Markov-Chain Approximations for Life-Cycle Models

Giulio Fella* Giovanni Gallipoli† Jutong Pan‡

May 3, 2017

Abstract

Non-stationary income processes are standard in quantitative life-cycle models, prompted by the observation that within-cohort income inequality increases with age. This paper generalizes Tauchen (1986) and Rouwenhorst's (1995) discretization methods to non-stationary AR(1) processes. We evaluate the performance of both methods in the context of a canonical finite-horizon, income-fluctuation problem with a non-stationary income process. We find that the generalized Rouwenhorst's method performs extremely well even with a relatively small number of states.

JEL classification: C63

Keywords: Numerical methods, finite state approximations

*Queen Mary University of London, CFM and IFS. E-mail: g.fella@qmul.ac.uk
†Vancouver School of Economics, UBC. E-mail: gallipol@mail.ubc.ca
‡Vancouver School of Economics, UBC. E-mail: panjutong@gmail.com

1 Introduction

In quantitative macroeconomic studies it is often necessary to approximate continuous stochastic processes using discrete state-space representations; e.g. Markov chains. Different methods are available to perform such approximations.[1] The properties of alternative discretization methods to approximate covariance-stationary AR(1) processes in the context of stationary infinite horizon problems have been studied in some detail by Kopecky and Suen (2010). They find that: (a) the choice of discretization method may have a significant impact on the model simulated moments; (b) the performance of Rouwenhorst's (1995) method is more robust, particularly for highly persistent processes.

While a covariance-stationary income process is convenient, it is not consistent with the fact, first highlighted by Deaton and Paxson (1994), that within-cohort income inequality increases with the age of a cohort. For this reason, most quantitative life-cycle analyses of consumption and income dynamics assume a non-stationary labor income process whose variance increases with age.[2] As a result, the difficulty of accurately approximating the income process with a small number of discrete states increases with age.

We show how to extend both Tauchen (1986) and Rouwenhorst's (1995) methods to discretize non-stationary AR(1) processes and compare their respective performance within the context of a life-cycle, income-fluctuation problem. Both extensions keep the number of states in each time period constant, but they allow the state vector and transition matrix to change over time. In both cases, some property of the original stationary counterpart are preserved: Tauchen's method matches the transition probabilities implied by the normality assumption, while Rouwenhorst's method matches the conditional and unconditional first and second moments of the original process.

We evaluate the performance of both methods in the context of a finite-horizon income-fluctuation problem with a unit-root income process with normal innovations.[3] We find that Rouwenhorst's method performs extremely well even with a relatively small number of grid-points.

Our paper is related to several studies (see, among others, those listed in footnote 1). However, to the best of our knowledge, it is the first one to formally study the approximation of non-stationary AR(1) processes. Papers studying quantitative life-cycle problems with non-stationary stochastic processes have typically approximated those processes using a variety of intuitively appealing approaches. Storesletten et al. (2004) use a binomial tree, Huggett (1996) uses a variant of Tauchen discretization with a different conditional distribution at the initial age, Kaplan (2012) uses an age-varying, equally-

[1] The seminal contributions are Tauchen (1986), Tauchen and Hussey (1991) and Rouwenhorst (1995). Adda and Cooper (2003), Flodén (2008) and Kopecky and Suen (2010) introduce improvements for stationary, univariate, AR(1) processes. Markov-chain approximations for stationary, vector autoregressive processes have been proposed by Galindev and Lkhagvasuren (2010), Terry and Knotek (2011) and Gospodinov and Lkhagvasuren (2014). Farmer and Toda (2016) propose a method that can be applied to stationary, non-linear, multivariate processes.

[2] Non-stationarity in the income process can take the form of distributional assumptions on the initial conditions as in Huggett (1996), a unit root component as in Storesletten, Telmer and Yaron (2004), or heteroskedasticity of the innovations as in Kaplan (2012).

[3] As we discuss in the main text, the advantage of using such a process for our benchmark is that the associated optimization problem can be solved using extremely accurate numerical techniques.

spaced grid with range and transition probabilities chosen to match some moments of the original continuous process. In most cases these methods are only partially documented, hence we know very little about their performance. Our work is meant to provide a more systematic treatment of this approximation problem.

The remainder of this paper is structured as follows. Section 2 discusses how to extend Tauchen (1986) and Rouwenhorst's (1995) methods to non-stationary AR(1) processes. Section 3 compares the accuracy of the two methods. Section 4 concludes.

