Online Appendix

Supplemental Material for

"A Moment-Matching Method for Approximating Vector Autoregressive Processes by Finite-State Markov Chains"

Nikolay Gospodinov
Federal Reserve Bank of Atlanta

Damba Lkhagvasuren
Concordia University
and CIREQ

August 3, 2013

A Quality of Approximation of Tauchen's Method for Highly Persistent Data

The VAR model that describes the dynamics of the underlying continuous-valued process is given by

$$\mathbf{y}_t = A\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{A.1}$$

where $\boldsymbol{\varepsilon}_t$ is i.i.d. $\mathcal{N}(0, \Omega)$, Ω is a diagonal matrix with an i-th diagonal element ω_i^2 and Σ is the unconditional covariance matrix of $\mathbf{y}_t$ with an i-th diagonal element σ_i^2.

Let $(\tilde{\mathbf{y}}_1^{(n)}, \tilde{\mathbf{y}}_2^{(n)}, \cdots, \tilde{\mathbf{y}}_t^{(n)}, \cdots, \tilde{\mathbf{y}}_T^{(n)})$ be a realization of the n-state Markov chain of length T, approximated over n grid points using the Tauchen's (1986) method with the standard normal CDF Φ_i, and $\tilde{\omega}_i$ denote the square root of the i-th diagonal element of the covariance matrix $\tilde{\Omega}$ of $\tilde{\boldsymbol{\varepsilon}}_t = \tilde{\mathbf{y}}_t^{(n)} - A\tilde{\mathbf{y}}_{t-1}^{(n)}$. In what follows, we keep n fixed and and perform the analysis as $T \to \infty$.

In Proposition 1 below, we show that calculating the transition probabilities using the continuous distribution functions does not always deliver meaningful approximations. In particular, Tauchen's (1986) method fails to approximate the variability in $\mathbf{y}_t$ as one or more of the roots of the underlying continuous-valued VAR process $\mathbf{y}_t$ approach the unit circle. This problem arises because most of the existing approximation methods, including the method by Tauchen (1986), target only the first conditional moment of the continuous-valued process $\mathbf{y}_t$.

Proposition 1. *For any set of integers $(N_1, N_2, \cdots, N_M)$ and any arbitrarily small positive number ϵ, there always exists a highly persistent vector autoregressive process for which $\tilde{\omega}_i/\omega_i < \epsilon$ for all i.*

Proof. Since we are interested in the behavior of highly persistent processes, it is convenient to reparameterize the matrix A as local-to-unity (see Phillips, 1987, for example). In particular, the matrix A is reparameterized as a function of T as (Elliott, 1998)

$$A = I_M - \frac{C}{T}, \tag{A.2}$$

where $C = \text{diag}(c_1, c_2, ..., c_M)$ with $c_i > 0$ being fixed constants for all $i = 1, ..., M$.[1] This is an artificial statistical device in which the parameter space for each individual process is

[1]We can also allow for non-zero off-diagonal elements of C (see Gospodinov, Maynard and Pesavento, 2011) provided that this does not induce nonstationarity and preserves the stability of the process. The proof that we present below goes through for this more general specification but at the cost of more complicated notation.

a shrinking neighborhood of one as T increases. This parameterization proves to be very useful for studying the properties of strongly dependent processes as $T \to \infty$.

