View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Munich Personal RePEc Archive

Discretization of highly persistent correlated AR(1) shocks

Lkhagvasuren, Damba and Galindev, Ragchaasuren
Concordia University, Queens University Belfast

23. November 2008

Online at http://mpra.ub.uni-muenchen.de/22523/
MPRA Paper No. 22523, posted 06. May 2010 / 06:42

Department of Economics

Working Paper Series

Discretization of Highly-Persistent Correlated AR(1) Shocks

08-012 | Damba Lkhagvasuren
Concordia University

Ragchaasuren Galindev
Queens University Belfast

Department of Economics, 1455 De Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8
Tel 514–848–2424 # 3900 · Fax 514–848–4536 · econ@alcor.concordia.ca · alcor.concordia.ca/~econ/repec

Discretization of Highly-Persistent Correlated AR(1) Shocks

Ragchaasuren Galindev* Queens University Belfast

Damba Lhkagvasuren† Concordia University

Abstract

The finite state Markov-chain approximation methods developed by Tauchen (1986) and Tauchen and Hussey (1991) are widely used in economics, finance and econometrics to solve functional equations in which state variables follow autoregressive processes. For highly persistent processes, the methods require a large number of discrete values for the state variables to produce close approximations which leads to an undesirable reduction in computational speed, especially in a multivariate case. This paper proposes an alternative method of discretizing multivariate autoregressive processes. This method can be treated as an extension of Rouwenhorsts (1995) method which, according to our finding, outperforms the existing methods in the scalar case for highly persistent processes.The new method works well as an approximation that is much more robust to the number of discrete values for a wide range of the parameter space.

Keywords: Finite State Markov-Chain Approximation, Discretization of Multivariate Autoregressive Processes, Transition Matrix, Numerical Methods, Value Function Iteration

*25 University Square, Queens University Management School, Belfast, BT7 1NN, UK; e-mail: g.ragchaasuren@qub.ac.uk.

†Corresponding author. Department of Economics, Concordia University, 1455 Maisonneuve Blvd. W, Montreal, Canada, H3G 1M8; e-mail: dlkhagva@alcor.concordia.ca.

1 Introduction

The finite state Markov-Chain approximation methods developed by Tauchen (1986) and Tauchen and Hussey (1991) are widely used in economics, finance and econometrics in solving for functional equations where state variables follow autoregressive processes. The methods choose discrete values for the state variables and construct transition probabilities so that the characteristics of the generated process mimic those of the underlying process. The accuracy of the approximation generated by these methods normally depends on the number of discrete values or grids for the state variables, called the fineness of the state spaces, and the persistence of the underlying process. According to Tauchen (1986), Tauchen and Hussey (1991), Zhang (2006) and Flodén (2008), the methods perform poorly for a process whose persistence is close to unity when the state space is moderately refined and hence require a finer state space to achieve a more accurate approximation. However, gaining a closer approximation at the cost of a finer state space may not always work, especially in a multivariate case.

This paper proposes a new method to approximate a particular multivariate autoregressive process, which is referred to as cross-correlated AR(1) shocks. Using appropriate transformations, any vector autoregressive processes can be converted into the process under consideration. The idea behind this method is to decompose the underlying process (carefully while maintaining its basic characteristics) into a set of AR(1) schemes, some of which are independent and the others are perfectly correlated with the independent ones in terms of their error terms. By virtue of the perfectly correlated error terms, the method amounts to constructing transition probabilities for each of the independent AR(1) processes and then generating the other AR(1) processes from the error terms of the independent processes. Using methods that work well in the scalar case, the independent AR(1) processes are accurately approximated. The new method generates accurate approximations for a wide range of the parameter space, without requiring a large number of grid points for the state variables.

The independent AR(1) processes under the new method can be approximated by

existing methods in the literature for the scalar case. As another contribution of the paper, we compare and contrast the numerical accuracy of these methods. Flodén (2008) examines the performance of the methods of Tauchen (1986), Tauchen and Hussey (1991) and Adda and Cooper (2003). Based on a poor performance of these three methods for highly persistent processes, Flodén modifies Tauchen and Hussey's method and obtains better results for a certain range of the parameter space. In addition to those in Flodén (2008), we include Rouwenhorst's (1995) method in our exercise which considers equispaced discrete values for the state variable and builds the probability transition matrix analytically. The persistence of the process we consider contains values that are sometimes significantly larger than those in Flodén (2008). We find that Rouwenhorst's method outperforms the others for highly persistent processes in the sense that the accuracy of its approximations are robust to the number of grids for the state variable. In general, Tauchen's method tends to overshoot their targets while those of Tauchen and Hussey and Adda and Cooper undershoot when the state space is not sufficiently fine. Moreover, we observe that some of the results in Flodén (2008) are reversed when the process is more persistent than the one he considered. Specifically, as the degree of persistence gets closer to unity, the original version of Tauchen and Hussey's method is able to generate some data which vary over time while Flodén's version of the method cannot.

In the scalar case, more accurate approximations can be achieved without increasing the number of grids for the state variable with all the methods except for Rouwenhorst's. One can use the monotonic relationship between targets and approximations - a one-to-one mapping - in the cases of both overshooting and undershooting. For example, when aiming for the persistence of a process with Tauchen's method, experiment with values smaller than the target and choose the one that yields the closest approximation; or experiment with higher values than the target for the methods that undershoot. However, in the multivariate case, it is difficult to establish the one-to-one mapping between the simulated and targeted parameters as one must experiment with

many different coefficients as well as the covariance matrix of the error terms.

The new method can be treated as a multivariate extension of the approximation methods which can work well in the scalar case. Rouwenhorst's method has not been extended to the multivariate case. Therefore, our method can be considered a multivariate extension of Rouwenhorst's method. Another interesting feature of the new method is that instead of applying one method to all the independent AR(1) processes in consideration, one can indeed mix different methods depending on the persistence of the individual processes. For instance, we can use the Tauchen and Hussey (1991) and Rouwenhorst (1995) methods simultaneously (with a moderate-sized state space) by applying the former to the AR(1) processes with sufficiently low degrees of persistence and the latter to highly persistent ones. The rationale of using Tauchen and Hussey's method for low persistent processes is that its approximations of the higher-order moments of the underlying process tend to be slightly more accurate than those of Rouwenhorst's method.

The paper is organized as follows. Section 2.1 shows the shortcoming of the existing methods through Tauchen's method.[1] Section 2.2 discusses the new method and its results in comparison with those in Section 2.1. Section 2.3 demonstrates how to use the new method to approximate VAR(1) processes. Section 3 applies both Tauchen's and the new methods to solve a functional equation of a simplified version of the Mortensen and Pissarides model and compare the results. Finally, Section 4 summarizes the conclusions of the paper.

[1] Considering a different method such as Tauchen and Hussey (1991) or a vector extension of Adda and Cooper (2003) is inconsequential for our purposes as all these methods perform poorly in the case of highly persistent uncorrelated AR(1) shocks, the special case of our multivariate autoregression.

2 Model

We consider the following multivariate autoregressive process:

$$\begin{array}{ccccc} x_{1,t} & = & \rho_1 x_{1,t-1} & + & \varepsilon_{1,t} \\ x_{2,t} & = & \rho_2 x_{2,t-1} & + & \varepsilon_{2,t} \\ \vdots & & \vdots & & \vdots \\ x_{n,t} & = & \rho_n x_{n,t-1} & + & \varepsilon_{n,t} \end{array} \tag{1}$$

where $|\rho_i| < 1$ for all $i \in \{1, 2, ..., n\}$, and the innovations, $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, ..., \varepsilon_{n,t})^T$, follow a multivariate normal distribution, $\varepsilon_t \sim \mathrm{N}(0, \Omega)$ with Ω being an $n \times n$ positive definite matrix. It is assumed that ε_t is serially uncorrelated. Given the above specifications, the process in (1) is referred to as cross-correlated AR(1) shocks for the rest of the paper. Using appropriate transformations, any vector autoregressive process can be converted into this process.

Before outlining the new method, we first discuss the disadvantage of the existing methods used in approximating the process in (1). We consider Tauchen's (1986) method as representative as they all perform poorly in the case of highly persistent uncorrelated AR(1) shocks which is a special case of (1).