2 Discrete approximations of AR(1) processes

Consider an AR(1) process of the following form,

$$y_t = \rho_t y_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{id}{\sim} N(0, \sigma_{\varepsilon t}) \tag{1}$$

with initial condition y_0, where y_0 can be deterministic or a random draw from some distribution. Let σ_t denote the unconditional standard deviation of y_t. It follows from equation (1) that

$$\sigma_t^2 = \rho_t^2 \sigma_{t-1}^2 + \sigma_{\varepsilon t}^2. \tag{2}$$

In general the above process is not covariance-stationary. Sufficient conditions for stationarity are that the process in equation (1) is restricted to

$$y_t = \rho y_{t-1} + \varepsilon_t, \quad |\rho| < 1, \ \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon) \tag{3}$$

with constant persistence ρ, constant innovation variance σ_ε and y_0 randomly drawn from the asymptotic distribution of y_t; namely, $N(0, \sigma)$ where $\sigma = \sigma_\varepsilon / \sqrt{1-\rho^2}$. We call this case the stationary case in what follows, to distinguish it from the general, unrestricted process[4] in equation (1).

The aim of these notes is to show how to adapt both Tauchen (1986) and Rouwenhorst's (1995) methods to discretize a *non-stationary* AR(1) of the general form in equation (1).

2.1 Tauchen's (1986) method

2.1.1 Stationary case

Tauchen (1986) proposes the following method to discretize a stationary AR(1) process. Construct a Markov chain with a time-independent, uniformly-spaced state space $Y^N = \{\bar{y}^1, \ldots, \bar{y}^N\}$ with

$$\bar{y}^N = -\bar{y}^1 = \Omega\sigma \tag{4}$$

where Ω is a positive constant.[5] If Φ denotes the cumulative distribution function for the standard normal distribution and $h = 2\Omega\sigma/(N-1)$ the step size between grid points,

[4] Note that the general process does not restrict ρ_t to lie inside the unit circle.

[5] Tauchen (1986) sets $\Omega = 3$. Kopecky and Suen (2010) calibrate it so that the standard deviation of the Markov chain coincides with that of the original AR(1) process.

the elements of the transition matrix Π^N satisfy

$$\pi^{ij} = \begin{cases} \Phi\left(\frac{\bar{y}^j - \rho\bar{y}^i + h/2}{\sigma_\varepsilon}\right) & \text{if } j = 1, \\ \Phi\left(\frac{\bar{y}^j - \rho\bar{y}^i - h/2}{\sigma_\varepsilon}\right) & \text{if } j = N, \\ \Phi\left(\frac{\bar{y}^j - \rho\bar{y}^i + h/2}{\sigma_\varepsilon}\right) - \Phi\left(\frac{\bar{y}^j - \rho\bar{y}^i - h/2}{\sigma_\varepsilon}\right) & \text{otherwise.} \end{cases}$$

Basically, the method constructs the transition probabilities π_{ij} to equal the probability (truncated at the extremes) that y_t falls in the interval $(\bar{y}^j - h/2, \bar{y}^j + h/2)$ conditionally on $y_{t-1} = \bar{y}^i$.

2.1.2 Non-stationary case

Our non-stationary extension of Tauchen (1986) constructs a state space $Y_t^N = \{\bar{y}_t^1, \ldots, \bar{y}_t^N\}$ with constant size N, but time-varying grid-points with

$$\bar{y}_t^N = -\bar{y}_t^1 = \Omega\sigma_t \tag{5}$$

and step size $h_t = 2\Omega\sigma_t/(N-1)$. The associated transition probabilities are

$$\pi_t^{ij} = \begin{cases} \Phi\left(\frac{\bar{y}_t^j - \rho\bar{y}_{t-1}^i + h_t/2}{\sigma_{\varepsilon t}}\right) & \text{if } j = 1, \\ \Phi\left(\frac{\bar{y}_t^j - \rho\bar{y}_{t-1}^i - h_t/2}{\sigma_{\varepsilon t}}\right) & \text{if } j = N, \\ \Phi\left(\frac{\bar{y}_t^j - \rho\bar{y}_{t-1}^i + h_t/2}{\sigma_{\varepsilon t}}\right) - \Phi\left(\frac{\bar{y}_t^j - \rho\bar{y}_{t-1}^i - h_t/2}{\sigma_{\varepsilon t}}\right) & \text{otherwise.} \end{cases}$$

The main difference between our extension and its stationary counterpart is that the range of the equidistant state space in equation (5) is time varying and, as a result, so are the transition probabilities.

