, y with strictly positive entries. From Theorem A.2, we thus deduce that α is a simple pole of $\mathrm{A}(z)^{-1}$, with residue given by

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

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

By Proposition 3.4, the MGF of W_T is given by $M_{W_T}(z) = \varpi^\top \mathrm{A}(z)^{-1}\mathrm{A}(0)1_N$ for $z \in \mathcal{S}_-$. Moreover, since $\mathrm{A}(z)^{-1}$ is meromorphic on Ω, this equation defines a meromorphic extension of $M_{W_T}(z)$ to Ω. As $z \to \alpha$, we obtain

$$(z-\alpha)M_{W_T}(z) = \varpi^\top (z-\alpha)\mathrm{A}(z)^{-1}\mathrm{A}(0)1_N \to \varpi^\top cxy^\top \mathrm{A}(0)1_N = c(\varpi^\top x)(y^\top \mathrm{A}(0)1_N).$$

The row sums of $\mathrm{A}(0)$ are bounded by those of Π, which are zero. Since $\mathrm{A}(0)$ is invertible, not all row sums are zero. Thus, $\mathrm{A}(0)1_N$ has nonpositive entries, with at least one strictly negative entry. The left eigenvector y has strictly positive entries, so $y^\top \mathrm{A}(0)1_N < 0$. Similarly, ϖ has non-negative entries, with at least one strictly positive entry, and the right eigenvector x has strictly positive entries, so $\varpi^\top x > 0$. As noted above, $c \neq 0$. Therefore, $\lim_{z\to\alpha}(z-\alpha)M_{W_T}(z) \neq 0$. This shows that α is a simple pole of $M_{W_T}(z)$. It now follows

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

from Theorem A.1 that the limits inferior and superior of $e^{\alpha w}\mathbb{P}(W_T > w)$ as $w \to \infty$ are strictly positive and finite. ■

Proof of Theorem 3.2. It suffices for us to show that α is the unique singularity of $M_{W_T}(z)$ on its axis of convergence, because in this case the application of Theorem A.1 at the end of the proof of Theorem 3.1 establishes that the limits inferior and superior of $e^{\alpha w}\mathbb{P}(W_T > w)$ as $w \to \infty$ are equal. We will do this by showing that $\zeta(\mathrm{A}(\alpha+it)) < 0$ for all real $t \neq 0$. The diagonal entries of $\mathrm{A}(\alpha+it)$ and $\mathrm{A}(\alpha)$ are $\mathrm{A}_{nn}(\alpha+it) = \psi_n(\alpha+it)+\pi_{nn}-\phi_n$ and $\mathrm{A}_{nn}(\alpha) = \psi_n(\alpha)+\pi_{nn}-\phi_n$, and thus by Lemma A.1 satisfy $|\exp \mathrm{A}_{nn}(\alpha+it)| \leq \exp \mathrm{A}_{nn}(\alpha)$, which we may rewrite as

$$\operatorname{Re} \mathrm{A}_{nn}(\alpha+it) \leq \mathrm{A}_{nn}(\alpha). \tag{A.5}$$

The off-diagonal entries of $\mathrm{A}(\alpha+it)$ and $\mathrm{A}(\alpha)$ are $\mathrm{A}_{nn'}(\alpha+it) = \pi_{nn'}\upsilon_{nn'}(\alpha+it)$ and $\mathrm{A}_{nn'}(\alpha) = \pi_{nn'}\upsilon_{nn'}(\alpha)$, and thus by Lemma A.1 satisfy

$$|\mathrm{A}_{nn'}(\alpha+it)| \leq \mathrm{A}_{nn'}(\alpha). \tag{A.6}$$

Moreover, under our non-lattice condition, Lemma A.1 implies that at least one of the inequalities in (A.5) or (A.6) must hold strictly for all real $t \neq 0$. Noting that $\mathrm{A}(\alpha)$ is irreducible Metzler, we see that $\mathrm{A}(\alpha+it)$ and $\mathrm{A}(\alpha)$ satisfy the conditions placed upon B and A in Theorem A.5, and deduce that $\zeta(\mathrm{A}(\alpha+it)) < \zeta(\mathrm{A}(\alpha)) = 0$ for all real $t \neq 0$. ■