2.1 Tauchen's method

The method developed in Tauchen (1986) is originally designed to approximate vector autoregressions with uncorrelated error terms. Since the elements of ε_t are cross-correlated, one must convert the process in (1) into Tauchen's form. For this purpose, let us consider the decomposition $\varepsilon_t = Ce_t$ where $e_t = (e_{1,t}, e_{2,t}, ..., e_{n,t})^T$ is an $n \times 1$ vector of white noise processes whose elements e_{it} are mutually independent with the standard normal distribution, $e_{it} \sim \mathrm{N}(0, 1)$ for all i, and C is the lower triangular matrix obtained from the Cholesky decomposition of Ω, $CC^T = \Omega$. Also, let R denote an n-dimensional diagonal matrix whose i-th diagonal entry is ρ_i. Then, we can rewrite

(1) as follows:

$$x_t = Rx_{t-1} + Ce_t. \tag{2}$$

Multiplying the both sides of (2) by C^{-1} and rearranging the outcome yields:[2]

$$y_t = Ay_{t-1} + e_t \tag{3}$$

where $y_t = C^{-1}x_t$ and $A = C^{-1}RC$. The expression in (3) is a VAR(1) process with uncorrelated error terms.[3] We can therefore apply Tauchen's method to it. First, using the grid points and the associated transition matrix, we simulate time series for y_t for τ time periods.[4] Let $\{\hat{y}_t\}_{t=1}^{\tau}$ denote the simulated time series. We then obtain the corresponding time series for x_t, $\{\hat{x}_t\}_{t=1}^{\tau}$, using the relation, $x_t = Cy_t$. The accuracy of the approximation can then be examined by estimating the key parameters of the initial process in (1). Following Tauchen (1986), we focus on the second order moments which are ρ_i and $\text{cov}(x_i, x_{i'})$ for all i and i'.

To evaluate the performance of the method for a highly persistent process, we consider the following set of parameter specifications: $n = 2$, $\sigma_{x_1}^2 = \sigma_{x_2}^2 = 1$, the variances of x_1 and x_2, $\rho_2 = 0.99$ and $\gamma \equiv \text{corr}(\varepsilon_{1,t}, \varepsilon_{2,t}) = 0.9$, but ρ_1 ranges from 0.5 to 0.9999. Given the persistence parameters, ρ_1 and ρ_2, and the correlation of the error terms, γ, we have $\alpha \equiv \text{corr}(x_{1,t}, x_{2,t}) = \gamma\sqrt{(1-\rho_1^2)(1-\rho_2^2)}/(1-\rho_1\rho_2)$. As in Tauchen (1986), we initially set $N_1 = N_2 = 9$, the number of discrete values that $y_{1,t}$ and $y_{2,t}$

[2]Under the assumption that Ω is a positive definite and symmetric matrix, C is invertible. Considering other decompositions that represent ε_t as a linear combination of i.i.d. normal random variables would not affect the main conclusions of the paper.

[3]It is straightforward to extend the method to a case with correlated error terms at the expense of multidimensional integration. This type of exercise is done by Knotek and Terry (2008). Nevertheless, the problem with highly persistent shocks still remains in their approximation. A simple way to see this is to realize that Tauchen's method and Knotek and Terry's version of the method deliver exactly the same approximation when applied to a VAR with uncorrelated error terms. Alternatively or more formally, one can see our analytical results in Appendix 1 which show that Tauchen's method performs poorly for highly persistent shocks as it calculates the transition matrix using the probability density function of the error terms. Since Knotek and Terry's version calculates the transition matrix the same way, the issue with highly persistent shocks remains in their approximation.

[4]When we simulate a particular time series, we draw the initial value from its unconditional distribution randomly. After simulating the time series, we discard the first one-tenth of the time periods before we estimate the parameters. Computer codes used in this paper are available upon request.

take on respectively from an interval, $[-3\sigma_{y_i}, 3\sigma_{y_i}]$ where σ_{y_i} is the standard deviation of y_i for $i = 1, 2$. We also consider two other cases in which the state space is much finer: $N_1 = N_2 = 19$ and $N_1 = N_2 = 49$.

Having generated $\{\hat{x}_{1,t}\}_{t=1}^{\tau}$ and $\{\hat{x}_{2,t}\}_{t=1}^{\tau}$ for $\tau = 500,000$ for a simulation, the parameters ρ_1, ρ_2, α, σ_{x_1} and σ_{x_2} are estimated. We repeat the same simulation 50 times before calculating the summary results displayed in Tables 1A and 1B. The former shows the mean of the estimated parameters relative to their targets while the latter shows the root mean squared error (RMSE) relative to their true values. However, to compare high persistent levels using fewer digits, we present our results on persistence in terms of $-\lg(1-\hat{\rho}_i)$ using the estimated persistence, $\hat{\rho}_i$ for $i = 1, 2$. The numbers closer to unity in Table 1A and zero in Table 1B imply better approximations. When the number of grids for the state variables are not sufficient, the approximations become less precise as (x_1, x_2) become more persistent. The reason is as follows.

First, higher persistence of x series means higher persistence of y series.[5] Second, given the linear transformation $x_t = Cy_t$, the quality of the approximation of x depends on that of y. Since Tauchen's method performs poorly in highly persistent cases,[6] the approximation of x will be less accurate. In Appendix 1, we study analytically why Tauchen's method performs poorly in highly persistent shocks. Our finding is that as persistence increases, the probability that the process switches from one state to any other state converges to zero much faster than it should. As a consequence, the generated time series exhibits much more persistence than the original continuous process.

The results appear to be much better in the cases where $N_1 = N_2 = 19$ and $N_1 = N_2 = 49$. However, such improvements come at the cost of very large transition

[5] In this particular case with $n = 2$, transforming (1) into (3) as outlined above yields the following VAR(1):

$$\begin{aligned} y_{1,t} &= \rho_1 y_{1,t-1} + e_{1,t} \\ y_{2,t} &= \tfrac{\gamma}{\sqrt{1-\gamma^2}}(\rho_2 - \rho_1) y_{1,t-1} + \rho_2 y_{2,t-1} + e_{2,t} \end{aligned}$$

where e_1 and e_2 are uncorrelated white noise processes. Therefore, persistence of y_1 and y_2 increases with that of the x series, at least in the absolute term.

[6] See Tauchen (1986), Tauchen and Hussey (1991), Zhang (2006) and Flodén (2008).

matrices. For instance, when $N_1 = N_2 = 9$, the size of the probability transition matrix is 81×81 and when $N_1 = N_2 = 49$, it becomes 2401×2401, etc. More importantly, Appendix 1 shows that no matter how large the number of grid points is there always exists a persistence level where Tauchen's method performs poorly. The situation becomes even worse as the dimension of the autoregressive process increases.

In summary, for highly persistent processes, Tauchen's method requires large transition matrices for which some computer memories may not be sufficient. An alternative would be to choose the parameters used in the approximation to minimize the distance between targeted and estimated parameters. This will, however, create serious computational issues. First, we have to simulate the model for a large number of periods and measure all the relevant parameters at every step of the minimization procedure. Second, the multi-dimensional minimization problem will become increasingly difficult as the number of variables increases. Third, depending on the minimization procedures, the resulting approximations may be very different from each other. The reason is that under Tauchen's method, changes in certain parameters have a non-monotonic impact on estimated parameters when it should not. For example, as we see in Figure 2, an increase in ρ_1 has a non-monotonic impact on $\hat{\rho}_2$. This means that in certain cases we may end up with different sets of estimated parameters for the same process.

2.2 New method

Having seen the shortcoming of the existing methods through Tauchen's method, we now discuss a possible solution - a new method. After outlining the new approximation method for the process in (1), we apply it to the same example considered in the previous section and contrast the estimated parameters to their targets. Then we discuss two special, yet very useful, cases of (1) for which the new method becomes even more straightforward.

The idea of the new method is to decompose the underlying process (carefully while maintaining its characteristics) into a set of AR(1) schemes: some are independent and

the others are perfectly correlated with the independent ones in terms of their error terms. Given the perfect correlation between the error terms, the method approximates only the independent AR(1) processes and uses their error terms to derive the others. Using the methods that work well in the scalar case, the independent AR(1) processes are accurately approximated.

Let $c_{i,j}$ denote the (i,j)-th entry of the lower triangular matrix C. Then, for any i, the process (1) can be decomposed as

$$x_{i,t} = \rho_i x_{i,t-1} + \sum_{j \leq i} c_{i,j} e_{j,t}. \tag{4}$$

Being a stationary process, $x_{i,t}$ in (4) can be rewritten as functions of only the innovations $e_{j,t}$ for all t as

$$\begin{aligned} x_{i,t} &= \sum_{j \leq i} c_{i,j} e_{j,t} + \rho_i \sum_{j \leq i} c_{i,j} e_{j,t-1} + \rho_i^2 \sum_{j \leq i} c_{i,j} e_{j,t-2} + ... \\ &= \sum_{j \leq i} c_{i,j} (e_{j,t} + \rho_i e_{j,t-1} + \rho_i^2 e_{j,t-2} + ...). \end{aligned} \tag{5}$$