2.2 Rouwenhorst's (1995) method

The Rouwenhorst method is best understood as determining the parameters of a two-state Markov chain, with equally-spaced state space, in such a way that the conditional first and second moments of the Markov chain coincide with the same moments of the original AR(1) process.[6]

2.2.1 Stationary case

In the case of the stationary AR(1) process in equation (3), the state space for the two-state Markov chain is $\bar{y}^2 = -\bar{y}^1$ and the transition matrix is written as

$$\Pi^2 = \begin{bmatrix} \pi^{11} & 1 - \pi^{11} \\ 1 - \pi^{22} & \pi^{22} \end{bmatrix}. \tag{6}$$

[6]In general, a Markov chain of order K is characterized by K^2 parameters (K states plus $(K^2 - K)$ linearly-independent transition probabilities) and can be uniquely identified by K^2 linearly-independent moment conditions. The Rouwenhorst method is, therefore, a special case of a general moment-matching procedure.

The moment condition for the expectation conditional on $y_{t-1} = \bar{y}^2$ is

$$\mathrm{E}(y_t|y_{t-1} = \bar{y}^2) = -(1 - \pi^{22})\bar{y}^2 + \pi^{22}\bar{y}^2 = \rho\bar{y}^2, \tag{7}$$

where the left hand side is the conditional expectation of the Markov chain and the right hand side its counterpart for the AR(1) process for y_{t-1} evaluated at the grid point $\bar{y}^2$. It follows that

$$\pi^{22} = \frac{1+\rho}{2} = \pi^{11}, \tag{8}$$

where the second equality follows from imposing the same condition for $y_{t-1} = \bar{y}^1 = -\bar{y}^2$.

The moment condition for the variance conditional on $y_{t-1} = \bar{y}^2$ is[7]

$$\mathrm{Var}(y_t|y_{t-1} = \bar{y}^2) = (1 - \pi^{22})\left(-\bar{y}^2 - \rho\bar{y}^2\right)^2 + \pi^{22}\left(\bar{y}^2 - \rho\bar{y}^2\right)^2 = \sigma_\varepsilon^2, \tag{9}$$

which, after replacing for π^{22} from equation (8), implies

$$\bar{y}^2 = \sigma. \tag{10}$$

Having determined Π^2, the method scales to an arbitrary number of grid points N in the following way.[8] The state space $Y^N = \{\bar{y}^1, \ldots, \bar{y}^N\}$ is equally-spaced with

$$\bar{y}^N = -\bar{y}^1 = \sigma\sqrt{N-1}. \tag{11}$$

For $N \geq 3$, the transition matrix satisfies the recursion

$$\Pi^N = \pi \begin{bmatrix} \Pi^{N-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} + (1-\pi) \begin{bmatrix} \mathbf{0} & \Pi^{N-1} \\ 0 & \mathbf{0}' \end{bmatrix} + \pi \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Pi^{N-1} \end{bmatrix} + (1-\pi) \begin{bmatrix} \mathbf{0}' & 0 \\ \Pi^{N-1} & \mathbf{0} \end{bmatrix}, \tag{12}$$

where $\pi = \pi^{11} = \pi^{22}$ and $\mathbf{0}$ is an $(N-1)$ column vector of zeros.

The main difference between Rouwenhorst and Tauchen's methods is that in the former the transition probabilities do not embody the normality assumption about the distribution of the shocks. Rather, Rouwenhorst matches exactly, by construction, the first and second conditional and, by the law of iterated expectations, unconditional moments of the continuous process independently from the shock distribution.

2.2.2 Non-stationary case

As for Tauchen, our non-stationary extension of Rouwenhorst (1995) constructs an equally-spaced, symmetric, state space $Y_t^N = \{\bar{y}_t^1, \ldots, \bar{y}_t^N\}$ with constant size N but time varying grid points and transition matrix Π_t^N. If $N = 2$, it follows that $\bar{y}_t^2 = -\bar{y}_t^1$ and the counterpart of the first-moment condition (7) becomes

$$\mathrm{E}(y_t|y_{t-1} = \bar{y}_{t-1}^2) = -(1 - \pi_t^{22})\bar{y}_t^2 + \pi_t^{22}\bar{y}_t^2 = \rho_t\bar{y}_{t-1}^2,$$

with unique solution

$$\pi_t^{22} = \frac{1}{2}\left(1 + \rho_t\frac{\bar{y}_{t-1}^2}{\bar{y}_t^2}\right) = \frac{1}{2}\left(1 + \rho_t\frac{\sigma_{t-1}}{\sigma_t}\right) = \pi_t^{11}, \tag{13}$$

[7] Symmetry implies that the second conditional-variance condition is linearly dependent with equation (9) and, therefore, satisfied.

