Finite State Markov-Chain Approximations to Highly Persistent Processes

Kopecky, Karen A. and Suen, Richard M. H.

University of California, Riverside

8 May 2009

Finite State Markov-Chain Approximations to Highly Persistent Processes*

Karen A. Kopecky† Richard M. H. Suen‡

First Version: May 2009

This Version: August 2009

Abstract

This paper re-examines the Rouwenhorst method of approximating first-order autoregressive processes. This method is appealing because it can match the conditional and unconditional mean, the conditional and unconditional variance and the first-order autocorrelation of any AR(1) process. This paper provides the first formal proof of this and other results. When comparing to five other methods, the Rouwenhorst method has the best performance in approximating the business cycle moments generated by the stochastic growth model. In addition, when the Rouwenhorst method is used, moments computed directly off the stationary distribution are as accurate as those obtained using Monte Carlo simulations.

Keywords: Numerical Methods, Finite State Approximations, Optimal Growth Model

JEL classification: C63.

*We thank Russell Cooper, John Rust, and conference participants at the UCR Conference on Business Cycles and the Workshop for Economists Working in Parallel for helpful comments and suggestions. We thank Yundong Tu for excellent research assistance.

†Department of Economics, Social Science Center, Room 4701, The University of Western Ontario, London, Ontario, N6A 5C2, Canada. Email: *kkopecky@uwo.ca*.

‡Corresponding author. Department of Economics, Sproul Hall, University of California, Riverside CA 92521-0427. Email: *richard.suen@ucr.edu*. Tel.: (951) 827-1502. Fax: (951) 827-5685.

1 Introduction

In macroeconomic models, the exogenous stochastic process is typically assumed to follow a stationary first-order autoregressive process. Two well-known examples are the asset pricing model à la Lucas (1978), and the standard real business cycle (RBC) model. In Lucas' model, the stochastic dividend stream is assumed to follow a Markov process. In the RBC model, the logarithm of the productivity shock is assumed to follow a Gaussian AR(1) process. In order to solve these models numerically, the continuous-valued autoregressive process is usually approximated by a discrete state-space Markov chain. To this end, researchers typically employ the approximation method proposed by Tauchen (1986), or the quadrature-based method developed in Tauchen and Hussey (1991). Although these methods differ substantially in details, the underlying idea is the same, that is to construct a discrete state-space Markov chain with transition probabilities that provide a good approximation for the conditional density of the autoregressive process. For AR(1) processes with low persistence, these methods can generate an accurate approximation even when a very coarse state space is used in the approximate Markov chain. However, the performance of these methods deteriorates when the serial correlation is very close to one, a problem that has been examined in the recent studies by Flodén (2008) and Lkhagvasuren and Galindev (2008).[1] These studies show that the accuracies of the Tauchen (1986) and the Tauchen-Hussey method are significantly lowered when the serial correlation of the underlying process is greater than 0.95, and that the problem persists even if one increases the number of states in the Markov chain.

This problem with the Tauchen (1986) and the Tauchen-Hussey method raises concerns because macroeconomic studies often employ highly persistent processes. This calls for a more reliable technique to approximate highly autocorrelated processes. The main objective of this paper is to consider such a technique. More specifically, the current study re-examines a discrete approximation method first proposed in Rouwenhorst (1995). Similar to the aforementioned methods, the Rouwenhorst method is about the construction of an approximate discrete state-space Markov chain. But unlike the other methods, the transition probabilities of the Markov chain are not intended to mimic the conditional distribution of the underlying AR(1) process. This might seem like a weakness at first, but the Rouwenhorst method has a number of desirable features that are not matched by

[1] This weakness is also acknowledged in the original papers. In Tauchen (1986, p.179), the author notes that "Experimentation showed that the quality of the approximation remains good except when λ [the serial correlation] is very close to unity." In Tauchen and Hussey (1991), the authors note that for processes with high persistence, "adequate approximation requires successively finer state spaces."

the other methods. First, only a few parameters are used in constructing the approximate Markov chain under this method. It is thus much more parsimonious and much easier to implement than the quadrature-based methods. Second, the constructed Markov chain can be calibrated to match five important statistics of *any* stationary AR(1) process. These are the conditional and unconditional mean, the conditional and unconditional variance, and the first-order autocorrelation. Thus, even though the transition probabilities of the Markov chain do not mimic the conditional distribution of the underlying AR(1) process, it can still exactly match the first two moments. Third, the Rouwenhorst method is particularly desirable for approximating Gaussian AR(1) processes. This is because the invariant distribution of the constructed Markov chain is a binomial distribution, which converges to the standard normal distribution when the number of states in the state space is sufficiently large.

Some of these features have been mentioned in Rouwenhorst (1995). But a formal proof of these results is still lacking. It is also unclear whether matching the moments of the AR(1) process is important in terms of solving dynamic general equilibrium models. In quantitative studies, obtaining a good approximation for the AR(1) process is seldom an end in itself. Thus a more appropriate metric for evaluating approximation methods in general would be their impact on the computed solutions of the general equilibrium models. Very few attempts have been made to assess the relative performance of the Rouwenhorst method and other approximation methods on this ground. Thus it remains unclear how the choice of approximation method would affect the accuracies of the computed solutions in these models. The current study is intended to fill these gaps.

The main contribution of this paper is two-fold. First, the paper provides formal proofs of all the results mentioned above. These results encompass the claims made in Rouwenhorst (1995). They also extend and generalize those claims in two ways. (i) Rouwenhorst mentions that when the transition matrix of the approximate Markov chain is symmetric, the invariant distribution is given by a binomial distribution. The current study shows that the invariant distribution is binomial even if the symmetric assumption is relaxed. (ii) Rouwenhorst also claims that in the symmetric case, the approximate Markov chain can be calibrated to match the unconditional mean, the unconditional variance and the first-order autocorrelation of any stationary AR(1) process. This paper shows that the Markov chain can also match the conditional mean and the conditional variance.

The second contribution of this paper is to compare the Rouwenhorst method to five other approximation methods that are commonly used in the literature. These include the Tauchen (1986)

method, the original quadrature-based method developed in Tauchen and Hussey (1991), two variations of this method considered in Flodén (2008), and the Adda-Cooper (2003) method. To achieve this, the prototypical stochastic neoclassical growth model without leisure is used as the analytical vehicle.[2] There are two main reasons why we choose this particular model. First, the neoclassical growth model is by far the most common analytical framework in macroeconomics. Variations of the original model have been used to study a wide range of economic issues. Second, it is possible to derive closed-form solutions for the neoclassical growth model under certain specifications. This property of the model provides tremendous convenience for evaluating the accuracy of the approximation methods.

The main criterion for evaluating the six approximation methods is the accuracy in approximating the business cycle moments as predicted by the stochastic growth model. Two approaches to generating these moments are considered. In the baseline approach, an approximation for the stationary distribution of the state variables is first derived. The moments of interest are then computed directly from this distribution. In the second approach, the business cycle moments are generated using the Monte Carlo simulation method. This involves simulating the model using the actual AR(1) process and the computed policy function, and thus does not require approximating the stationary distribution. One major difference between these two approaches is the sources of the errors that they introduce. While both methods suffer from errors in the computation of the policy function, under the baseline approach, additional errors arise when approximating the stationary distribution. However, this approach does not suffer from the sampling errors that the simulation method generates.

One important finding of this paper is that, regardless of which approach is taken, the choice of approximation method can have a large impact on the accuracy of the computed business cycle moments. Under the baseline approach, the choice of discretization method has a large impact on the accuracy of the stationary distribution approximation that is used to compute the moments. In general, a method that generates a good approximation for the moments of the AR(1) process also tends to yield an accurate approximation for the stationary distribution. The Rouwenhorst method has the best performance in this regard, followed by an improved version of the Tauchen (1986) method. In the sensitivity analysis, it is shown that the superior performance of the Rouwenhorst

[2] The same model is used in Taylor and Uhlig (1990) and the companion papers to illustrate and compare different solution methods. More recently, Aruoba, Fernández-Villaverde and Rubio-Ramírez (2006) use the same model, but with labor-leisure choice, to compare different solution methods.

method is robust under a wide range of parameter values.

When the Monte Carlo simulation method is used to generate the business cycle moments, no single method dominates all others in all cases. With a logarithmic utility function and full depreciation, the six methods yield almost identical results. When a more realistic value of the depreciation rate is used, the relative performance of the six methods depends on the number of states in the Markov chain. When a rather coarse state space is used, the Rouwenhorst method again has the best overall performance. However, when the fineness of the state space increases, the modified Tauchen (1986) method, the original Tauchen-Hussey method and its variation all perform as well as the Rouwenhorst method.

Another interesting finding is that the baseline approach, equipped with the Rouwenhorst method, performs as well as the simulation method in generating the business cycle moments. This result is of interest because the simulation method is considered standard practice in estimating unknown statistics of stochastic models. However, our results show that a high degree of accuracy in the business cycle moments generated from the neoclassical growth model can be achieved without simulation.

The current study is closest in spirit to Flodén (2008) and Lkhagvasuren and Galindev (2008). The main objective of Flodén (2008) is to compare the relative performance of various discretization methods in approximating univariate AR(1) processes. The author finds that existing methods, such as the Tauchen (1986) method and the Tauchen-Hussey method, are not suitable for approximating very persistent processes. He then proposes a variation of the original Tauchen-Hussey method which is more robust than the other methods. There are two major differences between this and the current study. First, Flodén does not consider the Rouwenhorst method. Second, this author does not consider the impact of the discretization procedure on the solutions of dynamic general equilibrium models. Lkhagvasuren and Galindev (2008) is another recent study on the same issue. The main objective of this paper is to develop an approximation method for vector autoregressive processes with correlated error terms. Under the proposed method, the original multivariate process is decomposed into a number of independent univariate AR(1) processes. These independent processes are then approximated using the conventional methods. Lkhagvasuren and Galindev show, through a few numerical examples, that the Rouwenhorst method outperforms other methods in approximating moments of univariate AR(1) processes. In contrast, the current study formally proves that the Rouwenhorst method can be used to match exactly a number of key statistics of

any stationary AR(1) process.

The rest of this paper is organized as follows. Section 2 presents the Rouwenhorst method and the analytical results pertaining to this method. Section 3 presents the numerical results. Section 4 concludes.

2 The Rouwenhorst Method

Consider the following AR(1) process

$$z_t = \rho z_{t-1} + \varepsilon_t, \qquad \text{with } |\rho| < 1, \tag{1}$$

and ε_t is a white noise with variance σ_ε^2. The AR(1) process is covariance-stationary with mean zero and variance

$$\sigma_z^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}. \tag{2}$$

In addition, if ε_t is normally distributed in each period, then z_t is also normally distributed.

Rouwenhorst (1995) proposes a discrete approximation to the AR(1) process in (1). This involves constructing an N-state Markov chain characterized by (i) a symmetric and evenly-spaced state space $Y_N = \{\overline{y}_1, ..., \overline{y}_N\}$, with $\overline{y}_1 = -\psi$ and $\overline{y}_N = \psi$, and (ii) a transition matrix Θ_N. For any $N \geq 2$, the transition matrix is determined by two parameters, $p, q \in (0,1)$, and is defined recursively as follows:

Step 1: When $N = 2$, define Θ_2 as

$$\Theta_2 = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}.$$

Step 2: For $N \geq 3$, first construct the N-by-N matrix

$$\begin{aligned} & p \begin{bmatrix} \Theta_{N-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} + (1-p) \begin{bmatrix} \mathbf{0} & \Theta_{N-1} \\ 0 & \mathbf{0}' \end{bmatrix} \\ & + (1-q) \begin{bmatrix} \mathbf{0}' & 0 \\ \Theta_{N-1} & \mathbf{0} \end{bmatrix} + q \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Theta_{N-1} \end{bmatrix}, \end{aligned} \tag{3}$$

where $\mathbf{0}$ is a $(N-1)$-by-1 column vector of zeros.

Step 3: Divide all but the top and bottom rows by two so that the elements in each row sum to one.

One problem with the Rouwenhorst method is that the matrix Θ_N generated by the three-step procedure is very difficult to work with analytically. For this reason, we begin our analysis by offering a new, analytically tractable procedure for generating the Rouwenhorst matrix. Using this new procedure, it is shown that a Markov chain with state space Y_N and transition matrix Θ_N has a unique invariant distribution in the form of a binomial distribution.