According to (5), each x_{it} can be represented as a weighted sum of i different AR(1) processes with the common persistence ρ_i but with different innovations, $(e_1, e_2, ..., e_i)$:

$$x_{i,t} = c_{i,1} u_{i,1,t} + c_{i,2} u_{i,2,t} + ... + c_{i,i} u_{i,i,t} \tag{6}$$

where $u_{i,j}$ for $j \leq i \leq n$ are determined by the following schemes:

$$u_{i,j,t} = \rho_i u_{i,j,t-1} + e_{j,t}. \tag{7}$$

According to (7), each $u_{i,j}$ is perfectly correlated with $u_{j,j}$ for $j < i$ in terms of e_j. For example, $u_{2,1}$ (and $u_{i,1}$ for $3 \leq i \leq n$) is correlated with $u_{1,1}$ as both have a common error term e_1 - i.e., $u_{1,1,t} = \rho_1 u_{1,1,t-1} + e_{1,t.}$ and $u_{2,1,t} = \rho_2 u_{2,1,t-1} + e_{1,t}$. Similarly, $u_{3,2}$ (and $u_{i,2}$ for $4 \leq i \leq n$) is correlated with $u_{2,2}$ as $u_{2,2,t} = \rho_2 u_{2,2,t-1} + e_{2,t}$ and

$u_{3,2,t} = \rho_3 u_{3,2,t-1} + e_{2,t}$. The implication is that we need only n independent processes and use their error terms to construct the remaining processes. We let $u_{i,i}$ for $i \leq n$ be the independent ones. Collecting $u_{i,j}$ for $j < i$, we rewrite (6) as follows:

$$x_{i,t} = v_{i,t} + c_{i,i} u_{i,i,t} \tag{8}$$

where

$$\begin{aligned} v_{1,t} &= 0 \\ v_{i,t} &= \rho_i v_{i,t-1} + \sum_{j<i} c_{i,j} e_{j,t} \quad \text{for } 2 \leq i \leq n. \end{aligned}$$

The intuition of this decomposition is that we can discretize only $u_{i,i}$ for $i \leq n$ by using any Markov-chain approximation methods and generate time series for $\{\hat{u}_{i,i,t}\}_{t=1}^{\tau}$. Then, we calculate the associated error terms as

$$\hat{e}_{i,t} = \hat{u}_{i,i,t} - \rho_i \hat{u}_{i,i,t-1}. \tag{9}$$

Given the simulated error terms, $\{\hat{e}_{i,t}\}_{t=0}^{\tau}$, we then construct time series for $\{\hat{v}_{i,t}\}_{t=1}^{\tau}$ in accordance with

$$\hat{v}_{i,t} = \rho_i \hat{v}_{i,t-1} + \sum_{j<i} c_{i,j} \hat{e}_{j,t}. \tag{10}$$

The expression in (10) implies that we know the value of $\hat{v}_{i,t}$ with certainty conditional on $\hat{v}_{i,t-1}$, $\{\hat{u}_{1,1,t-1}, \hat{u}_{2,2,t-1}, ..., \hat{u}_{i,i,t-1}\}$ and $\{\hat{u}_{1,1,t}, \hat{u}_{2,2,t}, ..., \hat{u}_{i,i,t}\}$. Given the time series for $\{\hat{u}_{i,i,t}\}_{t=1}^{\tau}$ and $\{\hat{v}_{i,t}\}_{t=1}^{\tau}$ for $i \leq n$, we can construct time series for $\{\hat{x}_{i,t}\}_{t=1}^{\tau}$ according to (8).

In summary, we have expressed n cross-correlated AR(1) shocks using $2n-1$ single AR(1) processes of which n are independent and the others are linear combinations of the error terms generated from these n independent processes.[7] As a consequence, we need n individual transition matrices (one for each $u_{i,i}$) to construct the transition

[7] In some cases, the number of AR(1) processes after the decomposition is smaller than $2n-1$. Below we show that n equally persistent cross-correlated AR(1) shocks are approximated by n individual AR(1) schemes.

probabilities for n cross-correlated shocks, $\{x_1, x_2, ..., x_n\}$. Under such circumstances, the quality of the simulated data is determined by the quality of the transition matrix built for each $u_{i,i}$.

2.2.1 On the methods used in the scalar case

In this section, we compare the performances of the existing methods used in the scalar case as another contribution of the paper. This exercise provides a rationale for choosing one or a set of different methods that can be used to approximate the independent AR(1) shocks under the new method to produce more precise approximations. For this purpose, we include Tauchen's (1986), the original and Flodén's versions of Tauchen and Hussey's (1991), Adda and Cooper's (2003) and Rouwenhorst's (1995) methods. Using these methods, we approximate an independent AR(1) process with zero mean and unit variance, and its persistence, ρ, ranges from 0.5 to 0.9999. We consider three choices for the number of discrete values: $N = 9$, $N = 19$ and $N = 49$. The process is simulated by each method for 50 times and each simulation contains 10,000,000 periods. Each simulation gives the estimates of the parameters, ρ, the standard deviation, σ, and the kurtosis, κ, of the process which are summarized in Tables 2A and 2B.

The results suggest that Rouwenhorst's method outperforms the other methods in all dimensions when the persistence is high. The reason is that it constructs the transition probabilities so as to match the unconditional mean, variance and the first-order autocorrelation of the underlying process.[8] The other methods, on the other hand, require a much finer state space for highly persistent processes to yield comparable results to the Rouwenhorst method in all three dimensions. When the state space is not sufficiently fine, Tauchen's and Flodén's version of Tauchen and Hussey's methods perform worse than other two. Flodén (2008) finds that his version of Tauchen and Hussey's method is more accurate than the original version of the method for highly persistent processes. Our results suggest that Flodén's conclusion is subjected to the

[8]Later works of Kopecky and Suen (2009) and Lkhagvasuren (2009) calculate other key moments of the AR(1) process generated by the Rouwenhorst method.

number of grids for the state variable when the process is more persistent than what he considered. In other words, when the number of grids is not sufficient, his conclusion is reversed. Specifically, as the degree of persistence gets closer to unity, the original version of Tauchen and Hussey's method can generate some data while Flodén's version of the method cannot.

The results in Tables 2A and 2B suggest that one could also use Tauchen's and Tauchen and Hussey's methods to simulate $u_{i,i,t}$ individually by either considering a sufficiently fine state space or exploiting the one-to-one mapping between the targets and approximations. Still, it would be numerically much more accurate than applying Tauchen's method to vector autoregressions discussed in Section 2.1. Another insightful observation from the results in Tables 2A and 2B is that, to improve the quality of the approximation along other dimensions such as higher order moments of the distribution of the underlying process, one can actually mix different methods to approximate the independent AR(1) shocks. Suppose that there are two shocks to be approximated - one has a sufficiently low degree of persistence and the other has extremely high one. In this case, one could use Tauchen and Hussey's method for the one with low persistence and Rouwenhorst's method for the other. The rationale for using Tauchen and Hussey's method for low persistent ones is that for higher order moments such as kurtosis, it preforms slightly better than Rouwenhorst's method (see Tables 2A and 2B for $\rho = 0.5$).

2.2.2 Example

This section examines the accuracy of the new method for the same process approximated by Tauchen's method in Section 2.1. We first approximate two independent AR(1) shocks, $u_{i,i}$ with persistence ρ_i and $\text{var}(u_{i,i}) = \frac{1}{1-\rho_i^2}$. In doing so, we specify the state spaces for $u_{1,1}$ and $u_{2,2}$ and obtain the corresponding transition probabilities using an accurate method. Given the transition probabilities, we simulate $u_{1,1,t}$ and

$u_{2,2,t}$ over τ time periods. Using the simulated $\{\hat{u}_{1,1,t}\}_{t=1}^{\tau}$, we then generate $\{\hat{v}_{2,t}\}_{t=1}^{\tau}$ as

$$\hat{v}_{2,t} = \rho_2 \hat{v}_{2,t-1} + \gamma\sqrt{1-\rho_2^2}\ (\hat{u}_{1,1,t} - \rho_1 \hat{u}_{1,1,t-1}). \tag{11}$$

Given $\hat{u}_{1,1,t}$, $\hat{u}_{2,2,t}$ and $\hat{v}_{2,t}$, we generate the time series for $\hat{x}_{1,t}$ and $\hat{x}_{2,t}$ according to the following decomposition:

$$x_{1,t} = \sqrt{1-\rho_1^2}\ u_{1,1,t} \tag{12}$$

$$x_{2,t} = v_{2,t} + \sqrt{1-\gamma^2}\sqrt{1-\rho_2^2}\ u_{2,2,t}. \tag{13}$$

Using the properties of $u_{1,1,t}$, $v_{2,t}$ and $u_{2,2,t}$, it is straightforward to show from (12) and (13) that the decomposition is consistent in the sense that it delivers $\sigma_{x_1}^2 = \sigma_{x_2}^2 = 1$ and $\alpha \equiv \text{corr}(x_{1,t}, x_{2,t}) = \frac{\text{cov}(\varepsilon_{1,t}, \varepsilon_{2,t})}{1-\rho_1\rho_2}$.