First note that using this reparameterization, the innovation variance matrix for the continuous-valued process can be expressed as

$$\Omega = \frac{C\Sigma + \Sigma C'}{T} - \frac{C\Sigma C'}{T^2}. \tag{A.3}$$

and the variance for the i-th innovation is

$$\omega_i^2 = \frac{2c_i\sigma_i^2}{T} - \frac{c_i^2\sigma_i^2}{T^2}. \tag{A.4}$$

For Tauchen's (1986) method, the probability that the process y_i switches from state j (corresponding to grid point $\bar{y}_i^{(j)}$) to any other state is given by

$$1 - \pi_{j,j}^{(i)} = 1 - \Pr\left(\left|\varepsilon_i - \frac{c_i\bar{y}_i^{(j)}}{T}\right| \leq 2\triangle_i\right), \tag{A.5}$$

where $\pi_{j,j}^{(i)}$ is the j-th diagonal element of the i-th $N_i \times N_i$ block of matrix Π and $\triangle_i$ denotes the distance between the grid points. As $T \to \infty$, the persistence of the process increases and $0 < \bar{y}_i^{(j)}/T < 2\triangle_i$ (for all j) with probability approaching one.[2] Therefore,

$$1 - \pi_{j,j}^{(i)} \leq 1 - \Pr\left(|\varepsilon_i| \leq 2\triangle_i\right) = 2\left(1 - \Phi\left(\frac{2\triangle_i}{\sqrt{2c_i\sigma_i^2/T - c_i^2\sigma_i^2/T^2}}\right)\right)$$

and thus,

$$\frac{1 - \pi_{j,j}^{(i)}}{\omega_i^2} < \frac{2\left(1 - \Phi\left(\frac{\triangle_i\sqrt{2T}}{\sigma_i\sqrt{c_i}}\right)\right)}{2c_i\sigma_i^2/T - c_i^2\sigma_i^2/T^2} \tag{A.6}$$

for all j. Since

$$\Phi\left(\frac{\triangle_i\sqrt{2T}}{\sigma_i\sqrt{c_i}}\right) \to 1 \text{ as } T \to \infty \tag{A.7}$$

by l'Hopital's rule,

$$\lim_{T\to\infty} \frac{1 - \pi_{j,j}^{(i)}}{\omega_i^2} = \frac{\triangle_i}{2\sigma_i^3 c_i^{3/2}\pi^{1/2}} \frac{1}{(1/T^{3/2} - c_i/T^{5/2})\exp(2c_i\,\triangle_i^2\,T/\sigma_i^2)} = 0. \tag{A.8}$$

Hence, since the limiting behavior of the conditional variance of the Markov-chain approxi-

[2]Note that Δ_i is fixed. While one can reduce the speed of the convergence by making m a decreasing function of the persistence, such an adjustment will severely distort the unconditional variances.

mation is determined by the limiting behavior of $1 - \pi_{j,j}^{(i)}$,

$$\frac{\tilde{\omega}_i^2}{\omega_i^2} \to 0 \text{ as } T \to \infty. \tag{A.9}$$

This completes the proof of Proposition 1.

B Asymptotic Validity of the MM Method

In this section, we establish the asymptotic validity of the proposed moment-matching method for approximating conditional expectations of nonlinear functions and solving functional equations. For notational simplicity, we present the results for a scalar continuous-valued process with conditional density $f(y'|y)$ although the results can be extended to the vector case $f(\mathbf{y}|\mathbf{x})$, where $\mathbf{y} \in \mathcal{R}^M$ and $\mathbf{x} = (\mathbf{y}_{-1}, ..., \mathbf{y}_{-L}) \in \mathcal{R}^{M \cdot L}$. Consider the function

$$e_g(y) = \int g(y') f(y'|y) dy, \tag{B.1}$$

where $g(y) \in C_0[a, b]$ and $C_0[a, b]$ denotes the space of continuous functions on $[a, b]$ with $a < b$ and both a and b are finite. Assume that the support of $f(y'|y)$ is a subset of $[a, b] \times [a, b]$ and $f(y'|y)$ is jointly continuous in y' and y. Let $\tilde{y}$ denote the n-state Markov-chain approximation proposed that takes on the discrete values $\{\bar{y}^{(1)}, \bar{y}^{(2)}, \cdots, \bar{y}^{(n)}\}$ and transition probabilities $\pi_{j,k}^{(n)} = \Pr(\tilde{y}' = \bar{y}^{(k)} | \tilde{y} = \bar{y}^{(j)})$. Let

$$e_{gn}(y) = \sum_{k=1}^{n} g(\bar{y}^{(k)}) \pi_{j,k}^{(n)}. \tag{B.2}$$

Following Tauchen and Hussey (1991), we need to show the uniform convergence result

$$\sup_{y \in [a,b]} |e_{gn}(y) - e_g(y)| \xrightarrow{p} 0 \tag{B.3}$$

as $n \to \infty$.