2.1 Reconstructing the Rouwenhorst Matrix

For any $p, q \in (0,1)$, and for any integer $N \geq 2$, define a system of polynomials as follows

$$\Phi(t; N, i) \equiv [p + (1-p)\, t]^{N-i} (1 - q + qt)^{i-1}, \tag{4}$$

for $i = 1, 2, ..., N$. The polynomials in (4) can be expanded to become

$$\Phi(t; N, i) = \sum_{j=1}^{N} \pi_{i,j}^{(N)} t^{j-1}, \qquad \text{for } i = 1, 2, ..., N. \tag{5}$$

Define an N-by-N matrix $\Pi_N = \left[\pi_{i,j}^{(N)}\right]$ using the coefficients in (5). Using the generating function in (4), one can derive the elements in Π_N recursively using the elements in Π_{N-1}, for $N - 1 \geq 2$. The details of this procedure are described in Appendix A. The main result of this subsection is Proposition 1 which states that the matrix Π_N is identical to the Rouwenhorst matrix Θ_N for any integer $N \geq 2$. All proofs can be found in Appendix B.

Proposition 1 *For any $N \geq 2$, and for any $p, q \in (0,1)$, the matrix Π_N defined above is identical to the Rouwenhorst matrix Θ_N generated by Steps 1-3.*

The next result states that Π_N is a stochastic matrix with non-zero entries. To begin with, set $t = 1$ in both (4) and (5) to obtain

$$\sum_{j=1}^{N} \pi_{i,j}^{(N)} = 1, \qquad \text{for } i = 1, 2, ..., N.$$

This means the elements in any row of Π_N sum to one. If, in addition, $\pi_{i,j}^{(N)} \geq 0$ for all i and j, then Π_N is a stochastic matrix. This is proved in the following lemma.

Lemma 2 *For any $N \geq 2$, the matrix Π_N defined above is a stochastic matrix with no zero entries.*

2.2 Discrete State-Space Markov Chain

Consider a Markov chain with a symmetric and evenly-spaced state space $Y_N = \{\overline{y}_1, ..., \overline{y}_N\}$ defined over the interval $[-\psi, \psi]$. The transition matrix of the Markov chain is given by $\Pi_N = \left[\pi_{i,j}^{(N)}\right]$ as defined above. The following result follows immediately from Lemma 2.

Proposition 3 *For any $N \geq 2$, the Markov chain with state space $Y_N = \{\overline{y}_1, ..., \overline{y}_N\}$ and transition matrix Π_N has a unique invariant distribution $\boldsymbol{\lambda}^{(N)} = \left(\lambda_1^{(N)}, ..., \lambda_N^{(N)}\right)$, where $\lambda_i^{(N)} \geq 0$ and $\sum_{i=1}^{N} \lambda_i^{(N)} = 1$.*

Rouwenhorst mentions that in the symmetric case where $p = q$, the unique invariant distribution is a binomial distribution with parameters $N - 1$ and $1/2$. The main result of this subsection is to show that the unique invariant distribution is binomial for any $p, q \in (0, 1)$.

Since the invariant distribution is unique, it can be solved by the guess-and-verify method. Let $s \equiv \frac{1-q}{2-(p+q)} \in (0, 1)$. The guess for $\boldsymbol{\lambda}^{(N)}$, represented by $\widehat{\boldsymbol{\lambda}}^{(N)}$, is a binomial distribution with parameters $N - 1$ and $1 - s$. This means

$$\widehat{\lambda}_i^{(N)} = \binom{N-1}{i-1} s^{N-i} (1-s)^{i-1}, \quad \text{for } i = 1, 2, ..., N. \tag{6}$$

It is easy to check that this is the actual solution when $N = 2$. The result for the general case is established in Proposition 4.

Proposition 4 *For any $N \geq 2$, the invariant distribution of the Markov chain defined above is a binomial distribution with parameters $N - 1$ and $1 - s$.*

Some of the conditional and unconditional moments of the Markov chain are listed in Table 1. The mathematical derivations of these results can be found in Appendix C.

Table 1: Selected Moments of the Markov Chain

Conditional Mean	$E\left(y_{t+1}\middle\vert y_t=\overline{y}_i\right)$	$(q-p)\,\psi+(p+q-1)\,\overline{y}_i$
Conditional Variance	$var\left(y_{t+1}\middle\vert y_t=\overline{y}_i\right)$	$\frac{4\psi^2}{(N-1)^2}\left[(N-i)\,(1-p)\,p+(i-1)\,q\,(1-q)\right]$
Unconditional Mean	$E\left(y_t\right)$	$\frac{(q-p)\psi}{2-(p+q)}$
Unconditional Second Moment	$E\left(y_t^2\right)$	$\psi^2\left\{1-4s\,(1-s)+\frac{4s(1-s)}{N-1}\right\}$
First-order Autocovariance	$Cov\left(y_t,y_{t+1}\right)$	$(p+q-1)\,var(y_t)$
First-order Autocorrelation	$Corr(y_t,y_{t+1})$	$p+q-1$

2.3 Approximating AR(1) Processes

The task at hand is to approximate a given stationary AR(1) process with an N-state Markov chain. Let $\{z_t\}$ be a stationary AR(1) process as defined in (1). Conditional on the realization of z_{t-1}, the mean and variance of z_t are given by

$$E\left(z_t|z_{t-1}\right)=\rho z_{t-1} \qquad \text{and} \qquad var\left(z_t|z_{t-1}\right)=\sigma_\varepsilon^2.$$

Next, define an N-state discrete Markov process $\{y_t\}$ as in Section 2.2 with

$$p=q=\frac{1+\rho}{2} \qquad \text{and} \qquad \psi=\sqrt{N-1}\sigma_\varepsilon. \tag{7}$$

Using the equations listed in Table 1, it is immediate to see that the resulting Markov chain has the same unconditional mean, unconditional variance and first-order autocorrelation as $\{z_t\}$. Suppose $y_t=\overline{y}_i$ for some $t\geq 0$ and for some $\overline{y}_i$ in the state space Y_N. The conditional mean and conditional variance of y_{t+1} are given by

$$E\left(y_{t+1}|y_t=\overline{y}_i\right)=\rho\overline{y}_i \qquad \text{and} \qquad var\left(y_{t+1}|y_t=\overline{y}_i\right)=\sigma_\varepsilon^2.$$

Thus the Markov chain $\{y_t\}$ has the same conditional mean and conditional variance as the AR(1) process $\{z_t\}$.

Two remarks regarding this procedure are worth mentioning. First, under the Rouwenhorst method, the approximate Markov chain is constructed using ρ and σ_ε^2 alone. In particular, the

transition matrix Π_N is not a discretized version of the conditional distribution of z_t. This is the fundamental difference between this method and the ones proposed by Tauchen (1986) and Tauchen and Hussey (1991). Second, the above procedure can be applied to *any* stationary AR(1) process, including those with very high persistence. Thus, unlike the other two methods, the one proposed by Rouwenhorst can always match the unconditional variance and the persistence of z_t.

Suppose now that the disturbance term ε_t in the AR(1) process is normally distributed in each period t. Then the distribution of z_t is a normal distribution. In this case, the invariant distribution of the Markov chain $\{y_t\}$ can provide a good approximation for the distribution of z_t. As shown in Proposition 4, the invariant distribution of y_t is always given by a binomial distribution. Under (7), the mean and variance of the invariant distribution are zero and $\sigma^2 \equiv \sigma_\varepsilon^2/\left(1-\rho^2\right)$, respectively. Thus the standardized process $\{y_t/\sigma\}$ would converge to the standard normal distribution when N is made sufficiently large. According to the Berry-Esséen Theorem, the rate of convergence is on the order of $N^{-1/2}$. This property is also mentioned in Rouwenhorst (1995).

3 Stochastic Neoclassical Growth Model

Consider the planner's problem in the stochastic neoclassical growth model,

$$\max_{\{C_t,K_{t+1}\}_{t=0}^{\infty}} E_0\left[\sum_{t=0}^{\infty}\beta^t U\left(C_t\right)\right]$$

subject to

$$C_t + K_{t+1} = A_t K_t^{\alpha} + (1-\delta)\,K_t,$$

$$C_t, K_{t+1} \geq 0,$$

where C_t denotes consumption at time t, K_t denotes capital and A_t is the stochastic technological factor. The function $U(\cdot)$ is the per-period utility function. The parameter $\beta \in (0,1)$ is the subjective discount factor, $\alpha \in (0,1)$ is the share of capital income in total output and $\delta \in (0,1]$ is the depreciation rate of capital. The logarithm of the technological shock, represented by $a_t \equiv \ln A_t$, is assumed to follow an AR(1) process,

$$a_{t+1} = \rho a_t + \varepsilon_{t+1}, \tag{8}$$

where $\varepsilon_{t+1} \sim$ i.i.d. $N\left(0, \sigma_{\varepsilon}^{2}\right)$ and $\rho \in (0,1)$. Conditional on $a_t = a$, the random variable a_{t+1} is normally distributed with mean ρa and variance σ_{ε}^{2}. Let $F(\cdot|a)$ be the conditional distribution function. For any given value of a, define $\overline{K}(a)$ by

$$\overline{K}(a) = \left[\frac{\exp(a)}{\delta}\right]^{\frac{1}{1-\alpha}}.$$

Then, conditional on $a_t = a$, the state space of capital can be restricted to $\mathcal{K}(a) = \left[0, \overline{K}(a)\right]$. The state space of the stochastic growth model is given by

$$\mathcal{S} = \{(K, a) : K \in \mathcal{K}(a), a \in \mathbb{R}\}.$$

The Bellman equation for the planner's problem can be written as

$$V(K, a) = \max_{K' \in \mathcal{K}(a)} \left\{ U\left[\exp(a) K^{\alpha} + (1-\delta) K - K'\right] + \beta \int V\left(K', a'\right) dF\left(a'|a\right) \right\}. \tag{9}$$

The solution of this problem includes a value function $V : \mathcal{S} \to \mathbb{R}$ and a policy function $g : \mathcal{S} \to \mathbb{R}$. The latter specifies the law of motion for capital. With logarithmic utility function and full depreciation, the policy function and the stationary distribution can be derived analytically. Specifically, the policy function for next-period capital (in logarithmic terms) is given by

$$k_{t+1} = g(k_t, a_t) \equiv \ln \alpha\beta + a_t + \alpha k_t. \tag{10}$$

The stationary distribution of (k, a) is a bivariate normal distribution with mean vector

$$\boldsymbol{\mu}' = \begin{bmatrix} \frac{\ln(\alpha\beta)}{1-\alpha} & 0 \end{bmatrix},$$

and variance-covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_k^2 & \sigma_{ka} \\ \sigma_{ka} & \sigma_a^2 \end{bmatrix},$$

where

$$\sigma_k^2 = \frac{(1+\alpha\beta)\sigma_a^2}{(1-\alpha^2)(1-\alpha\rho)},$$

$$\sigma_{ka} = \frac{\rho\sigma_a^2}{1-\alpha\rho}, \qquad \text{and} \qquad \sigma_a^2 = \frac{\sigma_{\varepsilon}^2}{1-\rho^2}.$$

Using these closed-form solutions, we can derive analytically the business cycle moments. These results are then used to assess the relative performance of six different discretization methods.

3.1 Discretizing the AR(1) Process

The first step in solving the Bellman equation is to devise an approximation for the integral in the objective function. This typically involves replacing the AR(1) process in (8) with a discrete state-space Markov chain. Formally, define an N-state Markov chain with state space $\mathcal{A} = \{\overline{a}_1, ..., \overline{a}_N\}$ and transition matrix $\Pi = [\pi_{i,j}]$. The Bellman equation can then be written as

$$\widetilde{V}(K, \overline{a}_i) = \max_{K' \in \mathcal{K}(\overline{a}_i)} \left\{ U\left[\exp(\overline{a}_i) K^{\alpha} + (1-\delta) K - K'\right] + \beta \sum_{j=1}^{N} \widetilde{V}\left(K', \overline{a}_j\right) \pi_{i,j} \right\}, \tag{11}$$

for every $\overline{a}_i$ in $\mathcal{A}$. The solution of this problem, $\widetilde{V}$, is an approximation of the *actual* value function. In the following section, six different methods for constructing the Markov chain are compared. These include the five described below and the Rouwenhorst method.