We choose Rouwenhorst's method to simulate $u_{1,1,t}$ and $u_{2,2,t}$. Again we consider $N_1 = N_2 = 9$, $N_1 = N_2 = 19$ and $N_1 = N_2 = 49$ for $u_{1,1,t}$ and $u_{2,2,t}$ as their discrete values. We simulate the process 50 times and each simulation generates 500,000 observations for $\hat{x}_{1,t}$ and $\hat{x}_{2,t}$ and gives the estimates of the parameters, ρ_1, ρ_2, α, σ_{x_1} and σ_{x_2}. Tables 3A and 3B display the results. The former shows the mean of the estimated parameters relative to their targets while the latter shows their RMSE relative to their true values. As can be seen from the results, the new method works much better than Tauchen's method and the approximations are very accurate even in the cases where Tauchen's method struggles. More importantly, the accuracy of the approximations by the new method is robust to the number of grids for the state variables. This is a highly desirable feature as it does not require large computational memories. Based on the results in Tables 1 and 3, Figure 2 provides a further piece of evidence on the performance of Tauchen's and the new methods where we use $N_1 = N_2 = 9$ for the both methods.

2.2.3 Special cases

The preceding sections deal with a general case where each process in (1) is allowed to have different degrees of persistence. We now consider two very useful special cases for which the new method is even simpler.

Equally-Persistent Shocks. When the underlying process is governed by equally persistent correlated AR(1) shocks - i.e., $\rho_i = \rho$ for all i, the expressions in (7) and (9) imply

$$\hat{u}_{i,j,t} - \hat{u}_{j,j,t} = \rho(\hat{u}_{i,j,t-1} - \hat{u}_{j,j,t-1})$$

for all $j < i$. Since $|\rho| < 1$, it implies that $\hat{u}_{i,j,t} = \hat{u}_{j,j,t}$ for all $j < i$. Consequently, the expression in (6) becomes

$$x_{i,t} = c_{i,1}u_{1,1,t} + c_{i,2}u_{2,2,t} + ... + c_{i,i}u_{i,i,t} \tag{14}$$

where each $u_{i,i,t}$ is an independent AR(1) shock with persistence ρ. In other words, we have expressed n cross-correlated AR(1) shocks as a linear combination of n equally-persistent independent AR(1) processes. If we discretize each process with the same number of grids, we will need to construct only one transition probability matrix of a single AR(1) shock for the entire system.

Equally-Persistent, Symmetric Shocks. Let us consider the following simple autoregressive process:

$$\begin{aligned} x_{1,t} &= \rho x_{1,t-1} + \varepsilon_{1,t} \\ x_{2,t} &= \rho x_{2,t-1} + \varepsilon_{2,t} \end{aligned} \tag{15}$$

where $\text{corr}(\varepsilon_1, \varepsilon_2) = \gamma$ and $\sigma^2_{x_1} = \sigma^2_{x_2} = 1$. The shocks x_1 and x_2 are symmetric in the sense that the moment conditions such as $\text{var}(x_1^2) = \text{var}(x_2^2)$ and $\text{var}(x_1^2 x_2) = \text{var}(x_1 x_2^2)$ hold. In the multivariate case, such symmetry can be easily distorted by discretization methods in the form of asymmetric grid points. Tauchen's method has

this disadvantage.

To show this, we apply Tauchen's method to (15) as outlined in Section 2.1. It follows that $y_{1,t}$ and $y_{2,t}$ follow an independent AR(1) process - i.e., $a_{11} = a_{22} = \rho$ and $a_{12} = a_{21} = 0$. Having discretized $y_{1,t}$ and $y_{2,t}$, we obtain $x_{1,t}$ and $x_{2,t}$ as $x_{1,t} = c_{11}y_{1,t}$ and $x_{2,t} = c_{21}y_{1,t} + c_{22}y_{2,t}$. Since y_1 and y_2 take pre-specified discrete values and the elements of C are real numbers, the grid points of $x_{1,t}$ can be different from those of x_2. The implication is that two shocks that have symmetric moment conditions in their continuous representation can have very different estimated moments due to the asymmetric grid points.

Unlike Tauchen's method, the new method allows us to preserve the underlying symmetry in the multivariate case. To discretize the process in (15), we can decompose x_1 and x_2 using three independent finite state AR(1) processes, u_1, u_2 and u_3, as:

$$x_1 = \sqrt{1 - |\gamma|}u_1 + \sqrt{|\gamma|}u_3$$

$$x_2 = \sqrt{1 - |\gamma|}u_2 + \frac{\gamma}{\sqrt{|\gamma|}}u_3.$$

First of all, if we choose the same state space for each of the three shocks, they will have the same transition matrix. Given the same absolute magnitude of the weights, $\sqrt{|\gamma|}$ and $\gamma/\sqrt{|\gamma|}$, the symmetry is always guaranteed by this decomposition along both grid points and transition probabilities.[9]

In order to support this argument, we consider $\rho = 0$ and $\gamma = 0.5$ for the process in (15). We choose $N = 8$ for Tauchen's method while $N = 4$ for the new method which are reasonable given the persistence of the process. The choice, $\rho = 0$, is deliberate as we want to show that the asymmetry in the simulated grids can arise primarily due to a underlying discretization method. Each method generates 50,000 observations for $\hat{x}_{1,t}$ and $\hat{x}_{2,t}$ - i.e., $\{\hat{x}_{1,t}, \hat{x}_{2,t}\}_{t=1}^{\tau}$ where $\tau = 50,000$ which are sufficient given the

[9]This technique of handling symmetric AR(1) shocks is used in Lkhagvasuren (2008) to simulate and estimate a dynamic stochastic model of internal migration where the correlation of the match-specific productivity shocks are assumed to be symmetric across different labor markets.

persistence of the process. Then we transform them monotonically into time series $\left\{\hat{x}^2_{1,t}, \hat{x}^2_{2,t}, \hat{x}^3_{1,t}, \hat{x}^3_{2,t}, \hat{x}^4_{1,t}, \hat{x}^4_{2,t}\right\}_{t=1}^{\tau}$. Using the standard deviation of each time series, we look at the following three ratios: $\frac{\text{std}(\hat{x}^2_1)}{\text{std}(\hat{x}^2_2)}$, $\frac{\text{std}(\hat{x}^3_1)}{\text{std}(\hat{x}^3_2)}$ and $\frac{\text{std}(\hat{x}^4_1)}{\text{std}(\hat{x}^4_2)}$. Given the underlying symmetry between x_1 and x_2, the true values of these three ratios are all one. We repeat this experiment 50 times. The results are summarized in Table 4 which shows that the new method captures the underlying symmetry much better than Tauchen's method. Since persistence is low, this difference is primarily due to the differences in how the two methods construct their grid points. To make the point clearer, we scatter $\hat{x}_1$ against $\hat{x}_2$ in Figure 3 for both methods. As can be seen, the grid points from new method is symmetric while those from Tauchen's method is not.

Let us now consider three equally-persistent AR(1) shocks with the following symmetry restrictions:

$$\Omega = \begin{pmatrix} 1 & \eta^2 & \eta^2 \\ \eta^2 & 1 & \eta^2 \pm \zeta^2 \\ \eta^2 & \eta^2 \pm \zeta^2 & 1 \end{pmatrix}$$

where $\eta^2 + \zeta^2 < 1$. In this case, we can use the following decomposition:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \eta u_1 \\ \eta u_1 \\ \eta u_1 \end{pmatrix} + \begin{pmatrix} 0 \\ \zeta u_2 \\ \pm\zeta u_2 \end{pmatrix} + \begin{pmatrix} \sqrt{1-\eta^2} u_3 \\ \sqrt{1-\eta^2-\zeta^2} u_4 \\ \sqrt{1-\eta^2-\zeta^2} u_5 \end{pmatrix}$$

where u_i for all i denotes an independent finite state AR(1) process. Analogously, one can choose the appropriate decompositions depending on the nature of the symmetry.

2.3 Approximating a VAR(1) process

It is important to note that Tauchen (1986) is, in fact, not written to approximate the cross-correlated AR(1) shocks, but rather designed to discretize a VAR(1) with uncorrelated error terms. The new method, introduced in the previous section, can also be applied to such a process. In this section, we suggest a procedure that converts

a VAR(1) process with uncorrelated error terms into a cross-correlated AR(1) process as in (1). Using this procedure, we apply the new method to some VAR(1) processes including the one considered in Tauchen (1986) and compare the results of the two methods.