The pointwise convergence of the conditional distribution of the Markov chain $\tilde{y}'$ given $\tilde{y} = \bar{y}^{(j)}$ to the conditional distribution of y' given $y = \mu(j)$ can be inferred from noting that the transition probability matrix for our method can be expressed in a polynomial form (see Kopecky and Suen, 2010) and by appealing to the Stone-Weierstrass approximation theorem. Finally, the condition that $e_{gn}(y)$ is uniformly bounded converts the pointwise convergence into uniform convergence. As a result, $e_{gn}(y)$ is equicontinuous which is a sufficient condition

for the uniform convergence result

$$\sup_{y\in[a,b]} |e_{gn}(y) - e_g(y)| \xrightarrow{p} 0 \text{ as } n \to \infty. \tag{B.4}$$

C Additional Numerical Results

In this section, we provide additional numerical results not reported in the paper. In particular, we consider the bivariate VAR(1) case ($M = 2$) with

$$\boldsymbol{\varepsilon}_t \sim i.i.d.\ \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}\right) \tag{C.1}$$

and $A = A_0^R$, where

$$A_0 = \begin{pmatrix} 0.9579 & 0.0505 \\ 0.0337 & 0.9242 \end{pmatrix} \tag{C.2}$$

and R is a positive integer set to 1 and 10.[3] It is straightforward to see that higher values of R imply lower persistence. As in Tauchen (1986), we choose nine grid points for each component: $\overline{N} = N_1 = N_2 = 9$. When using Tauchen's method, we set $m_i = 1.2 \ln N_i$ (Floden, 2008). (Below in subsection C.3, we consider different values for m_i while targeting unconditional variances as Kopecky and Suen (2010) do.)

C.1 Approximation accuracy

Let $\{\widetilde{\mathbf{y}}_t\}_{t=1}^{\tau}$ denote the simulated time series either from the Markov chain approximation by Tauchen (1986) or the method proposed in this paper. The accuracy of the two approximations can then be examined by estimating or computing the key parameters of the initial process. The parameters of interest are the unconditional variances of y_1 and y_2 (denoted by σ_1^2 and σ_2^2), the correlation coefficient between y_1 and y_2, and the persistence measures $1-\varsigma_1$ and $1-\varsigma_2$, where ς_1 and ς_2 are the two roots (eigenvalues) of matrix A. As in Tauchen (1986) and Tauchen and Hussey (1991), the simulated counterpart of A, $\hat{A}$, is obtained by fitting a VAR(1) to $\{\widetilde{\mathbf{y}}_t\}_{t=1}^{\tau}$.

The unconditional variances are directly calculated using the invariant mass distribution of $\widetilde{\mathbf{y}}_t$. The invariant distribution P (a vector of length N^*) is obtained by satisfying the

[3] The matrix A_0 is chosen for comparison purposes. Specifically, when $R = 10$,

$$A = A_0^{10} = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.5 \end{pmatrix}.$$

Therefore, the vector autoregressive process coincides with the one considered in Tauchen (1986).

following equation:

$$\Pi^{\mathrm{T}} P = P. \tag{C.3}$$

However, the evaluation of the approximation accuracy of the eigenvectors and the cross correlation coefficient is based on 1,000 Monte Carlo replications of length $\tau = 2,000,000$.[4] Columns "Tau." and "MM" in Table C.1 summarize the key moments generated by Tauchen's and the MM methods, respectively.