Tauchen (1986) method

Under this method, an evenly-spaced state space $\mathcal{A} = \{\overline{a}_1, ..., \overline{a}_N\}$ is used, with

$$\overline{a}_N = -\overline{a}_1 = \frac{M \sigma_{\varepsilon}}{\sqrt{1-\rho^2}}, \tag{12}$$

where M is a positive real number. The step between any two grid points is given by $h = (\overline{a}_N - \overline{a}_1)/(N-1)$. Let Φ be the probability distribution function for the standard normal distribution. For any $i = 1, ..., N$, the transition probabilities of the Markov chain are given by

$$\pi_{i,1} = \Phi\left(\frac{\overline{a}_1 - \rho \overline{a}_i + h/2}{\sigma_{\varepsilon}}\right),$$

$$\pi_{i,N} = 1 - \Phi\left(\frac{\overline{a}_N - \rho \overline{a}_i - h/2}{\sigma_{\varepsilon}}\right),$$

and

$$\pi_{i,j} = \Phi\left(\frac{\overline{a}_j - \rho \overline{a}_i + h/2}{\sigma_{\varepsilon}}\right) - \Phi\left(\frac{\overline{a}_j - \rho \overline{a}_i - h/2}{\sigma_{\varepsilon}}\right),$$

for $j = 2, ..., N-1$. Tauchen states that if the state space $\mathcal{A}$ is sufficiently fine, then the conditional distribution of the discrete process will converge to the conditional distribution function $F(a'|\overline{a}_i)$.

One drawback of this method is that its performance is strongly affected by the choice of M in (12). To the best of our knowledge, there is no established rule for determining this parameter. In Tauchen (1986), the author sets $M = 3$ but offers no justification for this choice. Flodén (2008) sets $M = 1.2\ln(N)$. Hence the width of the state space is increasing in the number of states. As explained in the results section, Flodén's choice of M is the main reason why he finds that the Tauchen (1986) method performs poorly in approximating highly persistent processes. In all the results reported below, M is calibrated to match the standard deviation of the original AR(1) process. This approach gives the method its best chance in approximating the AR(1) process. We choose to target σ_a instead of ρ because, relative to σ_a, the persistence parameter ρ is well approximated under this method for a range of values of M and degrees of persistence.

The Quadrature-Based Methods

This class of methods is based on the Gauss-Hermite quadrature. The general procedure is as follows. First, the elements of the state space $\mathcal{A}$ are determined by

$$\overline{a}_i = \sqrt{2}\sigma x_i, \qquad \text{for } i = 1, 2, ..., N,$$

where $\{x_i\}$ are the Gauss-Hermite nodes over $[-\infty, \infty]$. Let $\{\phi_j\}$ are the corresponding Gauss-Hermite weights. The elements in the transition matrix Π are then given by

$$\pi_{i,j} = \frac{f(\overline{a}_j|\overline{a}_i)}{f(\overline{a}_j|0)}\frac{\overline{w}_j}{s_i},$$

where $\overline{w}_j = \phi_j/\sqrt{\pi}$, the function $f(\overline{a}_j|\overline{a}_i)$ is the density function for a normal distribution with mean $\rho\overline{a}_i$ and variance σ^2, and

$$s_i = \sum_{n=1}^{N} \frac{f(\overline{a}_n|\overline{a}_i)}{f(\overline{a}_n|0)}\overline{w}_n.$$

The only difference between the original method considered in Tauchen and Hussey (1991) and the variations considered in Flodén (2008) is the choice of σ. In the original version, the standard deviation σ is taken to be σ_ε. In the first variation, the standard deviation of a_t is used instead, i.e., $\sigma = \sigma_a = \sigma_\varepsilon/\sqrt{1-\rho^2}$. In the second variation, σ is a weighted average of σ_a and σ_ε. In particular,

$\sigma = \omega\sigma_\varepsilon + (1-\omega)\,\sigma_a$, with $\omega = 0.5 + 0.25\rho$.

The Adda-Cooper Method

The first step of this method is to partition the real line into N intervals. These intervals are constructed so that the random variable a_t has an equal probability of falling into them. Formally, let $I_n = [x_n, x_{n+1}]$ be the nth interval with $x_1 = -\infty$ and $x_{N+1} = +\infty$. The cut-off points $\{x_n\}_{n=2}^N$ are obtained by solving the following system of equations:

$$\Phi\left(\frac{x_{n+1}}{\sigma_a}\right) - \Phi\left(\frac{x_n}{\sigma_a}\right) = \frac{1}{N}, \qquad \text{for } n = 1, 2, ..., N,$$

where Φ is the probability distribution function for the standard normal distribution. The nth element in the state space $\mathcal{A} = \{\overline{a}_1, ..., \overline{a}_N\}$ is then given by the mean value of the nth interval, i.e., $\overline{a}_n = E\left[a | a \in I_n\right]$. For any $i, j \in \{1, 2, ..., N\}$, the transition probability $\pi_{i,j}$ is defined as the probability of moving from interval I_i to interval I_j in one period. Formally, this is given by $\pi_{i,j} = \Pr\left[a' \in I_j | a \in I_i\right]$.

3.2 Experiments and Evaluation

The objective of this section is to evaluate the performance of different discretization methods. To achieve this, we focus on the business cycle moments generated by the stochastic growth model. The main criteria for evaluating the six discretization methods is the accuracy in approximating these moments.

Solution Method

The first step in computing the business cycle moments is to choose a specific form for the utility function and a set of values for the parameters $\{\alpha, \beta, \delta, \sigma_\varepsilon, \rho\}$. In the baseline model, the utility function is logarithmic and there is full depreciation. The full depreciation assumption is later relaxed in Section 3.4. The other parameter values are chosen to be the same as in King and Rebelo (1999): $\alpha = 0.33$, $\beta = 0.984$, $\sigma_\varepsilon = 0.0072$ and $\rho = 0.979$.

The next step is to discretize the state space $\mathcal{S}$. First, the AR(1) process in (8) is approximated using the methods mentioned above. The resulting N-state Markov chain is characterized by a state space $\mathcal{A} = \{\overline{a}_1, ..., \overline{a}_N\}$ and a transition matrix $\Pi = [\pi_{i,j}]$. Second, the continuous state space

for capital is replaced by an evenly-spaced grid. Define the variable $k \equiv \ln K$. The set of grid points for k is represented by $\mathcal{K} = \left\{\overline{k}_1, ..., \overline{k}_M\right\}$. The discretized state space can be expressed by

$$\widehat{\mathcal{S}} = \left\{\left(\overline{k}_m, \overline{a}_n\right) : \overline{k}_m \in \mathcal{K}, \overline{a}_n \in \mathcal{A}\right\}. \tag{13}$$

In the baseline case, the number of states in the Markov chain is set to five and the number of grid points for capital is 1000. After the discrete state space $\widehat{\mathcal{S}}$ is formed, the value function and the associated policy function are solved using the value-function iteration method described in Tauchen (1990) and Burnside (1999). The outcome of this procedure includes a set of $N \times M$ values of the policy function evaluated on $\widehat{\mathcal{S}}$. This set of values is represented by $\left\{\widehat{g}\left(\overline{k}_m, \overline{a}_n\right)\right\}$.

The final task is to compute the stationary distribution of the state variables (k, a). The first step to achieve this is to construct the transition matrix for these variables. Under the discrete state-space method, the probability of moving from state $\left(\overline{k}_m, \overline{a}_n\right)$ in $\widehat{\mathcal{S}}$ to state $\left(\overline{k}_l, \overline{a}_j\right)$ in $\widehat{\mathcal{S}}$ in one period is specified by

$$\Pr\left[\left(k', a'\right) = \left(\overline{k}_l, \overline{a}_j\right) | \left(k, a\right) = \left(\overline{k}_m, \overline{a}_n\right)\right] = \begin{cases} \pi_{n,j}, & \text{if } \overline{k}_l = \widehat{g}\left(\overline{k}_m, \overline{a}_n\right) \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

The resulting NM-by-NM transition matrix is denoted P. Let $\widehat{\boldsymbol{\pi}} = (\widehat{\pi}_1, ..., \widehat{\pi}_{NM})$ be the stationary distribution associated with P. Formally, this is defined by

$$\widehat{\boldsymbol{\pi}} P = \widehat{\boldsymbol{\pi}}.$$

In principle, $\widehat{\boldsymbol{\pi}}$ can be obtained as the eigenvector of P corresponding to eigenvalue 1, with the normalization $\sum_{i=1}^{NM} \widehat{\pi}_i = 1$. This method, however, is not practical when the number of grid points in the state space is large. In the following experiments, an approximation for the stationary distribution is obtained by iterating the equation

$$\widetilde{\boldsymbol{\pi}}^l P = \widetilde{\boldsymbol{\pi}}^{l+1}. \tag{15}$$

A good approximation for $\widehat{\boldsymbol{\pi}}$ can be obtained when l is sufficiently large. Given the approximate stationary distribution $\widetilde{\boldsymbol{\pi}}^l$ and the policy function $\widehat{g}$, the business cycle moments of interest can be computed. This process of computing the business cycle moments is referred to below as the

baseline approach.

An alternative route to compute the business cycle moments is to use Monte Carlo simulations. The standard procedure involves the following steps. Draw a sequence of pseudorandom numbers of length $T = 1,010,000$ for the disturbance term ε.[3] Construct the random variable a_t using the *actual* AR(1) process given in (8). The resulting sequence is denoted $\{\widetilde{a}_t\}_{t=0}^{T}$. Construct a sequence of capital $\left\{\widetilde{k}_t\right\}_{t=0}^{T}$ according to

$$\widetilde{k}_{t+1} = \widehat{g}\left(\widetilde{k}_t, \widetilde{a}_t\right), \quad \text{with } \widetilde{k}_0 \text{ given.}$$

In general, the generated values of $\widetilde{k}_t$ and $\widetilde{a}_t$ will not coincide with the grid points in $\widehat{\mathcal{S}}$. In this case, linear interpolation is used to compute the value of $\widehat{g}\left(\widetilde{k}_t, \widetilde{a}_t\right)$. To ensure that the generated values of $\widetilde{k}_t$ and $\widetilde{a}_t$ are drawn from the stationary distribution, the first $10,000$ observations in either sequence are deleted. Next, compute the sample variances and covariance as follows,

$$s_{xx} = \frac{1}{T}\sum_{t=0}^{T}\widetilde{x}_t^2 - \left(\frac{1}{T}\sum_{t=0}^{T}\widetilde{x}_t\right)^2, \quad \text{for } x = k, a,$$

$$s_{ak} = \frac{1}{T}\sum_{t=0}^{T}\widetilde{a}_t\widetilde{k}_t - \left(\frac{1}{T}\sum_{t=0}^{T}\widetilde{a}_t\right)\left(\frac{1}{T}\sum_{t=0}^{T}\widetilde{k}_t\right).$$

The sample moments, (s_{kk}, s_{aa}, s_{ak}), then serve as an estimate for the variance-covariance matrix of (k, a).[4] The moments for the other variables are obtained in the same fashion.

Baseline Results

Table 2 presents the baseline results. The six discretization methods are compared on three grounds: (i) the accuracy in approximating the AR(1) process, (ii) the precision in approximating the stationary distribution of the state variables, and (iii) the accuracy in approximating the business cycle moments. The table gives the ratio of the statistics computed following the above procedure to their *true* values. The true values are derived using the closed-form solutions as mentioned in Section 3.1.

[3] The generated sequence is first adjusted to remove any first-order serial correlation in it that may be introduced by the pseudorandom number generator. The resulting sequence is then transformed to one with mean zero and variance σ^2.

[4] Santos and Peralta-Alva (2005) suggest that generating one single sample path of extremely long length is more appropriate than having a large number of sample paths with much shorter length.

Panel (A) of Table 2 shows the performance of these methods in approximating the AR(1) process.[5] As explained in Section 2.3, the transition matrix in the Rouwenhorst method (R) can be calibrated to match exactly the persistence parameter, the standard deviation of ε and the standard deviation of a. Similarly the parameter M in the Tauchen (1986) method is calibrated to match exactly the standard deviation of a. The required value is $M = 1.6425$. With this choice of M, the Tauchen (1986) method has a relative error of about one percent in approximating the persistence parameter. These results are in stark contrast to those reported in Flodén (2008) Table 2. In this study, the author chooses $M = 1.9313$ when $N = 5$. As a result, the Tauchen (1986) method generates a 12 percent error in approximating σ_a and a 1.5 percent error in approximating ρ. This illustrates that the performance of this method is very sensitive to the choice of M.