Example 1. Tauchen (1986) considers a VAR(1) process of two variables in the form of (3) that is characterized by

$$A^1 = \begin{bmatrix} 0.7 & 0.3 \\ 0.2 & 0.5 \end{bmatrix} \tag{16}$$

and $\sigma^2_{e_1} = \sigma^2_{e_2} = 0.1$. Given this information set, the variance-covariance matrix of y_t is calculated as:

$$\Sigma^1 = \begin{bmatrix} 0.332 & 0.126 \\ 0.126 & 0.185 \end{bmatrix}. \tag{17}$$

First we apply Tauchen's method to this process. As in Tauchen (1986), we set $N_1 = N_2 = 9$, the number of discrete values that $y_{1,t}$ and $y_{2,t}$ take on respectively from an interval, $[-3\sigma_{y_i}, 3\sigma_{y_i}]$ for $i = 1, 2$. The method generates 5,000,000 observations for $\hat{y}_{1,t}$ and $\hat{y}_{2,t}$. The estimation based on the induced representation $\hat{y}_t = \hat{A}\hat{y}_{t-1} + \hat{e}_t$ reveals the following results:

$$\hat{A}^1_{\text{Tauchen}} = \begin{bmatrix} 0.699 & 0.298 \\ 0.200 & 0.497 \end{bmatrix}, \quad \hat{\Sigma}^1_{\text{Tauchen}} = \begin{bmatrix} 0.372 & 0.138 \\ 0.138 & 0.200 \end{bmatrix}$$

which are very close to those reported in Tauchen (1986), showing the accuracy of the method in the approximation of $\hat{A}^1_{\text{Tauchen}}$ to A^1. The approximation of $\hat{\Sigma}^1_{\text{Tauchen}}$ to Σ^1 is, on the other hand, not so accurate and it needs a finer state space. The state spaces determined by $N_1 = N_2 = 19$, for example, make $\hat{\Sigma}^1_{\text{Tauchen}}$ very close to Σ^1, without changing $\hat{A}^1_{\text{Tauchen}}$ significantly. This suggests that if one cares more about the accuracy of $\hat{\Sigma}^1_{\text{Tauchen}}$ to Σ^1, more refined state spaces are required.

To apply the new method to this process, we convert the VAR(1) in (3) into cross-correlated AR(1) shocks in (1). Given that A is diagonalizable, $A = VRV^{-1}$ where R is an $n \times n$ diagonal matrix and its diagonal elements, ρ_i for $i \in \{1, 2, ..., n\}$, are the eigenvalues of A and, V is an $n \times n$ matrix and its columns are the eigenvectors associated with eigenvalues ρ_1 to ρ_n. Thus the VAR(1) process in (3) can be rewritten as $y_t = VRV^{-1}y_{t-1}+e_t$. Multiplying the both sides by V^{-1} and rearranging the outcome yields the expression in (1) where we define $x_t = V^{-1}y_t$ and $\varepsilon_t = V^{-1}e_t$.

Given the procedure, A^1 in (17) and $\sigma^2_{e_1} = \sigma^2_{e_2} = 0.1$ imply

$$\begin{aligned} x_{1,t} &= 0.865x_{1,t-1} + \varepsilon_{1,t} \\ x_{2,t} &= 0.335x_{2,t-1} + \varepsilon_{2,t} \end{aligned}$$

where $\sigma^2_{x_1} = 0.41$, $\sigma^2_{x_2} = 0.117$, $\alpha \equiv \text{corr}(x_{1,t}, x_{2,t}) = 0.124$ and $\gamma \equiv \text{corr}(\varepsilon_{1,t}, \varepsilon_{2,t}) = 0.186$. Tauchen's VAR(1) process is now represented by the cross-correlated AR(1) shocks. Therefore, we now approximate the process by the new method. We set $N_1 = N_2 = 9$, the number of discrete values that $u_{1,1,t}$ and $u_{2,2,t}$ in (12) and (13) take on respectively and obtain 5,000,000 observations for $\hat{x}_{1,t}$ and $\hat{x}_{2,t}$. Using $y_t = Vx_t$, we convert the time series for $\hat{x}_{1,t}$ and $\hat{x}_{2,t}$ into those of $\hat{y}_{1,t}$ and $\hat{y}_{2,t}$. Estimating the induced representation $\hat{y}_t = \hat{A}^1\hat{y}_{t-1} + \hat{e}_t$ yields the following results:

$$\hat{A}^1_{\text{New}} = \begin{bmatrix} 0.6997 & 0.2995 \\ 0.1984 & 0.5014 \end{bmatrix}, \quad \hat{\Sigma}^1_{\text{New}} = \begin{bmatrix} 0.3296 & 0.1243 \\ 0.1243 & 0.1845 \end{bmatrix}.$$

As can be seen from the results, the new method gives more accurate approximations than Tauchen's method.

Example 2. In Example 1, the approximation $\hat{A}^1_{\text{Tauchen}}$ to the target A^1 is very accurate as the underlying process has a sufficiently low degree of persistence. Let us now apply the both methods to a process whose persistence is higher. Suppose that

the VAR(1) process of two variables is characterized by:

$$A^2 = \begin{bmatrix} 0.952 & 0.05 \\ 0.052 & 0.94 \end{bmatrix}, \quad \Sigma^2 = \begin{bmatrix} 10.574 & 9.126 \\ 9.126 & 8.77 \end{bmatrix} \tag{18}$$

where $\sigma^2_{\epsilon_1} = \sigma^2_{\epsilon_2} = 0.1$. The results from both methods are as follows:

$$\hat{A}^2_{\text{Tauchen}} = \begin{bmatrix} 0.9999 & 0.0001 \\ 0.0007 & 0.9991 \end{bmatrix}, \quad \hat{\Sigma}^2_{\text{Tauchen}} = \begin{bmatrix} 8.359 & 6.823 \\ 6.823 & 7.1 \end{bmatrix}$$

and

$$\hat{A}^2_{\text{New}} = \begin{bmatrix} 0.952 & 0.05 \\ 0.051 & 0.94 \end{bmatrix}, \quad \hat{\Sigma}^2_{\text{New}} = \begin{bmatrix} 10.3557 & 8.9229 \\ 8.9229 & 8.5798 \end{bmatrix}.$$

In this case, Tauchen's method yields large inaccuracies when compared to the performance of the new method in all dimensions. Moreover, if we consider $a_{11} = 0.953$ while keeping everything else equal, and simulate the process with Tauchen's method, the diagonal elements of $\hat{A}^1_{\text{Tauchen}}$ will be unity and the elements of $\hat{\Sigma}^1_{\text{Tauchen}}$ will therefore be nowhere near the target.

3 On solving functional equations

By construction, the simulated values $\{\hat{v}_{i,t}\}_{t=1}^{\tau}$ are not restricted to belong to a pre-specified finite state space. The explanation is the following. Let N_i be the number of grid points used to approximate each independent $u_{i,i}$ and M_i be the number of pre-specified grid points for v_i for all i. Now set the values of $u_{i,j}$ for all $j \leq i \leq n$ at some $u^1_{i,j}$ at time 1 - i.e., $\hat{u}_{i,j,1} = \hat{u}^1_{i,j}$. At any t, since the approximation of $\hat{u}_{i,i,t}$ takes on one of N_i different values, the error term $\hat{e}_{i,t} = \hat{u}_{i,i,t} - \rho\hat{u}_{i,i,t-1}$ takes on one of N_i^2 possible values. Given the law of motion in (9), the number of values that $\hat{u}_{i,j,1}$ can take on in period 1 will be N_j^2. But in period 2, it will rise to N_j^4 and so forth. In fact, the number of values that $\hat{u}_{i,j,t}$ for $j < i$ can take on increases exponentially

with t, leading to non-discrete state spaces. Therefore, unless $\rho_i = 0$ or $\rho_i = \rho_j$ for all $j < i$, the number of values that $\hat{v}_{it} = \sum_{j<i} c_{i,j}\hat{u}_{i,j,t}$ can take on will also increase exponentially with t. Consequently, the simulated values of $\hat{v}_{i,t}$ may not be restricted to the finite state space of M_i grid points.[10]

Since the realized values of v-s are not necessarily a finite set, it may suggest that we may need to evaluate the underlying functions over a certain interval when we apply the new method to solving functional equations. However, it should be noted that the distribution of $v_{i,t}$ given $v_{i,t-1}$ and $u_{i,i,t}$ for $1 \leq i \leq n$ is degenerate. Therefore, conditional on $v_{i,t-1}$ and $u_{i,i,t}$ where $1 \leq i \leq n$, one has to evaluate the underlying functions at only one point which is not necessarily one of the pre-specified grid points of v_i. Essentially, what we need for solving functional equations is conditional distributions rather than unconditional ones. Below, using a simple example, we demonstrate that one can effectively apply numerical interpolation to evaluate the underlying functions in such circumstances. Specifically, we solve a simple dynamic model using the method developed here and compare the simulated results with those obtained using Tauchen's method.

3.1 A simple dynamic model

We consider a simplified version of the Mortensen-Pissarides search and matching model (e.g., Mortensen and Pissarides, 1994). Our focus is on the discretization methods and their associated solutions derived from the model. Since we study the model under different persistence levels, some of the parameters we consider do not necessarily have empirical justification.