A.5. *Proofs of Results in Section 4.*

LEMMA A.2. *The system of equations (4.5) has a unique solution* $(b_1, \ldots, b_N) \in \mathbb{R}^N$. *The solution varies continuously with* $r > 0$, *with* $\min_n b_n \to \infty$ *as* r *decreases to zero.*

Proof. Fix a constant $k > 0$, and let $f : \mathbb{R}^N \to \mathbb{R}^N$ be the map with nth entry

$$f_n(b) = \frac{1}{k+1}\left(kb_n + y_n - \frac{1}{\gamma} + \frac{\rho}{\gamma r} - \frac{1}{\gamma r}\sum_{n'=1}^{N}\pi_{nn'}e^{\gamma(b_n-b_{n'})}\right). \tag{A.7}$$

The system of equations (4.5) may be rewritten as $b = f(b)$. By Cor. 2.1.1 of Zhang (2013), if we can find an order interval $[u, v] \subset \mathbb{R}^N$ such that f is increasing and concave on $[u, v]$ with $f_n(u) > u_n$ and $f_n(v) < v_n$ for all n, then the restriction of f to $[u, v]$ has a unique fixed point. To this end, we fix $\eta > 0$ and set

$$u = \left(\min_n y_n - \frac{1}{\gamma} + \frac{\rho}{\gamma r} - \eta\right)1_N, \quad v = \left(\max_n y_n - \frac{1}{\gamma} + \frac{\rho}{\gamma r} + \eta\right)1_N.$$

It is easily verified that $f_n(u) > u_n$ and $f_n(v) < v_n$ for all n. Observe that

$$\frac{\partial f_m(b)}{\partial b_n} = \begin{cases} \frac{1}{k+1}\left(k - \frac{1}{r}\sum_{n'\neq n}\pi_{nn'}e^{\gamma(b_n-b_{n'})}\right) & \text{if } m = n, \\ \frac{1}{(k+1)r}\pi_{mn}e^{\gamma(b_m-b_n)} & \text{if } m \neq n. \end{cases}$$

Since Π is Metzler and $[u, v]$ is bounded, we may choose k large enough that all partial derivatives are nonnegative on $[u, v]$, ensuring that f is increasing on $[u, v]$. Next, observe that, for any $\delta \in (0, 1)$ and $b, b' \in \mathbb{R}^N$, the Metzler property of Π and convexity of $x \mapsto e^{\gamma x}$ together imply that

$$\pi_{nn'}e^{\gamma((\delta b_n+(1-\delta)b'_n)-(\delta b_{n'}+(1-\delta)b'_{n'}))} \leq \delta\pi_{nn'}e^{\gamma(b_n-b_{n'})} + (1-\delta)\pi_{nn'}e^{\gamma(b'_n-b'_{n'})}$$

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

for all n, n', with equality when $n = n'$. It thus follows from (A.7) that f is concave. We deduce from the aforementioned result of Zhang (2013) that the restriction of f to $[u, v]$ has a unique fixed point, which (with a mild abuse of notation) we denote $b(r)$. It follows that $b(r)$ is the unique solution to (4.5) in $[u, v]$. Moreover, since $\eta > 0$ was arbitrary, $b(r)$ is in fact the unique solution to (4.5) in $\mathbb{R}^N$, and satisfies

$$\min_{n'} y_{n'} - \frac{1}{\gamma} + \frac{\rho}{\gamma r} \leq b_n(r) \leq \max_{n'} y_{n'} - \frac{1}{\gamma} + \frac{\rho}{\gamma r} \quad \text{for all } n.$$

We therefore see that $\min_n b_n(r) \to \infty$ as r decreases to zero. Furthermore, since $b(r)$ is the unique value of b maximizing $-\|f(b) - b\|$, which varies continuously with b and r, the Berge maximum theorem implies that $b(r)$ is continuous in r. ■

Proof of Proposition 4.1. Let $\mathcal{D}$ be the collection of all functions $V : \mathcal{N} \times \mathbb{R} \to \mathbb{R}$ such that $V_n(w)$ is a continuously differentiable function of $w \in \mathbb{R}$ for each $n \in \mathcal{N}$. The optimization problem to be solved is a special case of Example (c) in Sect. III.4 and III.7 of Fleming and Soner (2006), with x, z and u in their notation corresponding to W, J and C in our notation. It is shown there that the relevant Hamilton–Jacobi–Bellman (HJB) equation for our optimization problem is

$$\rho V_n(w) = \max_{c \in \mathbb{R}} \left\{ u(c) + V_n'(w)(rw - c + y_n) + \sum_{n'=1}^{N} \pi_{nn'} V_{n'}(w) \right\}. \tag{A.8}$$