The results suggest that our MM method dominates the method by Tauchen (1986) in terms of bias and RMSE for all parameters of interest across all degrees of persistence. For example, for the less persistent case ($R = 10$), the relative bias for of the estimated $1 - \varsigma_1$, σ_1^2 and σ_2^2, using data generated by Tauchen's (1986) method, is 3.5%, 6.6% and 4.4%, respectively, whereas the corresponding biases for the MM method are 1%, -0.8% and -0.5%. For the more persistent case ($R = 1$), the biases for the method of Tauchen (1986) become -19.3%, 35.6% and 28.7%, while those of the MM method remain almost constant at 1.9%, -0.7% and -0.9%, respectively. So, the advantages of our method become particularly striking when the underlying persistence increases.

It should be noted that for the degree of persistence that is much higher than those considered here, Tauchen's (1986) method fails to produce any time variation in the approximate Markov chain process, which is consistent with our theoretical results in Proposition 1 (also, see Fig. 1 in the text).

C.2 Conditional moments

As before, the distances between the targeted and the generated conditional moments are measured by $|\hat{\mu}_i(j) - \mu_i(j)|$ and $|\hat{\omega}_i^2(j)/\omega_i^2 - 1|$ for each i and j. To assess the overall accuracy of the conditional moments, we consider the weighted averages of these distances across the N^* states using the invariant distribution of $\widetilde{\mathbf{y}}_t$ as weights. The results are presented in the lower panel (Panel C) of Table C.1 and show that the MM method performs extremely well across all parameterizations. Again, this is not surprising since, by construction, this method targets the first two conditional moments of the underlying process. More importantly, the results show that calculating the transition probabilities using the conditional distribution, as in Tauchen (1986), generates a substantial bias in the conditional moments. This numerical finding lends support to our theoretical result in Proposition 1.

[4]Note that the length of the time series is much larger than that considered by Tauchen (1986). The main reason is that, for smaller number of observations, Tauchen's method fails to generate time-varying data for the examples considered here and, thus, renders the numerical evaluation of the methods impossible. Put differently, for shorter time series, the numerical results will be much more favorable for the method developed in this paper.

Table C.1: Approximation Accuracy

Moments		Less Persistence ($R = 10$) Tau.	Tau-adjust.	MM	More Persistence ($R = 1$) Tau.	Tau-adjust.	MM
Panel A. Moments computed using the invariant distribution							
$\hat{\sigma}_1^2$		0.066	0	-0.008	0.3559	0	-0.0071
$\hat{\sigma}_2^2$		0.044	0	-0.005	0.2866	0	-0.0094
Panel B. Moments measured from simulated data							
$\hat{\rho}_{1,2}$	RMSE	0.017	0.017	0.006	0.047	0.047	0.006
	Bias	-0.017	-0.017	-0.006	-0.047	-0.046	-0.005
	Std.	0.002	0.002	0.002	0.003	0.003	0.003
$1 - \hat{\varsigma}_1$	RMSE	0.035	0.035	0.010	0.193	0.193	0.019
	Bias	0.035	0.035	0.010	-0.192	-0.193	0.018
	Std.	0.003	0.003	0.003	0.007	0.007	0.008
$1 - \hat{\varsigma}_2$	RMSE	0.003	0.003	0.001	0.121	0.121	0.003
	Bias	0.003	0.003	0.000	-0.121	-0.121	0.001
	Std.	0.001	0.001	0.001	0.003	0.003	0.003
Panel C. Distance between simulated and true conditional moments							
$\hat{\mu}_1$		0.001	0.002	0.000	0.018	0.016	0.000
$\hat{\mu}_2$		0.001	0.001	0.000	0.004	0.004	0.000
$(\hat{\omega}_1/\omega_1)^2$		0.116	0.052	0.000	0.053	0.242	0.012
$(\hat{\omega}_2/\omega_2)^2$		0.060	0.022	0.000	0.343	0.058	0.001