[8] We refer the reader to Rouwenhorst (1995) and Kopecky and Suen (2010) for a rigorous derivation.

where the second equality follows from the counterpart of the second moment condition (9) which implies

$$\bar{y}_t^2 = -\bar{y}_t^1 = \sigma_t. \tag{14}$$

The third equality in equation (13) follows from the expression for the conditional first moment for $y_{t-1} = \bar{y}_{t-1}$.

As in the non-stationary version of Tauchen, the points of the state-space are a function of the time-dependent unconditional variance of y_t. Comparing equations (8) and (13) reveals that, relative to the stationary case, the probability of transiting from $\bar{y}_{t-1}^2$ to $\bar{y}_t^2$ depends on the rate of growth of the unconditional variance of y_t.

Equation (13) implies that the condition for the Markov chain to be well defined, and have no absorbing states, namely $0 < \pi_t^{11} = \pi_t^{22} < 1$, is equivalent to

$$\rho_t^2 \frac{\sigma_{t-1}^2}{\sigma_t^2} < 1. \tag{15}$$

It follows from equation (2) that this condition always holds. Therefore Rouwenhorst's approximation can be applied to any process of the type defined in equation (1).[9]

As in the stationary case, the approach scales to an N-dimensional, evenly-spaced state space Y_t^N by setting

$$\bar{y}^N = -\bar{y}^1 = \sigma\sqrt{N-1} \tag{16}$$

and Π_t^N to satisfy the recursion (12) with the transition matrices and the probability $\pi_t = \pi_t^{11} = \pi_t^{22}$ indexed by t.

3 Evaluation

This section assesses the performance of the two discretization methods above in solving a finite-horizon, income-fluctuation problem with a non-stationary labor income process.

Consider the following optimization problem in recursive form[10]

$$\begin{aligned} V_t(z_t,y_t) &= \max_{c_t,a_t} \log(c_t) + \beta \mathbb{E}_t V_{t+1}(z_{t+1}, y_{t+1}) \\ \text{s.t. } & z_t = (1+r)a_{t-1} + y_t \\ & a_t = z_t - c_t \\ & y_{t+1} = y_t \epsilon_t, \quad \log \epsilon_t \overset{i.i.d.}{\sim} N(0, \sigma_\varepsilon), \\ & a_t \geq 0, \quad a_t \text{ given.} \end{aligned} \tag{17}$$

Individuals start life at age 1, with initial wealth $a_0 = 0$ and $y_0 = 1$, and live until age 40. Each model period is a year. In the computation we set the discount rate β to 0.96 and the interest rate r to .04 which are standard values. We set the variance of the labor income process $\sigma_\varepsilon^2 = .0161$, as in Storesletten et al. (2004). The parameterization implies

[9] This is also trivially true for Tauchen's method.

[10] The lower bound of zero for the choice of next period's assets is without loss of generality. It is always possible to rewrite the problem so that the lower bound on, the appropriately translated, asset space is zero.

an aggregate wealth-income ratio of about 0.6, in line with the baseline calibration in Carroll (2009) for a similar model with no retirement and deterministic lifetime.

Since the above problem does not have a closed-form solution, we evaluate the accuracy of the two discretization methods by comparing simulated moments under the two methods to those generated by a very accurate benchmark solution.

The advantage of problem (17) is that, as first shown in Carroll (2004), the combination of unit-root (in logs) income process and CRRA felicity function implies that the problem can be normalized using (permanent) labor income y_t, thereby reducing the effective state space to the single variable $\hat{z}_t = z_t/y_t$.[11]. It follows that, under the assumptions that income innovations are log-normally distributed, one can approximate the expectation in equation (17) using Gaussian-Hermite quadrature.

This allows one to solve the model using a very accurate procedure—the endogenous gridpoint method—for the optimization step[12] and Gaussian-Hermite quadrature to approximate the expectation in (17). In particular, we compute the policy functions using an exponential grid G_z with 1,000 points for the normalized state variable $\hat{z}$ and 100 quadrature nodes for the shock $\log \varepsilon_t$. Given the well-known properties of quadrature,[13] the model solution using the endogenous gridpoint method and quadrature is extremely accurate.