The economy has an infinite number of firms. Each firm employs at most one worker. The objective of each firm is to maximize the expected discounted value of profits. A firm entering the market incurs a per-period vacancy cost δ while looking

[10]Earlier we showed that for equally persistent cases, i.e $\rho_i = \rho_0$ for all i in (1), all v-s are restricted to belong to a finite state space.

for a worker. Matches are formed randomly at an endogenous rate $q(\theta)$ where θ is the ratio of the aggregate measures of unemployed workers and vacancies, and dissolved at an exogenous rate λ. Per-period profit of a firm in a match is $p - w$ where p is labor productivity and w is the wage rate. We focus on two sources of shocks: the productivity, p, and the separation rate, λ. Specifically, we consider two strictly monotonic functions P and Λ such that $p = P(x_1)$ and $\lambda = \Lambda(x_2)$ where x_1 and x_2 evolve according to (1).

Each period consists of three stages. At the beginning of each period, some of the old matches are dissolved. In the second stage, the new values of p and λ are realized. Given the market condition, (p, λ), a firm decides whether to post a vacancy or not. In the third stage, matches are formed as a result of job search and vacancy posting. To remain focused on our numerical method, we make a simplifying assumption that wage is rigid, i.e. w is constant. The values of a filled job J and a vacancy V are given by

$$J(p, \lambda) = p - w + \beta(1 - \lambda)\mathrm{E}_{p,\lambda} J(p', \lambda') \tag{19}$$

$$V(p, \lambda) = -\delta + \beta\mathrm{E}_{p,\lambda} \left(q(\theta)J(p', \lambda') + (1 - q(\theta))V(p', \lambda')\right) \tag{20}$$

where β is the discount factor and $\mathrm{E}_{p,\lambda}$ is the mathematical expectation conditional on p and λ. Since there is an infinite number of firms, the value of entering the market is zero, i.e. $V(p, \lambda) = 0$ for all p and λ. Therefore, the expression in (20) becomes

$$\delta = \beta q(\theta)\mathrm{E}_{p,\lambda} J(p', \lambda'). \tag{21}$$

3.2 Numerical experiments

Given the firms' entry decision, one can study the extent to which the parameters in (1) affect the vacancy filling rate $q(\theta)$. Generally, the answer to this question is not available in a simple closed form. We approach the question numerically and solve the above functional equations using the value function iteration technique. For this

purpose, we consider the following specifications for P and Λ:

$$\begin{aligned} P(x_1) &= 1 + 0.01x_1 \\ \Lambda(x_2) &= 0.01(1 + \tfrac{2}{\pi}\arctan(\tfrac{x_2}{2})) \end{aligned} \tag{22}$$

where[11] x_1 and x_2 follow (1) with var(x_1)=var(x_2) = 1. We set $w = 0.9$ and $\beta = 0.99$ and experiment with different values for ρ_1, ρ_2 and $\gamma \equiv \text{corr}(\varepsilon_1, \varepsilon_2)$.

Let q_0 and J_0 be the steady state values of $q(\theta)$ and J, respectively. From (21), we obtain $\delta = \beta q_0 J_0$ where $J_0 = \frac{1-w}{1-\beta(1-0.01)}$. Using q_0, J_0 and (21), we derive

$$\frac{q(\theta)}{q_0} = \frac{J_0}{\text{E}_{p,\lambda} J(p', \lambda')}.$$

To evaluate the two methods, we focus on the volatility and serial autocorrelation of $r_t = \frac{q(\theta_t)}{q_0}$: $\text{cv}(r_t) = \frac{\text{std}(r_t)}{\text{mean}(r_t)}$ and $\text{corr}(r_t, r_{t+1})$. The numerical algorithm of solving the problem is as follows:

1. Construct the grid points and transition probabilities for $\{p, \lambda\}$ using those of $\{x_1, x_2\}$.

2. Apply the value function iteration technique for J using (19) until the differences in value functions between two consecutive iterations become less than 10^{-6} at each grid point.

3. Simulate the time series for $\{p_t, \lambda_t\}$ for $\tau = 2,000,000$ periods using the transition probabilities.

4. Given $\{p_t\}_{t=1}^{\tau}$ and $\{\lambda_t\}_{t=1}^{\tau}$ simulate $\{J_t\}_{t=1}^{\tau}$ and then $\{r_t\}_{t=1}^{\tau}$.

In order to approximate (x_1, x_2) with the new method, we generate three AR(1) shocks $(u_{1,1}, u_{2,2}, v_2)$ in which $u_{1,1}$ are $u_{2,2}$ are independent and v_2 is constructed as the error terms of $u_{1,1}$. Let N_1, N_2, and M_2 be the number of grid points used for

[11] This specification guarantees that $0 < \Lambda(x_2) < 1$ for any value of x_2.

discretizing $u_{1,1}$, $u_{2,2}$ and v_2, respectively. We set $N_1 = N_2 = M_2 \equiv \overline{N}$. Similarly we set $N_1 = N_2 \equiv \overline{N}$ when using Tauchen's method. The value function J has to be solved for $\overline{N}^4$ points in Tauchen's method while $\overline{N}^3$ points in the new method. Under the new method, when we evaluate the value function on the values of v_2 that are not one of $\overline{N}$ grid points, we use a linear interpolation technique. Given the degenerate conditional distribution of $\hat{v}_{2,t}$ and the grids for $\hat{u}_{1,1,t}$ and $\hat{v}_{2,t-1}$, there is a finite number of off-grid values of $\hat{v}_{2,t}$. On the other hand, the values obtained by linear interpolation is a weighted sum of the values of the functions on the grid points. Therefore, if we associate each of the finite number off-grid values to the values of the $\overline{N}$ grid points using a matrix constructed form the interpolation weights, evaluation of the conditional expectation in (19) amounts to simple matrix multiplication.[12]

The results are shown in Table 5. First of all, when the persistence is low there is not much difference between the two methods. Second, when the persistence is high, the estimated parameters from Tauchen's method are highly sensitive to the number of grids. Third, the approximation is very stable with the new method even when the persistence is very high. As we increase the number of grids in Tauchen's method, the two parameters are becoming closer to those obtained by the new method. This indicates that Tauchen's method is less robust to the number of grid points than the new method. In this exercise, we deliberately consider relatively lower levels of persistence than those reported in Tables 1 and 3. The obvious reason is that, for higher levels of persistence, Tauchen's method fails to generate data for the numbers of grid points considered here and therefore does not allow us to evaluate the two methods quantitatively.

[12]In Appendix 2, we describe the procedure of iterating J for each method.

4 Conclusion

In this paper, we develop a method which can be used to approximate both cross-correlated continuous AR(1) shocks and VAR(1) processes with uncorrelated error terms. The main idea of the method is to decompose the initial process into a set of AR(1) shocks of which some are purely independent while the rest are perfectly correlated with the independent ones in terms of their error terms. We simulate the independent processes with any methods that can generate accurate approximations. By virtue of the perfect correlation between the error terms, we then generate data for the dependent processes from the simulated error terms of the independent processes. Through this decomposition, the method yields a very accurate approximation to the initial process. The new method has been motivated by the fact that highly persistent vector autoregressions cannot be approximated accurately by the existing methods in the literature when the state spaces are moderate-sized. The paper has considered Tauchen's (1986) method as representative of those methods.

Another contribution of the paper is that it compares and contrasts the accuracy of existing methods in the literature for the scalar case. We include Rouwenhorst's (1995) method in addition to those considered in Flodén (2008), namely Tauchen (1986), different versions of Tauchen and Hussey (1991) and Adda and Cooper (2003) methods. We consider a broader range of persistence levels than Flodén (2008). Our findings suggest that Rouwenhorst's method gives much more accurate approximations than the others for high degrees of persistence. We reach a conclusion opposite to Flodén (2008) that the original Tauchen and Hussey method is better than Flodén's version of the method when the level of persistence is larger than what Flodén considered.

The new method can be understood as a multivariate extension of any methods that can work well in approximating independent AR(1) shocks. For example, the method in Rouwenhorst (1995), to our knowledge, has not been extended to a multivariate case. Our method is one way of extending Rouwenhorst (1995) to vector autoregressions. Moreover, as each independent process in our method is approximated individually,

one can mix different methods to gain a further improvement in higher-order moments. Suppose that one set of the shocks considered follows sufficiently low persistent AR(1) processes, while the other set follows highly persistent AR(1) processes. According to the new method, we can effectively apply Tauchen's or Tauchen and Hussey's methods to the former and Rouwenhorst's method to the latter.

Acknowledgement. We thank Gordon Fisher, Paul Gomme, Paul Klein, Patricia Ledesma, Ernst Shaumburg, Purevdorj Tuvaandorj, Lu Zhang, two anonymous referees and seminar participants at Queens University Belfast and Royal Economic Society Annual Conference in 2009 for useful discussions and comments. All remaining errors are our own. Damba Lkhagvasuren greatly acknowledges support from Social Sciences and Humanities Research Council of Canada (SSHRC) through Concordia University General Research Fund.