Notes. This table evaluates the performance of different approximation methods using the example considered in Section C. "Tau." denotes the approximation obtained by the method of Tauchen (1986), whereas "MM" denotes the Markov chain approximation method developed in this paper. "Tau-adjust." denotes the version of Tauchen (1986) where the grid points are adjusted to perfectly match the unconditional variances, σ_i^2, $i \in \{1, 2\}$. The accuracy of the approximation of the moments, except for $\hat{\mu}_1$ and $\hat{\mu}_2$, are reported as the percentage deviation from their true values. Panel A reports the moments calculated using the invariant multivariate distribution. Panel B summarizes the root mean squared error (RMSE), the bias and the standard deviation of cross correlation and eigenvalues. Panel C reports the distance between generated and true conditional moments. Specifically, the numbers in row $\hat{\mu}_i$ are the weighted average of $|\hat{\mu}_i(j) - \mu_i(j)|$ which uses the invariant distribution (not the simulated frequencies) of states $j = 1, 2, ..., N^*$ as weights. Analogously, the numbers in row $\hat{\omega}_i^2/\omega_i^2$ are the weighted average of $|\hat{\omega}_i^2(j)/\omega_i^2 - 1|$ using the same weigths. The numbers smaller than 0.0005 (0.05%) in absolute terms are denoted by 0.000 with their appropriate signs. In the case of perfectly matched moments, the approximation accuracy is denoted by 0.

C.3 Adjusting variances in Tauchen's method

Kopecky and Suen (2010) show that when using Tauchen's method for AR(1) shocks, one can perfectly match the unconditional variances by calibrating the grid points. In this section, we perform a similar analysis for the VAR(1) process. After considering several alternatives, we choose the approach that perfectly matches the unconditional variances without affecting the transition matrix.[5] Specifically, for each $i \in \{1, 2\}$, using equispaced grids on the interval $[-m_i\sigma_i, m_i\sigma_i]$, where $m_i = 1.2 \ln N_i$ (Floden, 2008), we first calculate the transition matrix Π and the associated invariant distribution P (see equation (C.3)). Second, we calculate the unconditional standard deviations of $\tilde{y}_i$, denoted by σ_i^{raw}, using the invariant distribution P. Then, we perfectly match the unconditional variances by replacing the grid points of $\tilde{y}_i$ with equispaced grid points on the interval $[-\tilde{m}_i\sigma_i, \tilde{m}_i\sigma_i]$, where $\tilde{m}_i = 1.2 \ln N_i \times \frac{\sigma_i}{\sigma_i^{\text{raw}}}$.

The moments associated with this modified Tauchen's method are summarized under columns "Tau-adjust." in Table C.1. The numerical results show that when the persistence is low, adjusting unconditional variances improves the conditional moments. However, in the case of high persistence, such adjustment can be counterproductive. Specifically, the approximation accuracy of the conditional variance of $\tilde{y}_1$ deteriorates from 5.3% to 24.2%. Therefore, the issue with approximating the conditional variance, using Floden's (2008) method, remains even after adjusting the unconditional variances, which is consistent with Kopecky and Suen (2010). What is more important for our analysis is that, regardless of the degree of persistence, the quality of the approximation is consistently higher for the MM method compared to the approximation by the adjusted Tauchen's method.

D Discretized Process and Simulated Shocks

The discretized process and a sample of simulated z and g shocks (see Section 4), constructed by the MM method, are available at `https://sites.google.com/site/dlkhagva/var_mmm`.

E MATLAB Programs

Here, we provide MATLAB codes for implementing the MM method described in the paper. The main program is provided in section E.1. The subsequent sections contain MATLAB functions used by the main program.