We simulate the model by generating 2,000,000[14] individual histories for y_t using Monte Carlo simulation of the *continuous* $AR(1)$ process and linearly interpolating the policy functions for points off the discretized state space G_z. Since, by construction, the non-normalized policy function $a_t(z_t, y_t) = \hat{a}_t(\hat{z}_t)y_t$ is linear in labor income, our benchmark simulation does not require any approximation with respect to labor income. Therefore, the simulated moments generated by our benchmark method constitute a highly accurate approximation to the true model moments.

Next, we compute the same set of moments by applying the same optimization method as in the benchmark but using either Tauchen or Rouwenhorst's methods to discretize the labor income process. To be precise, in each case we solve the (non-normalized) decision problem (17) by replacing the continuous income process with the appropriate Markov chain with age-dependent grids Y_t^N and transition matrices Π_t^N and using a common exponential grid G_z with 1,000 points for z_t. We consider three different values for the income grid size N; namely 5, 10 and 25.

Given the policy functions thus obtained, we compute the model moments using a Monte Carlo simulation which again generates 2,000,000 income histories. This is done in two different ways. In the first case, we generate the income histories using the discrete Markov chain approximation. The simulation involves interpolating the policy functions linearly only with respect to z. In the second case, as in the benchmark quadrature case, we generate income histories using the *continuous* $AR(1)$ process. We then interpolate linearly over both z *and* labor income y.

[11] The Appendix reports the derivation

[12] See Barillas and Fernández-Villaverde (2007) for an assessment of the accuracy of the endogenous grid method.

[13] Given n quadrature nodes, Gaussian quadrature approximates exactly the integral of any polynomial function of degree up to $2n - 1$.

[14] Increasing the number of individuals histories to 20,000,000 does not affect the results in any meaningful way.

The key difference between these two approaches has to do with the sources of the errors that they introduce. Both cases suffer from approximation errors for the policy function relative to quadrature due to: (a) the suboptimal approximation of the expectation in (17); (b) the fact that the policy functions solve the Euler equations exactly only at a relatively small number of grid points for labor income. Compared to the continuous $AR(1)$ simulation, the Markov chain simulation introduces an additional approximation error as the simulated policy functions are step, rather than piecewise-linear, functions along the income dimension.

Table 1: Ratio of model moments relative to their counterpart in the quadrature benchmark: (A) Markov chain simulation and (B) continuous random walk income process

		$N=5$			$N=10$			$N=25$		
		R	T_{Ω^*}	$T_{\Omega=3}$	R	T_{Ω^*}	$T_{\Omega=3}$	R	T_{Ω^*}	$T_{\Omega=3}$
					(A) Markov chain simulation					
Labor income (y_t)	Mean	0.9960	0.9939	1.0880	0.9975	0.9952	1.0543	0.9983	0.9969	1.0070
	SD	0.9208	0.8798	1.3993	0.9618	0.9085	1.2329	0.9842	0.9490	1.0159
	Gini	0.9574	0.9608	1.1255	0.9815	0.9817	1.1069	0.9928	0.9942	1.0169
Consumption (c_t)	Mean	0.9966	0.9882	1.0755	0.9978	1.0006	1.0546	0.9984	0.9988	1.0093
	SD	0.9253	0.8606	1.3517	0.9640	0.9242	1.2140	0.9850	0.9558	1.0213
	Gini	0.9630	0.9634	1.1392	0.9851	0.9750	1.1045	0.9949	0.9926	1.0148
Wealth (a_t)	Mean	1.0186	0.7385	0.5370	1.0083	1.2330	1.0706	1.0026	1.0795	1.1079
	SD	1.0611	0.7689	1.1376	1.0296	1.6243	0.9059	1.0110	1.2334	1.1987
	Gini	1.1088	1.3562	3.0492	1.0521	1.5599	0.6049	1.0198	1.2017	1.0961
Tot. inc. ($ra_{t-1}+y_t$)	Mean	0.9966	0.9882	1.0755	0.9978	1.0006	1.0546	0.9984	0.9988	1.0093
	SD	0.9232	0.8656	1.3723	0.9630	0.9190	1.2181	0.9846	0.9536	1.0190
	Gini	0.9607	0.9549	1.1071	0.9834	0.9885	1.0966	0.9938	0.9964	1.0174
Top 5% wealth share		1.0367	0.8449	1.6747	1.0217	1.3214	0.7299	1.0090	1.1451	1.0785
					(B) Random walk simulation					
Labor income (y_t)	Mean	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	SD	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Gini	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Consumption (c_t)	Mean	1.0005	0.9941	0.9886	1.0002	1.0065	1.0807	1.0001	1.0023	1.0025
	SD	1.0027	0.9802	0.9694	1.0013	1.0339	0.8926	1.0005	1.0176	1.0094
	Gini	0.9994	1.0020	1.0200	0.9998	0.9939	0.5290	0.9999	0.9986	0.9974
Wealth (a_t)	Mean	1.0232	0.7409	0.4950	1.0106	1.2889	1.0018	1.0039	1.1037	1.1111
	SD	1.0870	0.7862	0.9270	1.0418	2.2010	0.9888	1.0158	1.6298	1.3198
	Gini	1.1116	1.3533	2.8598	1.0527	1.7360	0.9890	1.0201	1.2945	1.1070
Tot. inc. ($ra_{t-1}+y_t$)	Mean	1.0005	0.9941	0.9886	1.0002	1.0065	1.0018	1.0001	1.0023	1.0025
	SD	1.0018	0.9861	0.9792	1.0009	1.0203	0.9877	1.0003	1.0104	1.0053
	Gini	1.0012	0.9944	0.9832	1.0006	1.0086	0.9942	1.0002	1.0030	1.0004
Top 5% wealth share		1.0670	0.8994	1.5462	1.0338	1.6214	0.7129	1.0130	1.2944	1.1086