Note that this is an N-state generalization of the HJB equation obtained by Achdou et al. (2021) in the two-state case. A function $V^* \in \mathcal{D}$ that solves (A.8) for every $(n, w) \in \mathcal{N} \times \mathbb{R}$ is called a classical solution. With our CARA specification of $u(c)$, elementary calculus may be used to verify that the maximum in (A.8) is achieved by

$$c = -\frac{1}{\gamma} \log V_n'(w), \tag{A.9}$$

and that a classical solution to (A.8) is given by $V_n^*(w) = -\mathrm{e}^{-\gamma(rw + b_n)}/\gamma r$, where $b_1, \ldots, b_n$ are the unique constants given by Lemma A.2.

A consumption flow $C \in \mathcal{C}$ satisfies the no-Ponzi condition (4.3), meaning that it belongs to $\mathcal{C}_1$, if and only if it satisfies the so-called transversality condition

$$\limsup_{t \to \infty} \mathrm{e}^{-\rho t} \mathbb{E}_0 V_{J_t}^*(W_t^C) = 0. \tag{A.10}$$

By Thm. 9.1 of Fleming and Soner (2006, p. 140), if there is a consumption flow $C^* \in \mathcal{C}_1$ with associated wealth process W^* such that

$$C_t^* \in \arg\max_{c \in \mathbb{R}} \left\{ u(c) + V_n^{*\prime}(W_t^*)(rW_t^* - c + y_n) + \sum_{n'=1}^{N} \pi_{nn'} V_{n'}^*(W_t^*) \right\} \quad \text{while } J_t = n, \tag{A.11}$$

then $U(C)$ attains its maximum over $\mathcal{C}_1$ at C^*. Moreover, any such C^* is the unique maximizer of $U(C)$ over $\mathcal{C}_1$ due to the strict concavity of $u(c)$.

We will show that the consumption flow C^* given by (4.4) belongs to $\mathcal{C}_1$ and satisfies (A.11). It is obvious that $C^* \in \mathcal{C}$, so to show that $C^* \in \mathcal{C}_1$, we need to verify the no-Ponzi condition (4.3). Since W^* is a Markov-modulated Lévy process, it follows from

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Proposition 2.1 that, when $J_0 = n$, the quantity $e^{-\rho t}\mathbb{E}_0(e^{-\gamma r W_t^*})$ appearing in (4.3) is the nth row sum of $\exp(tA)$, where

$$A = -\rho I - \gamma r \operatorname{diag}(y - b) + \Pi.$$

Let D be the $N \times N$ diagonal matrix with nth diagonal entry $e^{\gamma b_n}$. From (4.5), we see that DAD^{-1} is Metzler with each row summing to $-r$. Consequently, Theorem A.4 implies that $\zeta(DAD^{-1}) = -r$. Since A and DAD^{-1} have the same spectra, we deduce that $\zeta(A) = -r$, which implies that $\exp(tA) \to 0$ as $t \to \infty$. Thus (4.3) is satisfied and $C^* \in \mathcal{C}_1$. Finally, in view of the fact that (A.9) solves the maximization in (A.8), and that $C_t^* = -(1/\gamma)\log V_n^{*\prime}(W_t^*)$ while $J_t = n$, it is clear that C^* satisfies (A.11). We conclude that C^* uniquely maximizes $U(C)$ over $\mathcal{C}_1$. ∎

Proof of Proposition 4.3. For $r > 0$, let $g(r) = \varpi^\top(y - b(r))$. The general equilibrium condition (4.6) is satisfied by any $r > 0$ such that $g(r) = 0$. Lemma A.2 implies that $g(r)$ is a continuous function of $r > 0$ with $g(r) \to -\infty$ as r decreases to zero. Therefore, if we can show that g is somewhere non-negative, then the existence of $r > 0$ satisfying $g(r) = 0$ will follow from the intermediate value theorem. Subtracting y_n from either side of (4.5), multiplying by $-\gamma r \varpi_n$, and summing over n, we obtain

$$\gamma r g(r) = r - \rho + \sum_{n=1}^{N}\sum_{n'=1}^{N} \varpi_n \pi_{nn'} e^{\gamma(b_n(r) - b_{n'}(r))}.$$