Next, we consider the accuracies of these methods in approximating the stationary distribution of the state variables. Panel (B) of Table 2 shows the performance of these methods in approximating the standard deviation of k and the covariance between a and k. In general, a discretization method that generates an accurate approximation for σ_a also has high precision in approximating these two moments. Among these six methods, the Rouwenhorst method has the highest accuracy in approximating these two moments. The relative errors for the two are about 0.14 percent. The Tauchen (1986) method is the second best. These two methods outperform the others by a significant margin.

Next, we compare the performance of these methods in approximating the business cycle moments. In particular, we focus on the standard deviation of output, consumption and investment (in logarithmic terms) and the first-order autocorrelation of output (in logarithmic terms).[6] The results are shown in panel (C) of Table 2. Again the Rouwenhorst method has the best overall performance in terms of approximating all these moments. However, with $M = 1.6425$, the Tauchen (1986) method can produce highly accurate approximations that are comparable to those generated by the Rouwenhorst method. As mentioned above, the performance of this method is very sensitive to the choice of M. If we set $M = 1.9313$ as in Flodén (2008), then the Tauchen (1986) method would generate a 12-percent error in approximating the standard deviations.[7]

[5] The relative errors reported in panel (A) are directly comparable to those reported in Flodén (2008) Table 2 for $n = 5$ and $\rho = 0.98$ except for two differences. The first difference is that Flodén did not consider the Rouwenhorst method. The second difference lies in the choice of M in the Tauchen (1986) method. This point is elaborated below.

[6] The first-order autocorrelation of consumption and investment (in logarithmic terms), and the cross-correlation between output and these variables are not shown in the paper. These results are available from the authors upon request.

[7] These results are not shown in here but are available from the authors upon request.

Finally, two things can be observed when comparing across all three panels. First, the relative errors in approximating σ_a are very similar to those in approximating the standard deviation of capital, output, consumption and investment. Second, the relative errors in approximating ρ are close to those in approximating the first-order autocorrelation for output. These results suggest that a good approximation for the moments of the AR(1) process is important in obtaining an accurate approximation for the business cycle moments.

Error Analysis

The relative errors reported in Table 2 have a number of sources. For the purpose of this discussion, we classify these into two groups. The first group of errors arises when solving the Bellman equation in (9). This includes the errors that arise when we restrict the choice of next-period capital to a discrete set of values, and the truncation errors that emerge when we approximate the fixed point of the Bellman equation using a finite number of iterations. The second group of errors occurs during the computation of the stationary distribution of the state variables. First, the transition matrix P, constructed using the discrete Markov chain and the computed policy function, is an approximation of the actual transition function. Second, truncation errors arise when we approximate the stationary distribution using a finite number of iterations. The second group of errors would not occur if Monte Carlo simulations are used to generate the business cycle moments. In this case, however, a new source of error arises when we estimate the actual moments by a finite sample.

Using the actual policy function, it is possible to disentangle the two groups of errors. Consider the following experiment. Construct a discrete state space $\widehat{\mathcal{S}}$ as in (13) using one of the six discretization methods. Construct the transition matrix P as in (14) *but replace the computed policy function* $\widehat{g}(k,a)$ *with the actual one in (10).* Iterate equation (15) successively to obtain an approximation for the stationary distribution of the state variables. Finally, use the approximate stationary distribution and the *actual* policy function $g(k,a)$ to compute the business cycle moments. By replacing $\widehat{g}(k,a)$ with the actual policy function, this procedure effectively removes all the errors involved in solving the Bellman equation. The remaining errors are thus due to the approximation of the stationary distribution of the state variables. The results of this procedure are reported in panel (B) of Table 3. To facilitate comparison, the baseline results are shown in panel (A) of the same table.

It is immediate to see that the figures in the two panels are almost identical. Replacing the computed policy function with the actual one does not affect the approximation of the technology shock process. As a result, the approximated values for ρ, σ_ε and σ_a are identical in the two sets of results. As for the standard deviations of the endogenous variables, only minor discrepancies are observed in the two panels. In other words, even though we have removed all the errors in computing the policy function, the baseline results remain largely unchanged. This has two implications. First, this implies that almost all the relative errors in the baseline case are due to the approximation of the stationary distribution $\widehat{\boldsymbol{\pi}}$. Second, this means the choice of discretization method has only a relatively minor impact on the solution of the Bellman equation. In sum, this experiment illustrates that the choice of discretization method matters because it would significantly affect the approximation of the stationary distribution.

The same conclusion can be drawn from another experiment. Suppose now the business cycle moments are computed using Monte Carlo simulations. More specifically, after solving the dynamic programming problem in (9), the model is simulated using the actual AR(1) process and the computed policy function $\widehat{g}(k, a)$. Under this procedure, the choice of discretization method only affects the simulated moments through the computed policy function. Table 4 presents the relative errors obtained under this procedure alongside with the baseline results. The two methods of generating business cycle moments have produced very different results. When the model is simulated using the actual AR(1) process, all six discretization methods generate almost identical results. This again implies that the differences in the baseline results across the six discretization methods are due to the approximation of the stationary distribution $\widehat{\boldsymbol{\pi}}$.

Finally, when comparing between the two panels of Table 4, one can see that the baseline approach, when combined with the Rouwenhorst method, can generate estimated moments that are as accurate as those produced by the simulation method with one million draws.

3.3 Robustness Check

In this section, it is shown that the relative performance of the six discretization methods are robust to changes in (i) the number of points in the discrete state space N, (ii) the persistence parameter ρ, and (iii) the standard deviation of the white noise process σ_ε.

Changing the Number of States

Table 5 compares the performance of the six methods under different choices of N. Intuitively, increasing the number of states in the Markov chain should improve the performance of the discretization methods. This is true for the Rouwenhorst method, the Tauchen (1986) method, the original Tauchen-Hussey method, the F-2 method, and the Adda-Cooper method. However, this is not true for the F-1 method.

The results in Table 5 show that the superior performance of the Rouwenhorst method is robust even when there are only two states in the discrete Markov chain. The relative errors in approximating the standard deviations of output, capital, consumption and investment are similar in all three cases. In particular, increasing the number of states from five to ten increases the precision only marginally. The original Tauchen-Hussey method has the lowest precision among the six in all three cases. Even when the number of states is increased to ten, the Tauchen-Hussey method can only replicate 57 percent of the actual value of σ_y. The performance of this method is much better when approximating ρ_y but the precision is still the lowest among the six.

Next, we consider the performance of the Tauchen (1986) method. For each value of N, we adjust the parameter M so as to match the actual value of σ_a. The required values for $N = 2$ and $N = 10$ are 1.0000 and 1.9847, respectively. In other words, in order to match the standard deviation σ_a, a wider state space (i.e., a larger value of M) is needed when the number of states increases. When M is adjusted in this fashion, increasing the number of states in the Tauchen (1986) method increases the precision only marginally. For instance, the relative error in approximating σ_y reduces from 0.35 percent to 0.22 percent when N increases from five to ten.

Changing the Persistence Parameter

Table 6 compares the performance of the six methods under different values of ρ. The superior performance of the Rouwenhorst method is robust to changes in this parameter. In particular, increasing the persistence of the AR(1) process from 0.5 to 0.979 has very little impact on its precision. This shows that the Rouwenhorst method is a reliable technique for approximating stationary AR(1) process in general.

Similar to the results in Table 5, the parameter M in the Tauchen (1986) method is adjusted in each case so as to match the actual value of σ_a. The resulting values are shown in Table 6. In

general, a wider state space (i.e., a larger value of M) is needed for less persistent processes. When M is calibrated to match σ_a, the Tauchen (1986) method has better performance in approximating highly persistent processes. For instance, when $\rho = 0.5$ the relative errors in approximating σ_{ka} and ρ_y are 4.66 percent and 1.84 percent, respectively. These become 1.34 percent and 0.36 percent, respectively, when $\rho = 0.979$. The precision of this method in approximating the standard deviations is not sensitive to changes in ρ.

The performance of the three quadrature-based methods is very sensitive to the value of ρ. Similar to Flodén (2008), our results show that the quadrature-based methods work best in approximating AR(1) processes with low persistence. But unlike Flodén (2008) which only focuses on the parameters of the AR(1) process, the current study also considers the impact of these methods on the moments of the endogenous variables. When ρ equals to 0.5 or 0.6, the original Tauchen-Hussey method and its two variations can generate highly accurate approximations that are comparable to those generated by the Rouwenhorst method. The relative errors for the business cycle moments are all less than one percent. Within this range of ρ, the three quadrature-based methods are more accurate than the Tauchen (1986) method, especially in approximating σ_{ka} and ρ_y. However, the accuracies of the Tauchen-Hussey method and the F-2 method deteriorate quickly when the persistence parameter approaches one. For instance, the Tauchen-Hussey method has a relative error of 25 percent in approximating σ_y when ρ equals to 0.9 and an error of 61 percent when ρ is 0.979.

Finally, it is worth mentioning that the results of the two experiments conducted in the error analysis section are also robust to different values of the persistence parameter. These results are summarized as follow.[8] First, the figures reported in Table 6 are largely unaffected when we replace the computed policy function with the actual one. Second, when the business cycle moments are computed using Monte Carlo simulations, all six discretization methods generate very similar results.

Changing the Standard Deviation of the White Noise Process

The performance of the six methods under different values of σ_ε are shown in Table 7. In terms of approximating the AR(1) process, increasing the value of σ_ε from 0.001 to 0.1 does not seem to affect the performance of these methods. In terms of approximating the standard deviations of the endogenous variables and the covariance between a and k, the accuracies of the F-2 method and

[8] The numerical results are not shown in the paper but are available from the authors upon request.

the Adda-Cooper method improve when the AR(1) process is less volatile. The opposite is true for the Rouwenhorst method and the Tauchen (1986) method. The variations in the relative errors, however, are not significant. More specifically, increasing σ_ε from 0.001 to 0.1 changes the relative errors by less than two percentage points in most cases. Unlike the other methods, the performance of the F-1 method is more sensitive to the value of σ_ε. For instance, when σ_ε equals to 0.001 the relative errors in approximating σ_k and σ_{ka} are 0.5 percent and five percent, respectively. These become 1.6 percent and 2.3 percent, respectively, when σ_ε is 0.1. Finally, the precision of all six methods in approximating ρ_y is not sensitive to changes in the value of σ_ε.

3.4 Relaxing the Assumption of Full Depreciation

This section evaluates the performance of the six discretization methods in solving the stochastic growth model when the full depreciation assumption is relaxed. The rate of depreciation is now taken to be 2.5 percent, which is the same as in King and Rebelo (1999). All other parameters remain the same as in the baseline case. The same evaluation process is performed as in Section 3.2. For each of the six discretization methods, we compute the business cycle moments using the baseline approach and the Monte Carlo simulation method. Without full depreciation, however, a closed-form solution for the policy function is not available and the actual values of the business cycle moments are unknown. Thus we first derive a highly accurate approximation for the actual moments which is then used as our yardstick for comparison. To achieve this, we first construct an extremely fine state space with 2000 grid points for capital and 400 states in the Markov chain constructed by the Rouwenhorst method. The business cycle moments are then computed using the baseline approach described earlier. The rationale for this procedure is as follows. As explained in the error analysis section, the baseline approach involves two groups of errors: (i) errors that arise when solving the Bellman equation, and (ii) errors that arise when computing the stationary distribution. When the number of grid points in the discrete state space is sufficiently large, the value function iteration method is able to yield highly accurate solutions for the Bellman equation. Thus, by adopting an extremely fine state space, the above procedure should render the first group of errors very small. As for the second group of errors, our baseline results for the full depreciation case show that combining the Rouwenhorst method and the baseline approach can yield a highly accurate approximation for the stationary distribution. As a robustness check on this procedure, we double the size of the state space and find that it has no effect on the computed statistics. The

business cycle moments obtained under this procedure are referred to below as the true solutions.