Note: Parameter values: $\beta = 0.96,\ r = 0.04,\ \sigma_\varepsilon^2 = 0.0161$.
For columns T_{Ω^*}, $\Omega = 1.6919$ when $N = 5$, $\Omega = 2.0513$ when $N = 10$ and $\Omega = 2.5996$ when $N = 25$.

3.1 Results

We evaluate the accuracy of Tauchen and Rouwenhorst's discretization methods by comparing simulated moments obtained under the benchmark quadrature approach to those obtained under either of the two discretization methods. The moments are: (i) the unconditional mean; (ii) the unconditional standard deviation; and (iii) the Gini coefficient. Each set of moments is reported for the distributions of labor income, consumption, wealth and total income. Given the increasing interest in wealth concentration, at the bottom of each panel we also report the share of aggregate wealth held by the households in the top 5% of the wealth distribution.

Panel (A) and (B) in Table 1 report the ratio of the moments obtained from simulating the income process using Rouwenhorst and Tauchen's discretization methods to those computed for the Gaussian Hermite benchmark solution. In the table a value of one indicates that the approximation entails no error, relative to the benchmark solution. As shown in Flodén (2008) and Kopecky and Suen (2010), Tauchen's method is very sensitive to the choice of Ω. Tauchen (1986) originally sets $\Omega = 3$, while Kopecky and Suen (2010) calibrate Ω to match the variance of log income. The counterpart of the latter strategy for a non-stationary income process is not obvious. Hence we choose Ω to match the variance of log income over the whole population, and we report results both for this parametrization (columns T_{Ω^*}) and for the case in which $\Omega = 3$ (columns $\mathrm{T}_{\Omega=3}$).

Case 1: Markov chain simulation. Panel (A) shows results for the case in which the discretized income process is used both to compute the expectation in the decision problem and to simulate the model. In this case the Rouwenhorst method and the Tauchen method with "optimal" choice of Ω perform quite similarly in approximating the labor income moments and the first moment of the consumption distribution. As expected the Tauchen method with $\Omega = 3$ performs much worse. The Rouwenhorst method, though, is more accurate with respect to the standard deviation of consumption, and substantially more so with respect to the wealth distribution. In the latter case, the Rouwenhorst approximation has a maximum error (for any of the moments) of at most 11 per cent for $N = 5$ and of only 2 per cent for $N = 25$. In contrast, the Tauchen approximation is off by anywhere between 1/4 and 2 times relative to the benchmark quadrature method. The approximation is particularly poor, even with a large number of points, for both the top 5% share and the variance of wealth. Moreover, it is apparent that the approximation error does not necessarily shrink as the number of grid points increases. Intuitively, when comparing the range of the income grid for the Tauchen (equation (5)) and Rouwenhorst (equation (11)) methods, the range of the income grid increases faster with N for the latter method. This implies that, in the case of Tauchen, a larger number of simulated observations get piled onto the bounds relative to the benchmark method, reducing accuracy. This problem appears to be quite important when approximating the standard deviation of wealth holdings. This conjecture is confirmed by the fact that the Tauchen method with $\Omega = 3$, hence with a larger labor income range, performs better than the one with the "optimal choice" of Ω in this respect.