[5] Another way would be to modify the transition matrix for the purpose of reducing the probability of observing the finite state process at the extreme values of $\overline{y}_i$. However, this type of approach creates an undesirable situation in which any change in the transition matrix affects moments of different components of $\tilde{\mathbf{y}}$ through the dynamic correlation of the multivariate shocks.

E.1 Function var_Markov_MM

```
% This function constructs a finite state Markov chain approximation using
% the MM method in Gospodinov and Lkhagvasuren (2013) for a bivariate
% VAR(1) process considered in the numerical experiment of the paper.
% The VAR(1) process is: y'=Ay+epsilon,
% where var(epsilon) is given by a daigonal matrix Omega.
%
%   INPUT:
%          A0x stands for the 2X2 coefficient matrix A.
%          vex is the 2X2 diagonal matrix, Omega, i.e.
%               Omega(1,1)=omega_{1,1}^2 and Omega(2,2)=omega_{2,2}^2
%          nbar  is the number of grid points for each i.
%          ntune is the control variable, where
%                 setting ntune=0 performs the baseline method (MM0), while
%                 setting ntune>1 performs the full version of the method,
%                 MM. For the examples considered in the paper, ntune was
%                 set to 999. While higher values of ntune gives a better
%                 approximation, the gain becomes negligible beyond
%                 the value 10000.
%   OUTPUT:
%           PN is the N*-by-N* transition matrix, where N* = nbar^2. The
%                [row k, column j] element is the probability the system
%                switches from state j to state k. So, the elements of each
%                column add up to 1.
%           YN is the N*-by-2 matrix of the discrete values of y1 and y2
%                for N* states.
%
function [PN, YN]=var_Markov_MM(A0x,vex,nbar,ntune)
if ntune<0
    error('ntune has to be a positive integer (including zero)');
end
if mod(ntune,1)~=0
    error('ntune has to be a positive integer');
end
nx=ntune+1;
n=nbar;
```

```
n1=n;
n2=n;
[probtemp, z] = rouwen(0,0,1,n);
y1=z;
y2=y1;
A0=A0x;
% normalize the initial var so unconditional variances are 1
[A0new, vynew, vyold, venew]=var_norm(A0, vex);
vy=vyold;
A=A0new;
ve=venew;
pmat=zeros(2,n,n,n);
px=zeros(2,n,n);
for i=1:n
    for j=1:n
        for k=1:2
            mu=A(k,1)*y1(i)+ A(k,2)*y2(j);
            vact=ve(k,k);
            r=sqrt(1-vact);
            [prob1, z] = rouwen(r,0,1,n);
            [v1, p, na,nb, dummy_exceed]=cal_mu_fast(mu,vact,n,z);
            if nx<2
                if na==nb  % if mu is outside of the grids
                    pmat(k,i,j,:)=prob1(:,na);
                else % more relevant case
                    pmat(k,i,j,:)=p*prob1(:,na)+(1-p)*prob1(:,nb);
                end
            else
                if na==nb  % if mu is outside of the grids
                    pmat(k,i,j,:)=prob1(:,na);
                else  % begining of the more relevane
                    B=999*ones(nx,6);
                    ixx=0;
                    for ix=1:nx
               vactx=max(0.00000000000001, vact*(1.0-(ix-1)/(nx-1)));
              [v1x, px, nax,nbx, dummy_exceedx]=cal_mu_fast(mu,vactx,n,z);
```

```
                        if abs(dummy_exceedx)<0.5
                            ixx=ixx+1;
                            B(ixx,:)=[v1x px nax nbx dummy_exceedx vactx];
                        end
                    end
                    if ixx<1
                        pmat(k,i,j,:)=p*prob1(:,na)+(1-p)*prob1(:,nb);
                    else
                        bvectemp=B(:,1)-vact;
                        dif1=abs(bvectemp);
                        [difx, iz]=min(dif1);
                        pz=B(iz,2);
                        naz=B(iz,3);
                        nbz=B(iz,4);
                        vactz=B(iz,6);
                        rz=sqrt(1-vactz);
                        [probz, z] = rouwen(rz,0,1,n);
                        pmat(k,i,j,:)=pz*probz(:,naz)+(1-pz)*probz(:,nbz);
                    end  % end of the more relevane
                end
            end
        end
    end
end
% convert the transition probabilities into a conventional form
PN = bigPPP(pmat,n);
ynum=n*n;
ix=0;
Y=zeros(n*n,2);
for i=1:n
    for j=1:n
        ix=ix+1;
        Y(ix,:)=[y1(i) y2(j)];
    end
end
YN=[Y(:,1)*sqrt(vy(1,1)) Y(:,2)*sqrt(vy(2,2))];
```