Panel (A) of Table 8 shows the results obtained under the baseline approach for three different values of N and Panel (B) reports the simulation results. First, note that the superior performance of the Rouwenhorst discretization method is robust to relaxing the full depreciation assumption. Second, the overall performance of the other methods deteriorates significantly when δ is less than one. This is particularly true for the estimates of σ_{ka} and σ_i. For example, consider the Tauchen (1986) method which has the second highest precision in the full depreciation case. With only five states in the Markov chain and full depreciation, this method generates a relative error of 1.3 percent in approximating σ_{ka} and an error of about 0.5 percent in approximating σ_i (see Table 2). These become 6.4 percent and 4.0 percent, respectively, when δ equals 0.025. In contrast, relaxing the full depreciation assumption has only a negligible effect on the estimates of ρ_y.

Third, similar to the full depreciation case, increasing the number of states in the Markov chain usually improves the accuracy of the approximations. However, the drastic differences in the performance of the six methods remain even when N is large. For the Rouwenhorst method, a five-fold increase in the number of states only marginally affects the precision of the results. However, unlike the full depreciation case, increasing the number of states does not always improve the precision. In particular, the relatively large error in approximating σ_i remains even when there are 25 states. For the original Tauchen-Hussey method, its performance improves significantly when the fineness of the state space increases. However, even when there are 25 states, this method can only replicate 67 percent of the true value of σ_{ka} and 83 percent of the true value of σ_y. The overall performance of the F-1 method is also rather disappointing in this case. A five-fold increase in the number of states does not seem to have a significant impact on its precision. On the other hand, when N is large the Tauchen (1986) method and the F-2 method are able to yield highly accurate approximations that are comparable to those generated by the Rouwenhorst method. As for the Adda-Cooper method, relatively large errors remain even when there are 25 states. For instance, the relative errors in approximating σ_{ka} and σ_i are about five percent.

Unlike the full depreciation case, the six discretization methods do not generate near identical results under the Monte Carlo simulation approach. This can be seen by comparing the columns in Panel (B) of Table 8. Thus the choice of discretization method matters even when the business cycle moments are computed using Monte Carlo simulations. This is due to the following reason. In the absence of full depreciation, the policy function for next-period capital (in logarithms) is no longer

a linear function. Consequently, additional approximation errors arise when we compute $g(k_t, a_t)$ for values of k_t and a_t that are outside the discrete state space. The size of these errors depends on the location of the grid points and hence the choice of the discretization method. As the number of states in the Markov chain increases, the state space becomes finer and the errors associated with the interpolation procedure falls. For this reason, a five-fold increase in N significantly reduces the relative errors of the discretization methods. Under the Monte Carlo simulation approach, no single method dominates all others in all three choices of N. When there are five states in the Markov chain, the Rouwenhorst method has the best overall performance within the group. When there are 25 states, the Tauchen (1986) method, the original Tauchen-Hussey method and the F-2 method all perform equally well as the Rouwenhorst method.

Finally, it is worth mentioning that even in the absence of full depreciation, moments computed using the Rouwenhorst method and the baseline approach are as accurate as those obtained from Monte Carlo simulations with one million draws.

4 Conclusions

This paper re-examines the Rouwenhorst method of constructing a discrete-valued Markov chain to approximate a given first-order autoregressive process. Under this method, the constructed Markov chain can be calibrated to match the conditional and unconditional mean, the conditional and unconditional variance and the first-order autocorrelation of any stationary AR(1) process. Because of this distinctive feature, the Rouwenhorst method is more reliable than the Tauchen (1986) method and the Tauchen-Hussey method to approximate highly persistent processes. In this paper, a new and simpler procedure for generating the transition matrix in the Rouwenhorst method is developed and the first formal proof for all the important properties of the constructed Markov chain is provided.

In the quantitative analysis, the Rouwenhorst method is compared to five other discretization methods. These methods are evaluated based on their performance in approximating the business cycle moments generated by the standard neoclassical growth model without leisure. Two approaches to generate these moments are considered. In the baseline approach, an approximation for the stationary distribution of the state variables is first computed. In the second approach, the moments of interest are generated using Monte Carlo simulations. Our quantitative analysis shows

that, under both approaches, the choice of approximation method can have a large impact on the accuracy of the solutions. Under the baseline approach, an accurate approximation of the moments of the AR(1) process is important in accurately approximating the business cycle moments. The Rouwenhorst method has the best performance in this regard. Its superior performance is robust under a wide range of parameter values. Under the second approach, no single method dominates all others in all cases. When a realistic value of the depreciation rate is used, the Rouwenhorst method again has the best overall performance when there are only five states in the Markov chain. However, when the fineness of the state space increases, the Tauchen (1986) method, the original Tauchen-Hussey method and the F-2 method all perform equally well as the Rouwenhorst method. When comparing between the two approaches, it is found that combining the Rouwenhorst method and the baseline approach can yield highly accurate approximations that are similar to those obtained from Monte Carlo simulations with one million draws.

In this paper, we use a standard representative-agent model as our test model. We believe that similar results can be obtained in heterogeneous-agent economies. However, we leave a detailed exploration of these models for future research.

Table 2 Baseline Results

(A) Approximating the AR(1) process

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
ρ	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000
σ_ε	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000
σ_a	1.0000	0.4006	1.0215	0.7742	0.9471	1.0000

(B) Approximating the Variance-Covariance Matrix for State Variables

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
σ_k	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
σ_{ka}	1.0134	0.1401	1.0818	0.6071	0.8464	0.9986

(C) Approximating Business Cycle Moments

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
σ_y	1.0035	0.3880	1.0310	0.7763	0.9338	0.9995
σ_c	1.0026	0.3879	1.0295	0.7776	0.9343	1.0000
σ_i	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
ρ_y	1.0036	0.9538	1.0107	1.0063	0.9807	1.0000

T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, $\sigma_\varepsilon = 0.0072$, $\rho = 0.979$, $N = 5$, $M = 1.6425$.

Table 3 Error Analysis

(A) Using Computed Policy Function (Baseline case)

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
ρ	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000
σ_ε	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000
σ_a	1.0000	0.4006	1.0215	0.7742	0.9471	1.0000
σ_k	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
σ_{ka}	1.0134	0.1401	1.0818	0.6071	0.8464	0.9986
σ_y	1.0035	0.3880	1.0310	0.7763	0.9338	0.9995
σ_c	1.0026	0.3879	1.0295	0.7776	0.9343	1.0000
σ_i	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
ρ_y	1.0036	0.9538	1.0107	1.0063	0.9807	1.0000

(B) Using Actual Policy Function

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
ρ	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000
σ_ε	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000
σ_a	1.0000	0.4006	1.0212	0.7742	0.9471	1.0000
σ_k	1.0026	0.3880	1.0292	0.7777	0.9343	1.0000
σ_{ka}	1.0107	0.1400	1.0762	0.6104	0.8475	1.0000
σ_y	1.0026	0.3879	1.0292	0.7777	0.9343	1.0000
σ_c	1.0026	0.3879	1.0292	0.7777	0.9343	1.0000
σ_i	1.0026	0.3880	1.0292	0.7777	0.9343	1.0000
ρ_y	1.0036	0.9537	1.0107	1.0063	0.9807	1.0000

T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, σ_ε= 0.0072, $\rho = 0.979$, $N = 5$, $M = 1.6425$.

Table 4 Baseline Approach vs. Monte Carlo Simulations

(A) Baseline case

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
ρ	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000
σ_ε	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000
σ_a	1.0000	0.4006	1.0215	0.7742	0.9471	1.0000
σ_k	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
σ_{ka}	1.0134	0.1401	1.0818	0.6071	0.8464	0.9986
σ_y	1.0035	0.3880	1.0310	0.7763	0.9338	0.9995
σ_c	1.0026	0.3879	1.0295	0.7776	0.9343	1.0000
σ_i	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
ρ_y	1.0036	0.9538	1.0107	1.0063	0.9807	1.0000

(B) Monte Carlo Simulations

	Generated Values Relative to True Values					
	Tauchen	T-H	F-1	F-2	A-C	R
ρ	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
σ_ε	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
σ_a	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955
σ_k	0.9955	0.9955	0.9955	0.9955	0.9957	0.9955
σ_{ka}	0.9908	0.9908	0.9908	0.9908	0.9910	0.9908
σ_y	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955
σ_c	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955
σ_i	0.9955	0.9955	0.9955	0.9955	0.9957	0.9955
ρ_y	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, σ_ε= 0.0072, $\rho = 0.979$, $N = 5$, $M = 1.6425$.

Table 5 Changing the Number of States in the Markov Chain

	$N = 2$						$N = 5$ (Baseline)						$N = 10$					
	Generated Values Relative to True Values						Generated Values Relative to True Values						Generated Values Relative to True Values					
	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R
ρ	1.0214	0.7688	1.0215	1.0206	0.8879	1.0000	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000	0.9989	0.9867	1.0214	1.0006	1.0038	1.0000
σ_ε	0.0087	0.6584	0.0000	0.0805	1.9346	1.0000	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000	1.1318	0.9493	0.0225	0.8886	1.2781	1.0000
σ_a	1.0000	0.2039	1.0000	0.4071	0.7979	1.0000	1.0000	0.4006	1.0215	0.7742	0.9471	1.0000	1.0000	0.5860	1.0076	0.9558	0.9793	1.0000
σ_k	1.0040	0.1844	1.0040	0.4095	0.7718	0.9966	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986	0.9978	0.5794	1.0122	0.9583	0.9753	0.9988
σ_{ka}	1.0281	0.0283	1.0281	0.1705	0.5400	0.9966	1.0134	0.1401	1.0818	0.6071	0.8464	0.9986	0.9910	0.3260	1.0444	0.9190	0.9386	0.9988
σ_y	1.0065	0.1870	1.0065	0.4099	0.7682	0.9989	1.0035	0.3880	1.0310	0.7763	0.9338	0.9995	0.9978	0.5788	1.0144	0.9573	0.9743	0.9996
σ_c	1.0078	0.1882	1.0078	0.4101	0.7664	1.0000	1.0026	0.3879	1.0295	0.7776	0.9343	1.0000	0.9978	0.5784	1.0154	0.9569	0.9739	1.0000
σ_i	1.0040	0.1844	1.0040	0.4095	0.7718	0.9966	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986	0.9978	0.5794	1.0122	0.9583	0.9753	0.9988
ρ_y	1.0107	0.8752	1.0107	1.0103	0.9422	1.0000	1.0036	0.9538	1.0107	1.0063	0.9807	1.0000	0.9969	0.9817	1.0107	1.0015	0.9922	1.0000

Tau stands for the Tauchen (1986) method; T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, $\sigma_\varepsilon = 0.0072$, $\rho = 0.979$.

* For the Tauchen (1986) method, $M = 1.0000$ when $N = 2$, $M = 1.6425$ when $N = 5$ and $M = 1.9847$ when $N = 10$.

Table 6 Changing the Persistence Parameter

	$\rho = 0.5$						$\rho = 0.6$						$\rho = 0.7$					
	Generated Values Relative to True Values						Generated Values Relative to True Values						Generated Values Relative to True Values					
	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R
ρ	0.9680	0.9997	1.0007	1.0000	0.9310	1.0000	0.9725	0.9986	1.0039	0.9999	0.9471	1.0000	0.9774	0.9953	1.0174	0.9997	0.9665	1.0000
σ_ε	1.0129	0.9994	0.9998	0.9999	0.9737	1.0000	1.0207	0.9972	0.9977	0.9993	0.9888	1.0000	1.0331	0.9905	0.9819	0.9969	1.0112	1.0000
σ_a	1.0000	0.9990	0.9999	0.9999	0.9471	1.0000	1.0000	0.9950	0.9996	0.9993	0.9471	1.0000	1.0000	0.9793	0.9982	0.9963	0.9471	1.0000
σ_k	0.9941	1.0025	0.9973	1.0048	0.9364	1.0010	0.9934	0.9973	0.9988	1.0017	0.9345	0.9991	0.9934	0.9788	1.0085	0.9969	0.9243	0.9983
σ_{ka}	0.9534	0.9995	0.9971	1.0035	0.8050	1.0011	0.9537	0.9873	1.0021	0.9998	0.8072	0.9981	0.9552	0.9456	1.0238	0.9925	0.8050	0.9985
σ_y	0.9935	0.9996	0.9993	1.0009	0.9336	1.0002	0.9924	0.9948	1.0000	0.9997	0.9316	0.9996	0.9918	0.9769	1.0039	0.9963	0.9282	0.9996
σ_c	0.9934	0.9983	1.0004	0.9992	0.9324	1.0001	0.9922	0.9938	1.0006	0.9989	0.9304	1.0000	0.9911	0.9761	1.0018	0.9962	0.9302	1.0003
σ_i	0.9941	1.0025	0.9973	1.0048	0.9364	1.0010	0.9934	0.9973	0.9988	1.0017	0.9345	0.9991	0.9934	0.9788	1.0085	0.9969	0.9243	0.9983
ρ_y	0.9816	1.0002	0.9997	1.0009	0.9588	1.0001	0.9818	0.9985	1.0014	1.0003	0.9599	0.9999	0.9825	0.9943	1.0088	0.9997	0.9613	0.9998

Tau stands for the Tauchen (1986) method; T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, $\sigma_\varepsilon = 0.0072$, $N = 5$.