Case 2: Random walk simulation. Panel (B) in Table 1 reports the approximation errors obtained through Monte Carlo simulation using the continuous income process. By construction, there is no approximation error for the income process in this case.

As expected, the accuracy of both the Tauchen and Rouwenhorst methods generally improves relative to results for the Markov chain simulation. In fact, the accuracy of the Rouwenhorst method is extremely high even when $N = 5$.

Concerning the wealth moments, the performance of the Rouwenhorst method is similar to that obtained for the Markov chain simulation. The performance of the Tauchen method is, if anything, worse suggesting that, given the narrower income grid relative to Rouwenhorst, extrapolation along the income dimension increases the overall error relative to the Markov chain simulation. In fact, for N larger than 5 the Tauchen method with $\Omega = 3$, hence with a larger labor income range, performs better than the one with "optimal choice" of Ω.

In sum, the Rouwenhorst method exhibits considerable accuracy even when using a small number of grid points, and its performance is substantially more robust across all moments considered and for all numbers of grid points.

4 Conclusion

Approximating non-stationary processes is commonplace in quantitative studies of life-cycle behavior and inequality. In such studies it is important to reliably model the fanning out over age of the cross-sectional distribution of consumption, income and wealth. Large approximation errors may result in misleading inference and the problem appears to be especially severe when approximating the distribution of wealth.

In this paper we provide the first systematic examination of the performance of alternative methods to approximate non-stationary (state-dependent) income processes within a life-cycle setting. We begin by explicitly deriving new generalizations of the Tauchen and Rouwenhorst's approximation methods for the case of history-dependent state spaces, like the ones commonly employed in life-cycle economies. We then compare the relative performance of these approximation methods. For each method, we numerically solve a finite-lifetime, income-fluctuation problem, and compute a set of moments for the implied cross-sectional distributions of income, consumption and wealth. Next, we gauge the relative performance of the two methods by comparing these moments to the ones obtained from a quasi-exact solution of the same income-fluctuation problem.

The results of this comparison are quite clear and suggest that, in a life-cycle setting, Tauchen's method is generally much less precise than Roewenhorst's. This discrepancy is most severe when considering the distribution of wealth. Perhaps more worrying is the fact that adding grid points to the income approximation does not seem to significantly improve the performance of Tauchen's method. In contrast, increasing the number of grid points does improve the accuracy of the Roewenhorst approximation. However, we find that the latter method offers a very reliable approximation even with just 5 grid points.

Appendix

Normalized problem with unit-root labor income.

In the case in which the (log) income process has a unit root and the felicity function has the CRRA form $u(c) = c^{1-\gamma}/(1-\gamma)$, it is well known from Carroll (2004) that it is possible to normalize problem (17) by (permanent) labor income y_t, thereby reducing the effective state space to z_t.

To see this, replace for $c_t = z_t - a_t$ in (17) and consider the problem in the second-to-last period

$$\mathbb{V}_{T-1}(z_{T-1}, y_{T-1}) = \max_{a_{T-1}} u(z_{T-1} - a_{T-1}) + \beta \mathbb{E}_{T-1} u(z_T) \tag{18}$$

If one defines the state variables $\hat{z}_t = z_t/y_t$ and $\hat{a}_t = a_t/y_t$, equation (18) can be rewritten as

$$\begin{aligned} \mathbb{V}_{T-1}(z_{T-1}, y_{T-1}) &= \max_{\hat{a}_{T-1}} u(y_{T-1}(\hat{z}_{T-1} - \hat{a}_{T-1})) + \beta \mathbb{E}_{T-1} u(y_T \hat{z}_T) \\ &= y_{T-1}^{1-\gamma} \left\{ \max_{\hat{a}_{T-1}} u(\hat{z}_{T-1} - \hat{a}_{T-1}) + \beta \mathbb{E}_{T-1} \epsilon_T^{1-\gamma} u(\hat{z}_T) \right\} \end{aligned} \tag{19}$$

Note that by definition

$$\hat{z}_t = (1+r)\frac{a_{t-1}}{y_{t-1}\epsilon_t} + 1 = (1+r)\frac{\hat{a}_{t-1}}{\epsilon_t} + 1, \tag{20}$$

which implies that the curly bracket in (19) is equal to $V_{T-1}(\hat{z}_{T-1})$ where the latter satisfies the Bellman equation

$$V_{T-1}(\hat{z}_{T-1}) = \max_{\hat{a}_{T-1}} u(\hat{z}_{T-1} - \hat{a}_{T-1}) + \beta \mathbb{E}_{T-1} \epsilon_T^{1-\gamma} V_T(\hat{z}_T) \tag{21}$$

with $V_T(\hat{z}_T) = u(\hat{z}_T)$.