The Metzler property of Π and the inequality $e^x \geq 1 + x$ together imply that

$$\pi_{nn'} e^{\gamma(b_n(r) - b_{n'}(r))} \geq \pi_{nn'}(1 + \gamma(b_n(r) - b_{n'}(r)))$$

for all n, n', with equality when $n = n'$. We therefore have

$$\gamma r g(r) \geq r - \rho + \sum_{n=1}^{N} \varpi_n (1 + \gamma b_n(r)) \left(\sum_{n'=1}^{N} \pi_{nn'}\right) - \gamma \sum_{n'=1}^{N} \left(\sum_{n=1}^{N} \varpi_n \pi_{nn'}\right) b_{n'}(r).$$

Since $\sum_{n'} \pi_{nn'} = 0$ and $\sum_n \varpi_n \pi_{nn'} = 0$ (because ϖ is the stationary distribution corresponding to Π), it follows that $\gamma r g(r) \geq r - \rho$. Thus, $g(r) > 0$ for $r > \rho$, and we conclude that there exists $r \in (0, \rho]$ such that $g(r) = 0$.

When r satisfies (4.6), it must be the case that either $b_n(r) < y_n$ and $b_{n'}(r) > y_{n'}$ for some n, n', or $b(r) = y$. We will show that if $b(r) = y$ then all the values $y_1, \ldots, y_N$ are equal. Substituting y for b in the system of equations (4.5), we obtain

$$(\rho - r)e^{-\gamma y_n} = \sum_{n'=1}^{N} \pi_{nn'} e^{-\gamma y_{n'}}$$

for each n. Thus, $v = (e^{-\gamma y_1}, \ldots, e^{-\gamma y_N})^\top$ is a right eigenvector of Π with corresponding eigenvalue $\rho - r$. The stationary distribution ϖ is a left eigenvector of Π with corresponding eigenvalue zero, so we have $0 = \varpi^\top \Pi v = (\rho - r)\varpi^\top v$, implying that $r = \rho$. Thus the eigenvalue corresponding to the right eigenvector v is zero. Since Π is irreducible Metzler with $\zeta(\Pi) = 0$ (by Theorem A.4), Theorem A.3 implies that there is a unique (up to scalar multiplication) right eigenvector of Π corresponding to its zero eigenvalue. All row sums of Π are zero, so one such right eigenvector is 1_N. Thus v must be a scalar multiple of 1_N, meaning that $y_1 = \cdots = y_N$. ∎

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

REFERENCES

Achdou, Y., J. Han, J.M. Lasry, P.L. Lions, & B. Moll (2021) Income and wealth distribution in macroeconomics: A continuous-time approach. To appear in the *Review of Economic Studies*.

Albornoz, F., S. Fanelli & J. C. Hallak (2016) Survival in export markets. *Journal of International Economics* 102, 262–281.

Aoki, S. & M. Nirei (2017) Zipf's law, Pareto's law, and the evolution of top incomes in the United States. *American Economic Journal: Macroeconomics* 9, 36–71.

Arkolakis, C. (2016) A unified theory of firm selection and growth. *Quarterly Journal of Economics* 131, 89–155.

Asmussen, S. (2003) *Applied Probability and Queues*, *2nd Edition*. Springer.

Beare, B. K. & W. K. Seo (2020) Representation of I(1) and I(2) autoregressive Hilbertian processes. *Econometric Theory* 36, 773–802.

Beare, B. K. & A. A. Toda (2020) On the emergence of a power law in the distribution of COVID-19 cases. *Physica D: Nonlinear Phenomena* 412, 132649.

Beare, B.K. & A.A. Toda (2021) Determination of Pareto exponents in economic models driven by Markov multiplicative processes. Preprint. https://arxiv.org/abs/1712.01431.

Benhabib, J., A. Bisin & S. Zhu (2016) The distribution of wealth in the Blanchard-Yaari model. *Macroeconomic Dynamics* 20, 466–481.

Cao, D. & W. Luo (2017) Persistent heterogeneous returns and top end wealth inequality. *Review of Economic Dynamics* 26, 301–326.