* For the Tauchen (1986) method, $M = 1.9241$ when $\rho = 0.5$, $M = 1.9128$ when $\rho = 0.6$ and $M = 1.8917$ when $\rho = 0.7$.

Table 6 (continued) Changing the Persistence Parameter

	$\rho = 0.9$						$\rho = 0.95$						$\rho = 0.979$ (baseline)					
	Generated Values Relative to True Values						Generated Values Relative to True Values						Generated Values Relative to True Values					
	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R
ρ	0.9884	0.9689	1.0911	0.9986	1.0060	1.0000	0.9981	0.9550	1.0524	1.0025	1.0067	1.0000	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000
σ_ε	1.1027	0.9379	0.4217	0.9379	1.1403	1.0000	1.0964	0.9101	0.0586	0.8142	1.2822	1.0000	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000
σ_a	1.0000	0.7701	1.0008	0.9347	0.9471	1.0000	1.0000	0.5904	1.0135	0.8639	0.9471	1.0000	1.0000	0.4006	1.0215	0.7742	0.9471	1.0000
σ_k	0.9920	0.7564	1.0259	0.9337	0.9258	1.0024	0.9957	0.5753	1.0326	0.8660	0.9346	0.9948	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
σ_{ka}	0.9642	0.5453	1.1319	0.8716	0.8248	1.0024	0.9836	0.3108	1.1082	0.7530	0.8398	0.9948	1.0134	0.1401	1.0818	0.6071	0.8464	0.9986
σ_y	0.9918	0.7558	1.0297	0.9341	0.9291	1.0008	0.9960	0.5750	1.0324	0.8658	0.9328	0.9983	1.0035	0.3880	1.0310	0.7763	0.9338	0.9995
σ_c	0.9917	0.7556	1.0316	0.9344	0.9307	1.0000	0.9962	0.5748	1.0323	0.8657	0.9319	1.0000	1.0026	0.3879	1.0295	0.7776	0.9343	1.0000
σ_i	0.9920	0.7564	1.0259	0.9337	0.9258	1.0024	0.9957	0.5753	1.0326	0.8660	0.9346	0.9948	1.0053	0.3882	1.0342	0.7734	0.9330	0.9986
ρ_y	0.9870	0.9702	1.0454	0.9996	0.9718	1.0000	0.9944	0.9602	1.0260	1.0030	0.9765	0.9999	1.0036	0.9538	1.0107	1.0063	0.9807	1.0000

Tau stands for the Tauchen (1986) method; T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, $\sigma_\varepsilon = 0.0072$, $N = 5$.

* For the Tauchen (1986) method, $M = 1.7683$ when $\rho = 0.9$, $M = 1.6963$ when $\rho = 0.95$, $M = 1.6425$ when $\rho = 0.979$.

Table 7 Changing the Standard Deviation of the White Noise Process

	σ_ε= 0.001						σ_ε= 0.01						σ_ε= 0.1					
	Generated Values Relative to True Values						Generated Values Relative to True Values						Generated Values Relative to True Values					
	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R
ρ	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000
σ_ε	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000
σ_a	1.0000	0.4006	1.0213	0.7742	0.9471	1.0000	1.0000	0.4006	1.0216	0.7742	0.9471	1.0000	1.0000	0.4006	1.0212	0.7742	0.9471	1.0000
σ_k	1.0265	0.4166	1.0052	0.7911	0.9546	0.9910	1.0041	0.3864	1.0318	0.7798	0.9328	1.0022	1.0026	0.3879	0.9838	0.7775	0.9342	1.0000
σ_{ka}	1.0342	0.1497	1.0503	0.6203	0.8651	0.9905	1.0122	0.1395	1.0794	0.6121	0.8462	1.0022	1.0107	0.1400	1.0236	0.6103	0.8475	1.0000
σ_y	1.0103	0.3965	1.0211	0.7819	0.9405	0.9969	1.0031	0.3875	1.0303	0.7784	0.9338	1.0007	1.0026	0.3879	1.0130	0.7776	0.9342	1.0000
σ_c	1.0028	0.3875	1.0290	0.7778	0.9340	1.0000	1.0026	0.3880	1.0296	0.7777	0.9343	1.0000	1.0026	0.3879	1.0297	0.7777	0.9343	1.0000
σ_i	1.0265	0.4166	1.0052	0.7911	0.9546	0.9910	1.0041	0.3864	1.0318	0.7798	0.9328	1.0022	1.0026	0.3879	0.9838	0.7775	0.9342	1.0000
ρ_y	1.0037	0.9554	1.0107	1.0063	0.9810	0.9999	1.0036	0.9536	1.0107	1.0063	0.9807	1.0000	1.0036	0.9537	1.0107	1.0063	0.9807	1.0000

Tau stands for the Tauchen (1986) method; T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 1$, $\alpha = 0.33$, $\beta = 0.984$, ρ= 0.979, $N = 5$.

* For the Tauchen (1986) method, $M = 1.6425$ in all three cases.

Table 8 Results obtained when $\delta = 0.025$.

(A) Baseline Approach

	$N = 5$						$N = 10$						$N = 25$					
	Generated Values Relative to True Values						Generated Values Relative to True Values						Generated Values Relative to True Values					
	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R
ρ	1.0097	0.9453	1.0215	1.0096	0.9993	1.0000	0.9989	0.9867	1.0214	1.0006	1.0038	1.0000	0.9997	0.9980	1.0176	1.0000	1.0012	1.0000
σ_ε	0.8167	0.8905	0.0002	0.5019	1.5599	1.0000	1.1318	0.9493	0.0225	0.8886	1.2781	1.0000	1.0389	0.9877	0.4192	0.9994	1.0958	1.0000
σ_a	1.0000	0.4006	1.0212	0.7742	0.9471	1.0000	1.0000	0.5860	1.0076	0.9558	0.9793	1.0000	1.0000	0.8481	0.9998	0.9996	0.9937	1.0000
σ_k	0.9986	0.3306	1.0123	0.7430	0.8816	0.9907	0.9894	0.5457	1.0090	0.9572	0.9528	1.0027	0.9985	0.8381	1.0508	0.9996	0.9864	0.9983
σ_{ka}	1.0652	0.0803	1.2629	0.6479	0.6580	0.9906	0.9424	0.2538	1.2444	0.9357	0.8461	1.0033	0.9850	0.6721	1.2313	0.9983	0.9518	0.9957
σ_y	1.0125	0.3506	1.0616	0.7828	0.8882	0.9970	0.9874	0.5502	1.0514	0.9605	0.9532	1.0010	0.9969	0.8359	1.0502	0.9994	0.9858	0.9990
σ_c	1.0502	0.2903	1.2024	0.8405	0.7934	1.0035	0.9699	0.5000	1.1752	0.9772	0.9135	1.0051	0.9941	0.8183	1.1401	1.0033	0.9759	1.0032
σ_i	0.9191	0.6467	0.7902	0.6459	1.2673	1.0109	1.0791	0.7899	0.7217	0.9341	1.1336	1.0245	1.0374	0.9395	0.7716	1.0161	1.0564	1.0134
ρ_y	1.0036	0.9411	1.0101	1.0061	0.9779	0.9999	0.9967	0.9790	1.0102	1.0015	0.9915	1.0000	0.9991	0.9958	1.0084	1.0000	0.9975	1.0000

Tau stands for the Tauchen (1986) method; T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 0.025$, $\alpha = 0.33$, $\beta = 0.984$, ρ= 0.979, σ_ε= 0.0072.

* For the Tauchen (1986) method, $M = 1.6425$ when $N = 5$, $M = 1.9847$ when $N = 10$ and $M = 2.5107$ when $N = 25$.

Table 8 Results obtained when $\delta = 0.025$.

(B) Monte Carlo Simulations

	$N = 5$						$N = 10$						$N = 25$					
	Generated Values Relative to True Values						Generated Values Relative to True Values						Generated Values Relative to True Values					
	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R	Tau*	T-H	F-1	F-2	A-C	R
ρ	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
σ_ε	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
σ_a	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955	0.9955
σ_k	0.9396	0.5373	0.8201	0.8098	1.1383	0.9888	1.0309	0.7985	0.8222	0.9673	1.0544	0.9990	0.9985	0.9838	0.8668	0.9946	1.0147	0.9961
σ_{ka}	0.9359	0.5100	0.8131	0.8055	1.1313	0.9816	1.0235	0.7796	0.8139	0.9621	1.0487	0.9946	0.9933	0.9776	0.8706	0.9892	1.0094	0.9910
σ_y	0.9780	0.8478	0.9399	0.9372	1.0403	0.9928	1.0061	0.9306	0.9402	0.9865	1.0139	0.9966	0.9963	0.9915	0.9567	0.9950	1.0014	0.9956
σ_c	0.9928	0.9892	0.9805	0.9912	1.0177	0.9978	1.0017	0.9994	0.9806	0.9950	1.0049	0.9979	0.9973	1.0011	0.9863	0.9959	0.9995	0.9968
σ_i	0.9606	0.7478	0.8399	0.8943	1.1553	1.0053	1.0434	0.9278	0.8426	0.9842	1.0671	1.0118	1.0088	1.0203	0.8979	1.0007	1.0263	1.0062
ρ_y	0.9995	0.9959	0.9986	0.9986	1.0008	0.9998	1.0001	0.9984	0.9986	0.9997	1.0003	0.9999	0.9999	0.9998	0.9990	0.9999	1.0000	0.9999

Tau stands for the Tauchen (1986) method; T-H stands for the original Tauchen-Hussey method; F-1 stands for the first variation of T-H; F-2 stands for the second variation; A-C stands for the Adda-Cooper method; R stands for the Rouwenhorst method.

Parameter values: $\delta = 0.025$, $\alpha = 0.33$, $\beta = 0.984$, ρ= 0.979, σ_ε= 0.0072.

* For the Tauchen (1986) method, $M = 1.6425$ when $N = 5$, $M = 1.9847$ when $N = 10$ and $M = 2.5107$ when $N = 25$.

Appendix A

Fix $N \geq 3$. The objective of this section is to derive a set of equations that can be used to describe the elements in Π_N. The proof of Proposition 1 is built upon these equations.

To begin with, the elements in the first and the last rows of Π_N can be obtained by expanding the polynomials $[p + (1-p)\,t]^{N-1}$ and $(1 - q + qt)^{N-1}$, respectively. Using the binomial formula, we can obtain

$$\pi_{1,j}^{(N)} = \binom{N-1}{j-1} p^{N-j} (1-p)^{j-1}, \tag{16}$$

and

$$\pi_{N,j}^{(N)} = \binom{N-1}{j-1} (1-q)^{N-j} q^{j-1}, \tag{17}$$

for $j = 1, 2, ..., N$.