Appendix 1: Overshooting

In this appendix, we discuss the overshooting problem generated by Tauchen's method for highly persistent processes. For simplicity, we present our discussion for the scalar case. It is straightforward to extend our results to the multivariate case. Consider the following scalar autoregressive scheme:

$$y_t = \rho y_{t-1} + \varepsilon_t \tag{23}$$

where $0 < \rho < 1$ and ε_t is a white noise process with variance σ_ε^2. Without loss of generality, assume that $\mathrm{E}(\varepsilon_t) = 0$ and normalize the standard deviation of y_t to one so that $\sigma_\varepsilon^2 = 1 - \rho^2$. Since we focus on highly persistent shocks, we set $\rho = 1 - \frac{1}{K}$ where K is a large positive number.

Tauchen's method uses equispaced grid points for y and the transition probabilities are calculated as areas under the probability density function of the error terms ε. Let

$\overline{y}^1 < \overline{y}^2 < ... < \overline{y}^N$ denote the grid points. Let $w = (\overline{y}^2 - \overline{y}^1)/2$ - i.e., $2w$ is the distance between two subsequent points. According to Tauchen's method, the probability that the process switches from state j to any other state is given by

$$Q_j^T = 1 - \text{Prob}(|\varepsilon - \frac{\overline{y}^j}{K}| < w).$$

Let K be large enough that $0 < \frac{\overline{y}^j}{K} < w$ for all j. Then, it is straightforward to show that

$$Q_j^T \leq 1 - \text{Prob}(|\varepsilon| < w) = 2(1 - \Phi(\frac{w}{\sqrt{1-\rho^2}})) < 2\left(1 - \Phi\left(w\sqrt{K/2}\right)\right)$$

for any j where Φ denotes the CDF of the standard normal distribution. The result suggests that as persistence increases or equivalently, as K increases, the probability that the process switches from a particular state to any other state goes to zero. This is not surprising as higher persistence means a higher probability that the current state repeats itself. What is relevant to our discussion is how fast Q_j^T goes to zero as K increases. For this purpose, we consider Rouwenhorst's method discussed in Section 2 as a benchmark. The main reason is that Rouwenhorst's method also uses equispaced grid points and its transition probabilities are constructed so that the persistence of the underlying process is perfectly matched.

Using Rouwenhorst's transition matrix and $\rho = 1 - \frac{1}{K}$, it can be shown that the probability that the current state repeats itself is $(1 - \frac{1}{2K})^N + D_0 \frac{1}{K^2}$ where D_0 is some nonnegative, finite number. Therefore, with Rouwenhorst's method, the probability that the process switches from a particular state to all other states is $Q_j^R = \frac{N}{2K} + \frac{D_1}{K^2}$ where $|D_1| < \infty$.

Comparing Q_j^T and Q_j^R and using l'Hôspital's rule one can show that

$$\lim_{K\to\infty} \frac{Q_j^T}{Q_j^R} < \lim_{K\to\infty} \frac{w}{2\sqrt{\pi}} \frac{1}{\left(\frac{1}{2K^{3/2}} + \frac{2D_1}{K^{5/2}}\right) e^{\frac{w^2 K}{4}}} = 0.$$

This shows that, for any N, the probability that the process switches from one state to any other state decreases exponentially in Tauchen's method relative to that in Rouwenhorst's method as ρ approaches to unity. Therefore, as persistence increases, all the diagonal elements of the transition matrix constructed by Tauchen's method go to unity much faster than that constructed by Rouwenhorst's method. This is why Tauchen's method delivers much higher persistence than targeted and thus sometimes generates no transition at all when ρ is high (See Figure 1). Using this result, it is also straightforward to see that no matter how large N is, there always exists a high persistence level where Tauchen's method performs poorly.

Appendix 2: Value Function Iteration

Tauchen's method

Substituting $x_1 = c_{11}y_1$ and $x_2 = c_{21}y_1 + c_{22}y_2$ into (22), we obtain the following two functions:

$$\tilde{P}_1(y_1, y_2) = P(c_{11}y_1) - w$$

$$\tilde{\Lambda}_1(y_1, y_2) = \beta(1 - \Lambda(c_{21}y_1 + c_{22}y_2)).$$

Let $\overline{y}_i^1 < \overline{y}_i^2 < ... < \overline{y}_i^{\overline{N}}$ denote the grid points for y_i, $i \in \{1, 2\}$. Then, the firm's asset pricing equation can be rewritten in the discrete space as

$$J_1(\overline{y}_1^i, \overline{y}_2^j) = \tilde{P}_1(\overline{y}_1^i, \overline{y}_2^j) + \tilde{\Lambda}_1(\overline{y}_1^i, \overline{y}_2^j) \sum_{i'=1}^{\overline{N}} \sum_{j'=1}^{\overline{N}} J_1(\overline{y}_1^{i'}, \overline{y}_2^{j'}) \Pi(i', j'|i, j)$$

where $\Pi(i', j'|i, j)$ is the probability that the process switches to any state (i', j') conditional on the current state (i, j). The size of the transition probability matrix is $\overline{N}^2 \times \overline{N}^2$.

New method

For brevity, let $u_i = u_{i,i}$ for $i = 1, 2$. Then substituting (8) into (22), we obtain

$$\tilde{P}_2(u_1, u_2, v_2) = P(c_{1,1}u_1) - w$$

$$\tilde{\Lambda}_2(u_1, u_2, v_2) = \beta(1 - \Lambda(v_2 + c_{2,2}u_2)).$$

Let $\{\overline{u}_1^i, \overline{u}_2^i, \overline{v}_2^i\}_{i=1}^{\overline{N}}$ denote the grid points for u_1, u_2 and v_2. Then, the firm's asset pricing equation can be rewritten in the discrete space as:

$$J_2(\overline{u}_1^i, \overline{u}_2^j, \overline{v}_2^k) = \tilde{P}_2(\overline{u}_1^i, \overline{u}_2^j, \overline{v}_2^k) + \tilde{\Lambda}_2(\overline{u}_1^i, \overline{u}_2^j, \overline{v}_2^k) \sum_{i'=1}^{\overline{N}} \sum_{j'=1}^{\overline{N}} J_2(\overline{u}_1^{i'}, \overline{u}_2^{j'}, v_2^{i,i',k}) \Pi_1(i'|i) \Pi_2(j'|j)$$

where $v_2^{i,i',k} = \rho_2 \overline{v}_2^k + c_{2,1}(\overline{u}_1^{i'} - \rho_1 \overline{u}^i)$ and Π_1 and Π_2 denote the transition probabilities of u_1 and u_2 respectively. The size of the transition probability matrices of u_i, $i \in \{1, 2\}$ is $\overline{N} \times \overline{N}$. When $\rho_1 = \rho_2$, $v_2 = c_{2,1}u_1$ and thus there is no need for interpolation.

An alternative specification

We now present an alternative way of using the new method which simplifies its application. Let us denote $d = v_2 - c_{2,1}u_1$. Then we can write

$$\tilde{P}_3(u_1, u_2, d) = P(c_{1,1}u_1) - w$$

$$\tilde{\Lambda}_3(u_1, u_2, d) = \beta(1 - \Lambda(d + c_{2,1}u_1 + c_{2,2}u_2)).$$

Let $\overline{d}^1 < \overline{d}^2 < ... < \overline{d}^{\overline{N}}$ denote the grid points for d. The functional equation, in this case, becomes

$$J_3(\overline{u}_1^i, \overline{u}_2^j, \overline{d}^k) = \tilde{P}_3(\overline{u}_1^i, \overline{u}_2^j, \overline{d}^k) + \tilde{\Lambda}_3(\overline{u}_1^i, \overline{u}_2^j, \overline{d}^k) \sum_{i'=1}^{\overline{N}} \sum_{j'=1}^{\overline{N}} J_3(\overline{u}_1^{i'}, \overline{u}_2^{j'}, d^{i,k}) \Pi_1(i'|i) \Pi_2(j'|j)$$

where $d^{i,k} = \rho_2 \overline{d}^k + c_{2,1} \overline{u}_1^i (\rho_2 - \rho_1)$. As is seen $d^{i,k}$ is determined only by the current values of d and u_1. Therefore, using d instead of v_2 makes the method numerically even simpler by reducing the number of grid points over which the function is interpolated. When $\rho_1 = \rho_2$, $d = 0$ and thus there is no need for interpolation.

References

[1] Adda, J., Cooper, W, C., 2003. Dynamic Economics. MIT Press, Cambridge, MA.

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

[3] Knotek, E, S., Terry, T., 2008. Markov-Chain Approximations of Vector Autoregressions: An Application of General Multivariate-Normal Integration Techniques. Working Paper.

[4] Kopecky, K., Suen, R., 2009. Finite State Markov-Chain Approximations to Highly Persistent Processes. Working Paper.

[5] Lkhagvasuren, D., 2008. Wage Differences Between Movers and Stayers: Implications on Cross Sectional Volatility of Individual Income Processes. Working Paper.

[6] Lkhagvasuren, D., 2009. Key Moments in the Rouwenhorst Method. Mimeo. Concordia University.

[7] Mortensen, D, T., Pissarides, C, A., 1994. Job Creation and Job Destruction in the Theory of Unemployment. Review of Economic Studies 61(3), 397-415.