E.2 Function cal_mu_fast

```
% cal_mu_fast
% This function calculates the conditional variance of the mixture
% distribution given the conditional mean mu and the conditional variance
% v0 of the mass distributions on the n grids given by z.
%
% For details, see Nikolay Gospodinov and Damba Lkhagvasuren, 2013
%

function [v1, p, na,nb, dummy_exceed]=cal_mu_fast(mu,v0,n,z)
r=sqrt(1-v0);
zm=z*r;
if mu>=zm(n)
    dummy_exceed=1;
    na=n;
    nb=n;
    p=0;
    v1=v0;
elseif mu<=zm(1)
    dummy_exceed=-1;
    na=1;
    nb=1;
    p=1;
    v1=v0;
else
    dummy_exceed=0;
    na=1+floor((mu-zm(1))/(zm(2)-zm(1)));
    nb=na+1;
    zax=zm(na);
    zbx=zm(nb);
    p=(zbx-mu)/(zbx-zax);
    v1=v0+p*(1-p)*(zbx-zax)^2;
end
```

E.3 Function var_norm

```
% var_norm
% This code is written to normalize unconditional variance of components
% of a VAR(1) process: y'=Ay+epsilon.
%   INPUT: ve - covariance matrix of the error term. This is a diagonal
%               matrix where the i-th diagonal element is var(epsilon_i).
%          A  - the coef. matrix.
%   OUTPUT: vynew - cov. matrix of normalized y
%          vyold - initial cov. matrix of y
%          venew - cov. matrix of the new error term
%          Anew  - the new coef. matrix
function [Anew, vynew, vyold, venew]=var_norm(A, ve)
dif=100;
temp=size(A);
nx=temp(1,1);
V0=zeros(nx,nx);
while dif>0.00000000001
    V=A*V0*A'+ve;
    dif=max(max(V-V0));
    V0=V;
end
vyold=V0;
venew=zeros(nx,nx);
Anew=zeros(nx,nx);
for i=1:nx
    venew(i,i)=ve(i,i)/vyold(i,i);
    for j=1:nx
        Anew(i,j)=A(i,j)*sqrt(vyold(j,j))/sqrt(vyold(i,i));
    end
end
vynew=zeros(nx,nx);
for i=1:nx
    for j=1:nx
        vynew(i,j)=vyold(i,j)/(sqrt(vyold(i,i))*sqrt(vyold(j,j)) );
    end
end
```