For all other rows, i.e., $i = 2, ..., N-1$, the elements in Π_N can be defined recursively using the elements in Π_{N-1}. Begin with the system for $N - 1 \geq 2$. The system of polynomials is given by

$$\Phi(t; N-1, i) = [p + (1-p)\,t]^{N-1-i} (1 - q + qt)^{i-1} = \sum_{j=1}^{N-1} \pi_{i,j}^{(N-1)} t^{j-1},$$

for $i = 1, ..., N-1$. There are two ways to relate this system to the one for N:

$$\Phi(t; N, i) = [p + (1-p)\,t]\,\Phi(t; N-1, i), \tag{18}$$

for $i = 1, ..., N-1$, and

$$\Phi(t; N, i) = (1 - q + qt)\,\Phi(t; N-1, i-1), \tag{19}$$

for $i = 2, ..., N$. Substituting (5) into (18) gives

$$\begin{aligned} \sum_{j=1}^{N} \pi_{i,j}^{(N)} t^{j-1} &= [p + (1-p)\,t] \sum_{j=1}^{N-1} \pi_{i,j}^{(N-1)} t^{j-1} \\ &= \sum_{j=1}^{N-1} p\pi_{i,j}^{(N-1)} t^{j-1} + \sum_{j=1}^{N-1} (1-p)\,\pi_{i,j}^{(N-1)} t^{j}, \end{aligned}$$

for $i = 1, ..., N-1$. Similarly, substituting (5) into (19) would give

$$\begin{aligned}\sum_{j=1}^{N} \pi_{i,j}^{(N)} t^{j-1} &= (1-q+qt) \sum_{j=1}^{N-1} \pi_{(i-1),j}^{(N-1)} t^{j-1} \\ &= \sum_{j=1}^{N-1} (1-q)\, \pi_{(i-1),j}^{(N-1)} t^{j-1} + \sum_{j=1}^{N-1} q\pi_{(i-1),j}^{(N-1)} t^{j},\end{aligned}$$

for $i = 2, ..., N$. The following can be obtained by comparing the coefficients for $i = 1, 2, ..., N-1$,

$$\pi_{i,1}^{(N)} = p\pi_{i,1}^{(N-1)} = (1-q)\, \pi_{(i-1),1}^{(N-1)} \tag{20}$$

$$\begin{aligned}\pi_{i,j}^{(N)} &= p\pi_{i,j}^{(N-1)} + (1-p)\, \pi_{i,(j-1)}^{(N-1)} \\ &= (1-q)\, \pi_{(i-1),j}^{(N)} + q\pi_{(i-1),(j-1)}^{(N)}, \quad \text{for } j = 2, ..., N-1,\end{aligned} \tag{21}$$

and

$$\pi_{i,N}^{(N)} = (1-p)\, \pi_{i,(N-1)}^{(N-1)} = q\pi_{(i-1),N}^{(N-1)}. \tag{22}$$

Appendix B

Proof of Proposition 1

Fix $N \geq 2$. The elements in the Rouwenhorst matrix $\Theta_N = \left[\theta_{i,j}^{(N)}\right]$ are governed by the following sets of equations:

For the elements in the first row,

$$\theta_{1,j}^{(N)} = \begin{cases} p\theta_{1,j}^{(N-1)} & \text{if } j = 1 \\ p\theta_{1,j}^{(N-1)} + (1-p)\, \theta_{1,(j-1)}^{(N-1)} & \text{if } j = 2, ..., N-1 \\ (1-p)\, \theta_{1,(j-1)}^{(N-1)} & \text{if } j = N. \end{cases} \tag{23}$$

For the elements in the final row,

$$\theta_{N,j}^{(N)} = \begin{cases} (1-q)\,\theta_{(N-1),j}^{(N-1)} & \text{if } j = 1 \\ (1-q)\,\theta_{(N-1),j}^{(N-1)} + q\theta_{(N-1),(j-1)}^{(N-1)} & \text{if } j = 2, ..., N-1 \\ q\theta_{(N-1),(j-1)}^{(N-1)} & \text{if } j = N. \end{cases} \tag{24}$$

For the elements in row $i = 2, ..., N-1$,

$$\theta_{i,j}^{(N)} = \begin{cases} \frac{1}{2}\left[p\theta_{i,j}^{(N-1)} + (1-q)\,\theta_{(i-1),j}^{(N-1)}\right] & \text{if } j = 1 \\ \frac{1}{2}\left[(1-p)\,\theta_{i,(j-1)}^{(N-1)} + q\theta_{(i-1),(j-1)}^{(N-1)}\right] & \text{if } j = N, \end{cases} \tag{25}$$

and for $j = 2, ..., N-1$,

$$\begin{aligned} \theta_{i,j}^{(N)} &= \frac{1}{2}\left[p\theta_{i,j}^{(N-1)} + (1-p)\,\theta_{i,(j-1)}^{(N-1)} + (1-q)\,\theta_{(i-1),j}^{(N-1)}\right. \\ &\quad \left. + q\theta_{(i-1),(j-1)}^{(N-1)}\right], \end{aligned} \tag{26}$$

For any given Θ_{N-1}, the system of equations (23)-(26) defines a unique Θ_N. Similarly, for any given Π_{N-1}, the system of equations (16)-(22) defines a unique Π_N. Since $\Theta_2 = \Pi_2$, it suffice to show that the elements in Π_N generated by (16)-(22) satisfies the system (23)-(26).

Consider the first row (i.e., $i = 1$) in Π_N. According to (16),

$$\pi_{11}^{(N)} = p^{N-1} = p\pi_{11}^{(N-1)},$$

and

$$\pi_{1,N}^{(N)} = (1-p)^{N-1} = (1-p)\,\pi_{1,(N-1)}^{(N-1)}.$$

For $j = 2, ..., N-1$, since

$$\pi_{1,j}^{(N-1)} = \binom{N-2}{j-1} p^{N-1-j}\,(1-p)^{j-1},$$

$$\pi_{1,(j-1)}^{(N-1)} = \binom{N-2}{j-2} p^{N-j}\,(1-p)^{j-2},$$

and

$$\binom{N-1}{j-1} = \binom{N-2}{j-1} + \binom{N-2}{j-2},$$

we have

$$\pi_{1,j}^{(N)} = p\pi_{1,j}^{(N-1)} + (1-p)\,\pi_{1,(j-1)}^{(N-1)}.$$

This shows that the elements in the first row of Π_N satisfies (23). Using (17) and the same procedure, one can show that the elements in the last row of Π_N satisfies (24).

The rest of the proof follows immediately from (20)-(22). For any row $i = 2, ..., N-1$ in Π_N, (20) implies

$$\pi_{i,1}^{(N)} = \frac{1}{2}\left[p\pi_{i,1}^{(N-1)} + (1-q)\,\pi_{(i-1),1}^{(N-1)}\right].$$

Similarly, (21) and (22) imply

$$\pi_{i,N}^{(N)} = \frac{1}{2}\left[(1-p)\,\pi_{i,(N-1)}^{(N-1)} + q\pi_{(i-1),N}^{(N-1)}\right],$$

and

$$\begin{aligned}\pi_{ij}^{(N)} &= \frac{1}{2}\left[p\pi_{ij}^{(N-1)} + (1-p)\,\pi_{i,(j-1)}^{(N-1)} + (1-q)\,\pi_{(i-1),j}^{(N-1)}\right. \\ &\quad \left. +q\pi_{(i-1),(j-1)}^{(N-1)}\right], \end{aligned} \tag{27}$$

for $j = 2, ..., N-1$, respectively. Thus all the elements in row $i = 2, ..., N-1$ in Π_N satisfies (25) and (26). This completes the proof.

Proof of Lemma 2

It suffice to check that all the elements of Π_N are strictly positive. From (16) and (17), it is obvious that the elements in the first and the last rows are strictly positive. For the other rows, a simple induction argument is used. First, Π_2 is a stochastic matrix with non-zero entries. Suppose the result is true for $N-1 \geq 2$. It follows from (20)-(22) that $\pi_{ij}^{(N)} > 0$ for $i = 2, ..., N-1$ and for $j = 1, 2, ..., N$. This completes the proof.

Proof of Proposition 4

As mentioned in the proof of Proposition 1, the first column of Π_N is given by

$$\pi_{i,1}^{(N)} = p^{N-i}\left(1-q\right)^{i-1},$$

for $i = 1, 2, ..., N$. Define $\widehat{\lambda}_i^{(N)}$ as in (6). Then

$$\begin{aligned}
\sum_{i=1}^{N} \widehat{\lambda}_i^{(N)} \pi_{i,1}^{(N)} &= \sum_{i=1}^{N} \binom{N-1}{i-1} s^{N-i}\left(1-s\right)^{i-1} p^{N-i}\left(1-q\right)^{i-1} \\
&= \sum_{i=1}^{N} \binom{N-1}{i-1} \left(sp\right)^{N-i}\left(1-s\right)^{i-1}\left(1-q\right)^{i-1} \\
&= \left[sp + \left(1-s\right)\left(1-q\right)\right]^{N} \\
&= s^N = \widehat{\lambda}_1^{(N)}.
\end{aligned}$$

For all other columns except the first one, an induction argument is used to prove the result. As mentioned in the text, the guess is correct when $N = 2$. Suppose the guess is correct for some $N \geq 2$, i.e.,

$$\widehat{\lambda}_j^{(N)} = \sum_{i=1}^{N} \widehat{\lambda}_i^{(N)} \pi_{i,j}^{(N)}, \qquad \text{for } j = 1, 2, ..., N. \tag{28}$$

We have already proved that this is true when $j = 1$, so proceeds to $j = 2, ..., N+1$.

Using (6), the following can be derived

$$\widehat{\lambda}_i^{(N+1)} = \begin{cases} s\widehat{\lambda}_i^{(N)} & \text{for } i = 1 \\ s\widehat{\lambda}_i^{(N)} + \left(1-s\right)\widehat{\lambda}_{i-1}^{(N)} & \text{for } i = 2, ..., N, \\ \left(1-s\right)\widehat{\lambda}_{i-1}^{(N)} & \text{for } i = N+1. \end{cases} \tag{29}$$

Using these one can obtain

$$\begin{aligned}
&\sum_{i=1}^{N+1}\widehat{\lambda}_i^{(N+1)}\pi_{i,j}^{(N+1)} \\
= \;& \widehat{\lambda}_1^{(N+1)}\pi_{1,j}^{(N+1)}+\sum_{i=2}^{N}\widehat{\lambda}_i^{(N+1)}\pi_{i,j}^{(N+1)}+\widehat{\lambda}_{N+1}^{(N+1)}\pi_{(N+1),j}^{(N+1)} \\
= \;& s\widehat{\lambda}_1^{(N)}\pi_{1,j}^{(N+1)}+\sum_{i=2}^{N}\left[s\widehat{\lambda}_i^{(N)}+(1-s)\,\widehat{\lambda}_{i-1}^{(N)}\right]\pi_{i,j}^{(N+1)}+(1-s)\,\widehat{\lambda}_N^{(N+1)}\pi_{(N+1),j}^{(N+1)} \\
= \;& \sum_{i=1}^{N}s\widehat{\lambda}_i^{(N)}\pi_{i,j}^{(N+1)}+\sum_{i=1}^{N-1}(1-s)\,\widehat{\lambda}_i^{(N)}\pi_{(i+1),j}^{(N+1)}+(1-s)\,\widehat{\lambda}_N^{(N+1)}\pi_{(N+1),j}^{(N+1)} \\
= \;& \sum_{i=1}^{N}s\widehat{\lambda}_i^{(N)}\pi_{i,j}^{(N+1)}+\sum_{i=1}^{N}(1-s)\,\widehat{\lambda}_i^{(N)}\pi_{(i+1),j}^{(N+1)}. \qquad (30)
\end{aligned}$$

Based on (21), the following can be obtained

$$\pi_{i,j}^{(N+1)}=p\pi_{i,j}^{(N)}+(1-p)\,\pi_{i,j-1}^{(N)},$$

and

$$\pi_{i+1,j}^{(N+1)}=(1-q)\,\pi_{i,j}^{(N)}+q\pi_{i,(j-1)}^{(N)},$$

for $j=2,3,...,N$. Substituting these into (30) gives

$$\begin{aligned}
&\sum_{i=1}^{N+1}\widehat{\lambda}_i^{(N+1)}\pi_{i,j}^{(N+1)} \\
= \;& s\sum_{i=1}^{N}\widehat{\lambda}_i^{(N)}\left[p\pi_{i,j}^{(N)}+(1-p)\,\pi_{i,(j-1)}^{(N)}\right]+(1-s)\sum_{i=1}^{N}\widehat{\lambda}_i^{(N)}\left[(1-q)\,\pi_{i,j}^{(N)}+q\pi_{i,(j-1)}^{(N)}\right] \\
= \;& [sp+(1-s)\,(1-q)]\sum_{i=1}^{N}\widehat{\lambda}_i^{(N)}\pi_{i,j}^{(N)}+[s\,(1-p)+(1-s)\,q]\sum_{i=1}^{N}\widehat{\lambda}_i^{(N)}\pi_{i,(j-1)}^{(N)}.
\end{aligned}$$

Using the induction hypothesis (28), the following can be obtained

$$\begin{aligned}
\sum_{i=1}^{N+1}\widehat{\lambda}_i^{(N+1)}\pi_{i,j}^{(N+1)} &= [sp+(1-s)\,(1-q)]\,\widehat{\lambda}_j^{(N)}+[s\,(1-p)+(1-s)\,q]\,\widehat{\lambda}_{j-1}^{(N)} \\
&= s\widehat{\lambda}_j^{(N)}+(1-s)\,\widehat{\lambda}_{j-1}^{(N)} \\
&= \widehat{\lambda}_j^{(N+1)},
\end{aligned}$$

for $j = 2, 3, ..., N$. The last line is obtained by using (29). Since $\sum_{i=1}^{N+1} \widehat{\lambda}_i^{(N+1)} = 1$ and $\sum_{j=1}^{N+1} \pi_{i,j}^{(N+1)} = 1$, the remaining equation

$$\sum_{i=1}^{N+1} \widehat{\lambda}_i^{(N+1)} \pi_{i,j}^{(N+1)} = \widehat{\lambda}_j^{(N+1)}, \qquad \text{for } j = N+1,$$

must be satisfied. This completes the proof.