Equations (19) and (21) imply that $\mathbb{V}_{T-1}(z_{T-1}, y_{T-1}) = y_{T-1}^{1-\gamma} V(\hat{z}_{T-1})$. The same logic implies that this holds also for any $t < T-1$.

Therefore the Bellman equation for the problem in normalized form satisfies

$$V_t(\hat{z}_t) = \max_{\hat{a}_t} u(\hat{z}_t - \hat{a}_t) + \beta \mathbb{E}_t \epsilon_{t+1}^{1-\gamma} V_{t+1}(\hat{z}_{t+1}), \tag{22}$$

for all t. It follows from (20) and the envelope condition that the associated Euler equation is

$$u'(\hat{c}_t) = \beta R \mathbb{E}\left[\epsilon_{t+1}^{-\rho} u'(\hat{c}_{t+1})\right] \tag{23}$$

The advantage of the normalized problem (21) is that one can solve for the saving function $\hat{a}_t(\hat{z}_t)$ which is independent of the income realization y_t and use $a_t(z_t, y_t) = \hat{a}_t(\hat{z}_t) y_t$ to recover the policy function for a_t.

Under the assumption that ϵ_t is i.i.d. and log-normally distributed the expectation in equation (21) can be computed using Gaussian Hermite quadrature.

References

Adda, Jerome and Cooper, Russell W. (2003), *Dynamic Economics: Quantitative Methods and Applications*, MIT Press.

Barillas, Francisco and Fernández-Villaverde, Jésus (2007), 'A generalization of the endogenous grid method', *Journal of Economic Dynamics and Control* **31**, 2698–2712.

Carroll, Christopher (2004), Theoretical foundations of buffer stock saving, Working Paper 10867, National Bureau of Economic Research.

Carroll, Christopher D. (2009), 'Precautionary saving and the marginal propensity to consume out of permanent income', *Journal of Monetary Economics* **56**, 780–790.

Deaton, Angus and Paxson, Christina (1994), 'Intertemporal choice and inequality', *Journal of Political Economy* **102**(3), 437–67.

Farmer, Leland E. and Toda, Alexis Akira (2016), 'Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments', *Quantitative Economics* . forthcoming.

Flodén, Martin (2008), 'A note on the accuracy of Markov-chain approximations to highly persistent AR(1) processes', *Economics Letters* **99**(3), 516 – 520.

Galindev, Ragchaasuren and Lkhagvasuren, Damba (2010), 'Discretization of highly persistent correlated AR(1) shocks', *Journal of Economic Dynamics and Control* **34**(7), 1260 – 1276.

Gospodinov, Nikolay and Lkhagvasuren, Damba (2014), 'A moment-matching method for approximating vector autoregressive processes by finite-state markov chains', *Journal of Applied Econometrics* **29**(5), 843–859.

Huggett, Mark (1996), 'Wealth distribution in life-cycle economies', *Journal of Monetary Economics* **38**(3), 469–494.

Kaplan, Greg (2012), 'Inequality and the life cycle', *Quantitative Economics* **3**(3), 471–525.

Kopecky, Karen and Suen, Richard (2010), 'Finite State Markov-chain Approximations to Highly Persistent Processes', *Review of Economic Dynamics* **13**(3), 701–714.

Rouwenhorst, Geert (1995), Asset pricing implications of equilibrium business cycle models, *in* Thomas F.Cooley, ed., 'Frontiers of business cycle research', Princeton University Press, chapter 10.

Storesletten, Kjetil, Telmer, Christopher I. and Yaron, Amir (2004), 'Consumption and risk sharing over the life cycle', *Journal of Monetary Economics* **51**(3), 609–633.

Tauchen, George (1986), 'Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions', *Economics Letters* **20**(2), 177–181.

Tauchen, George and Hussey, Robert (1991), 'Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models', *Econometrica* **59**(2), 371–96.

Terry, Stephen J. and Knotek, Edward S. II (2011), 'Markov-chain approximations of vector autoregressions: Application of general multivariate-normal integration techniques', *Economics Letters* **110**(1), 4 – 6.