Champernowne, D. G. (1953) A model of income distribution. *Economic Journal* 63, 318–351.

Chen, T. & B. Francis (1995) *Optimal Sampled-Data Control Systems*. Springer.

Deutsch, E. (1975) The spectral abscissa of partitioned matrices. *Journal of Mathematical Analysis and Applications* 50, 66–73.

Fleming, W. H. & H. M. Soner (2006) *Controlled Markov Processes and Viscosity Solutions*. Springer.

Franchi, M. & P. Paruolo (2020) Cointegration in functional autoregressive processes. *Econometric Theory* 36, 803–839.

Gabaix, X., J. M. Lasry, P. L. Lions & B. Moll (2016) The dynamics of inequality. *Econometrica* 84, 2071–2111.

Gouin-Bonenfant, É. & A.A. Toda (2020) Pareto extrapolation: An analytical framework for studying tail inequality. Preprint. https://ssrn.com/abstract=3260899.

Hartman, P. (1982) *Ordinary Differential Equations*, *2nd Edition*. Birkhauser.

Horn, R.A. & C.R. Johnson (2013) *Matrix Analysis*, *2nd Edition*. Cambridge University Press.

Howland, J. S. (1971) Simple poles of operator-valued functions. *Journal of Mathematical Analysis and Applications* 36, 12–21.

Jones, C. I. & J. Kim (2018) A Schumpeterian model of top income inequality. *Journal of Political Economy* 126, 1785–1826.

Kasa, K. & X. Lei (2018) Risk, uncertainty, and the dynamics of inequality. *Journal of Monetary Economics* 94, 60–78.

Lukacs, E. (1970) *Characteristic Functions*, *2nd Edition*. Griffin.

Nakagawa, K. (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.

Norris, J. R. (1997) *Markov Chains*. Cambridge University Press.

Nussbaum, R. D. (1986) Convexity and log convexity for the spectral radius. *Linear Algebra and its Applications* 73, 59–122.

Reed, W. J. (2001) The Pareto, Zipf and other power laws. *Economics Letters* 74, 15–19.

Rutherford, R. S. G. (1955) Income distributions: A new model. *Econometrica* 23, 277–294.

Sargan, J. D. (1957) The distribution of wealth. *Econometrica* 25, 568–590.

Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

Schumacher, J.M. (1986) Residue formulas for meromorphic matrices. In C.I. Byrnes & A. Lindquist (eds.), *Computational and Combinatorial Methods in Systems Theory*, pp. 97–111. North-Holland.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press

Schumacher, J.M. (1991) System-theoretic trends in econometrics. In A.C. Antoulas (ed.), *Mathematical System Theory: The Influence of R.E. Kalman*, pp. 559–577. Springer.
Simon, H. A. (1955) On a class of skew distribution functions. *Biometrika* 42, 425–440.
Simon, H. A. & C. P. Bonini (1958) The size distribution of business firms. *American Economic Review* 48, 607–617.
Smith, H. L. (1995) *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*. American Mathematical Society.
Stachurski, J. & A. A. Toda (2019) An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity. *Journal of Economic Theory* 182, 1–24.
Steinberg, S. (1968) Meromorphic families of compact operators. *Archive for Rational Mechanics and Analysis* 31, 372–379.
Toda, A. A. (2014) Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory* 154, 310–348.
Toda, A. A. (2017) Huggett economies with multiple stationary equilibria. *Journal of Economic Dynamics and Control* 84, 77–90.
Toda, A. A. & K. Walsh (2015) The double power law in consumption and implications for testing Euler equations. *Journal of Political Economy* 123, 1177–1200.
Wang, N. (2003) Caballero meets Bewley: The permanent-income hypothesis in general equilibrium. *American Economic Review* 93, 927–936.
Widder, D. V. (1941) *The Laplace Transform*. Princeton University Press.
Wold, H. O. A. & P. Whittle (1957) A model explaining the Pareto distribution of wealth. *Econometrica* 25, 591–595.
Yaari, M. E. (1965) Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies* 32, 137–150.
Zhang, Z. (2013) *Variational, Topological, and Partial Order Methods with Their Applications*. Springer.

https://doi.org/10.1017/S0266466621000268 Published online by Cambridge University Press