Appendix C

The objective of this section is to derive the moments listed on Table 1. Since it is understood that these are moments for an N-state Markov chain, the notations $\pi_{i,j}^{(N)}$ and $\lambda_j^{(N)}$ are simplified to become $\pi_{i,j}$ and λ_j, respectively.

Preliminaries

The following result is used in deriving the conditional mean for the Markov chain.

Lemma 5 *For any $N \geq 2$, and for $i = 1, ..., N$,*

$$\sum_{j=1}^{N} \pi_{i,j} (j-1) = (1-p)(N-i) + (i-1) q, \tag{31}$$

$$\sum_{j=1}^{N} \pi_{1,j} (j-1)^2 = \left[\sum_{j=1}^{N} \pi_{i,j} (j-1) \right]^2 + (N-i)(1-p) p + (i-1) q (1-q). \tag{32}$$

Proof. Recall the following expression

$$[p + (1-p) t]^{N-i} (1 - q + qt)^{i-1} = \sum_{j=1}^{N} \pi_{i,j} t^{j-1}, \tag{33}$$

for $i = 1, ..., N$. Equation (31) can be obtained in two steps: (i) Differentiate both sides of (33) with respect to t. (ii) Set $t = 1$.

Equation (32) can be obtained as follows: Fix $i = 1, ..., N$. Differentiate both sides of (33) with respect to t twice and set $t = 1$. This gives

$$\begin{aligned}\sum_{j=1}^{N} \pi_{i,j} (j-1)(j-2) &= \sum_{j=1}^{N} \pi_{i,j} (j-1)^2 - \sum_{j=1}^{N} \pi_{i,j} (j-1) \\ &= [(N-i)(1-p) + (i-1)q]^2 - (N-i)(1-p)^2 - (i-1)q^2 \\ &= \left[\sum_{j=1}^{N} \pi_{i,j} (j-1)\right]^2 - (N-i)(1-p)^2 - (i-1)q^2.\end{aligned}$$

Equation (32) can be obtained by combining this and equation (31). This completes the proof of Lemma 5. ■

The following equations are useful in deriving the other moments. For a binomial distribution with parameters $N-1$ and $1-s$, the first two moments are given by

$$\sum_{i=1}^{N} \binom{N-1}{i-1} s^{N-i} (1-s)^{i-1} (i-1) = (N-1)(1-s), \tag{34}$$

$$\begin{aligned}&\sum_{i=1}^{N} \binom{N-1}{i-1} s^{N-i} (1-s)^{i-1} (i-1)^2 \\ = \;& (N-1)(1-s)s + (N-1)^2 (1-s)^2. \end{aligned} \tag{35}$$

Conditional Mean

We are now ready to compute the conditional means. Conditional on $y_t = \overline{y}_i$, the mean value of y_{t+1} is given by

$$\begin{aligned}E(y_{t+1} | y_t = \overline{y}_i) &= \sum_{j=1}^{N} \pi_{i,j} \overline{y}_j = \sum_{j=1}^{N} \pi_{i,j} \left[-\psi + \frac{2\psi}{N-1}(j-1)\right] \\ &= -\psi + \frac{2\psi}{N-1} \sum_{j=1}^{N} \pi_{i,j} (j-1).\end{aligned}$$

It follows from (31) that

$$\begin{aligned}\sum_{j=1}^{N} \pi_{i,j} (j-1) &= (1-p)(N-i) + (i-1)q \\ &= (1-p)(N-1) + (q+p-1)(i-1).\end{aligned}$$

Hence

$$\begin{aligned} E\left(y_{t+1}|y_t=\overline{y}_i\right) &= -\psi+\frac{2\psi}{N-1}\left[(1-p)(N-1)+(q+p-1)(i-1)\right] \\ &= -\psi+2\psi(1-p)+(q+p-1)\frac{2\psi}{N-1}(i-1) \\ &= (q-p)\psi+(q+p-1)\overline{y}_i. \end{aligned} \tag{36}$$

Conditional Variance

Conditional on $y_t=\overline{y}_i$, the variance of y_{t+1} is given by

$$var\left(y_{t+1}|\overline{y}_i\right)=\sum_{j=1}^{N}\pi_{i,j}\overline{y}_j^2-\left(\sum_{j=1}^{N}\pi_{i,j}\overline{y}_j\right)^2,$$

where

$$\sum_{j=1}^{N}\pi_{i,j}\overline{y}_j^2=\psi^2-\frac{4\psi^2}{N-1}\sum_{j=1}^{N}\pi_{i,j}(j-1)+\frac{4\psi^2}{(N-1)^2}\sum_{j=1}^{N}\pi_{i,j}(j-1)^2,$$

and

$$\left(\sum_{j=1}^{N}\pi_{ij}\overline{y}_j\right)^2=\psi^2-\frac{4\psi^2}{N-1}\sum_{j=1}^{N}\pi_{i,j}(j-1)+\frac{4\psi^2}{(N-1)^2}\left[\sum_{j=1}^{N}\pi_{i,j}(j-1)\right]^2.$$

It follows from (32) that

$$var\left(y_{t+1}|\overline{y}_i\right)=\frac{4\psi^2}{(N-1)^2}\left[(N-i)(1-p)p+(i-1)q(1-q)\right].$$

Unconditional Mean

The unconditional mean of the Markov chain is given by

$$\begin{aligned} \sum_{i=1}^{N}\lambda_i\overline{y}_i &= \sum_{i=1}^{N}\lambda_i E\left(y_{t+1}|y_t=\overline{y}_i\right) \\ &= \sum_{i=1}^{N}\lambda_i\left[(q-p)\psi+(q+p-1)\overline{y}_i\right] \\ &= (q-p)\psi+(q+p-1)\sum_{i=1}^{N}\lambda_i\overline{y}_i. \end{aligned}$$

Hence

$$\sum_{i=1}^{N} \lambda_i \overline{y}_i = \frac{(q-p)\,\psi}{2-(p+q)} \equiv \mu. \tag{37}$$

Unconditional Second Moment

$$\begin{aligned}
\sum_{i=1}^{N} \lambda_i \overline{y}_i^2 &= \sum_{i=1}^{N} \lambda_i \left[-\psi + \frac{2\psi}{N-1}(i-1)\right]^2 \\
&= \sum_{i=1}^{N} \lambda_i \left[\psi^2 - \frac{4\psi^2}{N-1}(i-1) + \frac{4\psi^2}{(N-1)^2}(i-1)^2\right] \\
&= \psi^2 - \frac{4\psi^2}{N-1}\sum_{i=1}^{N} \lambda_i (i-1) + \frac{4\psi^2}{(N-1)^2}\sum_{i=1}^{N} \lambda_i (i-1)^2 .
\end{aligned}$$

Using (34) and (35), we have

$$\begin{aligned}
\sum_{i=1}^{N} \lambda_i \overline{y}_i^2 &= \psi^2 - 4\psi^2(1-s) + \frac{4\psi^2(1-s)\,s}{N-1} + 4\psi^2(1-s)^2 \\
&= \psi^2\left[1 - 4(1-s)\,s + \frac{4(1-s)\,s}{N-1}\right].
\end{aligned}$$

First-order Autocovariance

First consider the following expression,

$$\begin{aligned}
E(y_t y_{t+1}) &= \sum_{i=1}^{N} \lambda_i E(y_{t+1} y_t | y_t = \overline{y}_i) \\
&= \sum_{i=1}^{N} \lambda_i \overline{y}_i E(y_{t+1} | y_t = \overline{y}_i).
\end{aligned}$$

Using (36), we have

$$\begin{aligned}
E(y_t y_{t+1}) &= \sum_{i=1}^{N} \lambda_i \overline{y}_i \left[(q-p)\,\psi + (q+p-1)\,\overline{y}_i\right] \\
&= (q-p)\,\psi \sum_{i=1}^{N} \lambda_i \overline{y}_i + (q+p-1)\sum_{i=1}^{N} \lambda_i \overline{y}_i^2 .
\end{aligned} \tag{38}$$

Let σ_y^2 be the unconditional variance of the Markov chain so that

$$\sigma_y^2 = \sum_{i=1}^{N} \lambda_i \overline{y}_i^2 - \mu^2,$$

where μ is the unconditional mean defined in (37). Substituting this into (38) gives

$$\begin{aligned} & E\left(y_t y_{t+1}\right) \\ = \; & (q-p)\,\psi\mu + (q+p-1)\left[\sigma_y^2 + \mu^2\right] \\ = \; & \left[(q-p)\,\psi + (q+p-1)\,\mu\right]\mu + (q+p-1)\,\sigma_y^2, \end{aligned}$$

where

$$(q-p)\,\psi + (q+p-1)\,\mu = \frac{(q-p)\,\psi}{2-(p+q)} = \mu.$$

Hence

$$E\left(y_t y_{t+1}\right) = \mu^2 + (q+p-1)\,\sigma_y^2.$$

Thus the first-order autocovariance is given by

$$E\left[\left(y_t - \mu\right)\left(y_{t+1} - \mu\right)\right] = E\left(y_t y_{t+1}\right) - \mu^2 = (q+p-1)\,\sigma_y^2.$$

References

[1] Adda, J., Cooper, R., 2003. Dynamic Economics: Quantitative Methods and Applications. MIT Press, Cambridge, MA.

[2] Aruoba, S., Fernández-Villaverde, J., Rubio-Ramírez, J., 2006. Comparing Solution Methods for Dynamic Equilibrium Economies. Journal of Economic Dynamics and Control 30, 2477-2508.

[3] Burnside, C., 1999. Discrete State-Space Methods for the Study of Dynamic Economies. In: Marimon, R., Scott, A. (Ed.), Computational Methods for the Study of Dynamic Economies. Oxford University Press, Oxford, 95-113.

[4] Flodén, M., 2008. A Note on the Accuracy of Markov-chain Approximations to Highly Persistent AR(1) Processes. Economic Letters 99, 516-520.

[5] King, R.G., Rebelo, S.T., 1999. Resuscitating Real Business Cycles. In: Taylor, B.J., Woodford, M. (Ed.), Handbook of Macroeconomics, Volume 1. Elsevier, Amsterdam, 927-1007.

[6] Lkhagvasuren, D., Galindev, R., 2008. Discretization of Highly-Persistent Correlated AR(1) Shocks. Unpublished manuscript, Concordia University.

[7] Lucas, R.E., 1978. Asset Prices in an Exchange Economy. Econometrica 46, 1429-1445.

[8] Rouwenhorst, K.G., 1995. Asset Pricing Implications of Equilibrium Business Cycle Models. In: Cooley, T.F. (Ed.), Frontiers of Business Cycle Research. Princeton University Press, Princeton, NJ, 294-330.

[9] Santos, M.S., Peralta-Alva, A., 2005. Accuracy of Simulations for Stochastic Dynamic Models. Econometrica 73, 1939-1976.

[10] Tauchen, G., 1986. Finite State Markov-chain Approximations to Univariate and Vector Autoregressions. Economic Letters 20, 177-181.

[11] Tauchen, G., 1990. Solving the Stochastic Growth Model by using Quadrature Methods and Value-Function Iterations. Journal of Business and Economic Statistics 8, 49-51.

[12] Tauchen, G., Hussey, R., 1991. Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models. Econometrica 59, 371-396.

[13] Taylor, J., Uhlig, H., 1990. Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods. Journal of Business and Economic Statistics 8, 1-18.