E.4 Function rouwen

```
% rouwen
% Rouwenhorst's method (1995) to approximate an AR(1) process using
% a  finite state Markov process.
% For details, see Rouwenhorst, G., 1995: Asset pricing  implications of
% equilibrium business cycle models, in Thomas Cooley (ed.), Frontiers of
% Business Cycle Research, Princeton University Press, Princeton, NJ.
%
% Suppose we need to approximate the following AR(1) process:
%                   y'=rho_Rouw*y+e
% where abs(rho_Rouw)<1, sig_uncond=std(e)/sqrt(1-rho_Rouw^2) and
% mu_uncond denotes E(y), the unconditional mean of y. Let n_R be the
% number of grid points. n_R must be a positive integer greater than one.
%
% [P_Rouw, z_Rouw] = rouwen(rho_Rouw, mu_uncond, sig_uncond, n_R) returns
% the discrete state space of n_R grid points for y, z_Rouw, and
% the transition matrix P_Rouw.
%
function [P_Rouw, z_Rouw] = rouwen(rho_Rouw, mu_uncond, sig_uncond, n_R)
% CHECK IF abs(rho)<=1
if abs(rho_Rouw)>1
    error('Persitence, rho, must be less than one in absolute value.');
end
% CHECK IF n_R IS AN INTEGER GREATER THAN ONE.
if n_R <1.50001 %| mod(n_R,1)~=0
    error('n_R has to be an integer greater than one.');
end
% CHECK IF n_R IS AN INTEGER.
if mod(n_R,1)~=0
    warning('the number of the grid points is not an integer.')
    warning('The method rounded n_R to its nearest integer.')
    n_R=round(n_R);
    disp('n_R=');
    disp(n_R);
end
```

```
% GRIDS
step_R = sig_uncond*sqrt(n_R - 1);
z_Rouw=[-1:2/(n_R-1):1]';
z_Rouw=mu_uncond+step_R*z_Rouw;

% CONSTRUCTION OF THE TRANSITION PROBABILITY MATRIX
p=(rho_Rouw + 1)/2;
q=p;
P_Rouw=[ p  (1-p);
        (1-q) q];
for i_R=2:n_R-1
    a1R=[P_Rouw zeros(i_R, 1); zeros(1, i_R+1)];
    a2R=[zeros(i_R, 1) P_Rouw; zeros(1, i_R+1)];
    a3R=[zeros(1,i_R+1); P_Rouw zeros(i_R,1)];
    a4R=[zeros(1,i_R+1); zeros(i_R,1) P_Rouw];
    P_Rouw=p*a1R+(1-p)*a2R+(1-q)*a3R+q*a4R;
    P_Rouw(2:i_R, :) = P_Rouw(2:i_R, :)/2;
end
P_Rouw=P_Rouw';
for i_R = 1:n_R
    P_Rouw(:,i_R) = P_Rouw(:,i_R)/sum(P_Rouw(:,i_R));
end
```

E.5 Function bigPPP

```
% bigPPP
% This function is used by the main code var_Markov_MM.
function PPP = bigPPP(pmatxxx,n)
PPP=zeros(n^2,n^2);
ix2=0;
for i1=1:n
    for i2=1:n
        ix2=ix2+1;
        for i3=1:n
            for i4=1:n
                ix1=(i3-1)*n+i4;
```

```
                PPP(ix1,ix2)=pmatxxx(1,i1,i2,i3)*pmatxxx(2,i1,i2,i4);
            end
        end
    end
end
for i = 1:n*n
    PPP(:,i) = PPP(:,i) / sum(PPP(:,i));
end
```

References

Elliott G. 1998. On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* **66**: 149–158.

Floden M. 2008. A note on the accuracy of Markov-chain approximations to highly persistent AR(1) processes. *Economics Letters* **99**: 516–520.

Gospodinov N, Maynard A, Pesavento E. 2011. Sensitivity of impulse responses to small low-frequency comovements: Reconciling the evidence on the effects of technology shocks. *Journal of Business and Economic Statistics* **29**: 455–467.

Kopecky KA, Suen RM. 2010. Finite state Markov-chain approximations to highly persistent processes. *Review of Economic Dynamics* **13**: 701–714.

Phillips PCB. 1987. Towards a unified asymptotic theory for autoregression. *Biometrika* **74**: 535–547.

Tauchen G. 1986. Finite state Markov-chain approximations to univariate and vector autoregressions. *Economics Letters* **20**: 177–181.

Tauchen G, Hussey R. 1991. Quadrature-based methods for obtaining approximate solutions to linear asset pricing models. *Econometrica* **59**: 371–396.