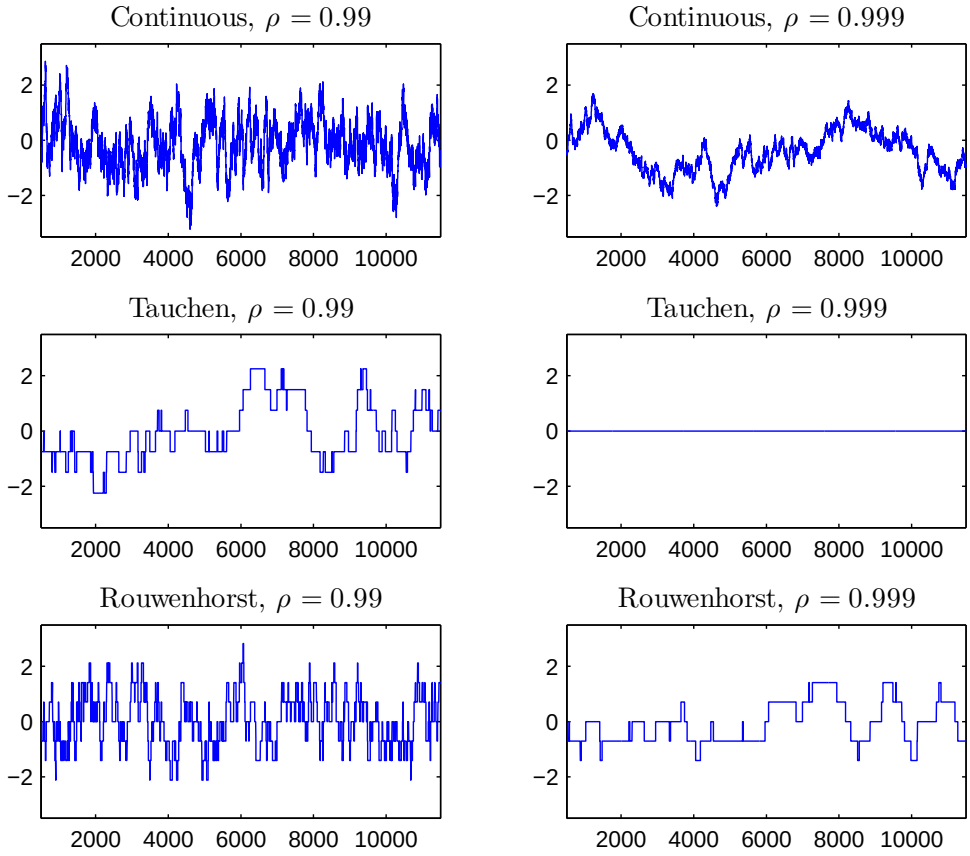
[8] Rouwenhorst, G., 1995. Asset pricing implications of equilibrium business cycle models, in: Cooley, T. (Ed.), Frontiers of Business Cycle Research. Princeton University Press, Princeton, NJ, pp. 294-330.

[9] Tauchen, G., 1986. Finite State Markov-Chain Approximation to Univariate and Vector Autoregression. Economics Letters 20, 177-181.

[10] Tauchen, G., Hussey, R., 1991. Quadrature-Based Methods for Obtaining Approximate Solutions to Linear Asset Pricing Models. Econometrica 59(2), 371-396.

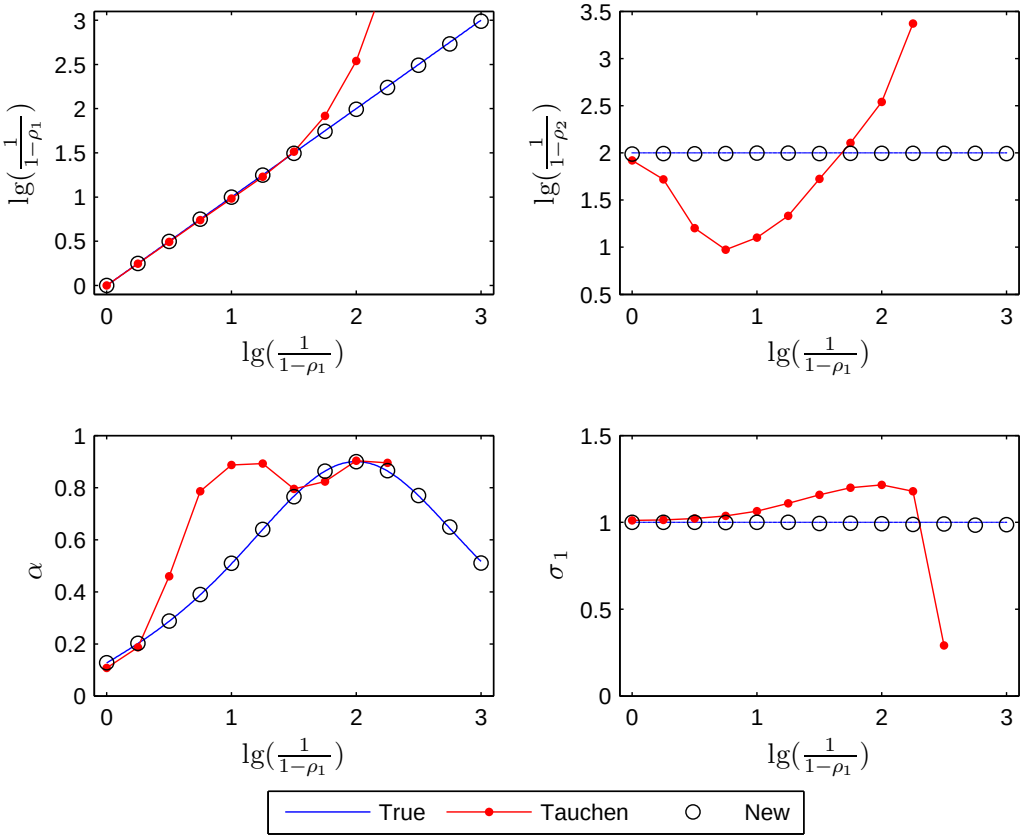
[11] Zhang, L., 2005. Value Premium. Journal of Finance 60, 67-103.

Figure 1. A highly-persistent AR(1) process



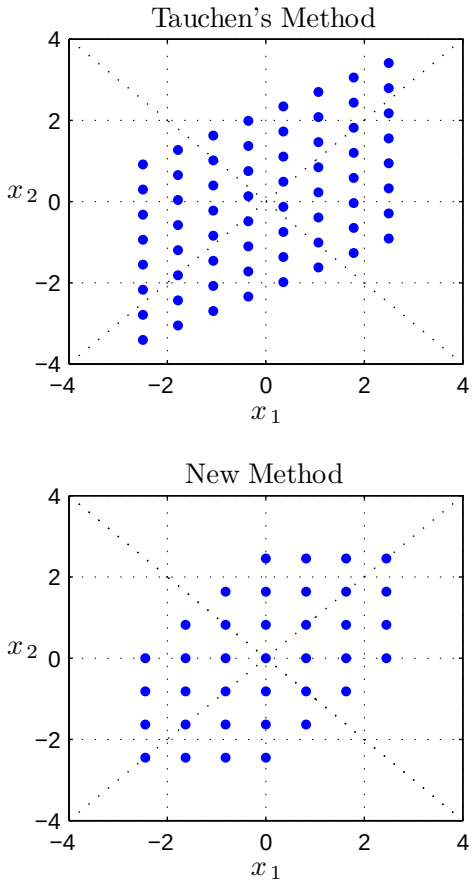
Notes. It compares the simulated time series of a continuous AR(1) process with those generated by Tauchen's and Rouwenhorst's methods for two different levels of persistence: $\rho = 0.99$ and $\rho = 0.999$. The number of grid points for the state variable, y, is nine - i.e., $N = 9$ and $\text{std}(y) = 1$ in all cases.

Figure 2. A highly-persistent vector autoregression



Notes. Using the results in Tables 1 and 3, we plot the approximations by both Tauchen's and the new methods against their targets.

Figure 3. Discretization of symmetric shocks



Notes. These are the scatter diagrams of the series generated by both Tauchen's and the new methods. See the discussion in Section 2.2.3 for details.

Table 1A. Approximation by Tauchen's method: Mean

N	ρ	$\frac{\lg(1-\hat{\rho}_1)}{\lg(1-\rho_1)}$	$\frac{\lg(1-\hat{\rho}_2)}{\lg(1-\rho_2)}$	$\frac{\hat{\alpha}}{\alpha}$	$\frac{\hat{\sigma}_{x_1}}{\sigma_{x_1}}$	$\frac{\hat{\sigma}_{x_2}}{\sigma_{x_2}}$
9	0.5	0.990	0.919	0.785	1.015	0.893
	0.9	0.983	0.861	0.851	1.065	0.742
	0.99	1.272	1.281	1.002	1.220	1.227
	0.999	NA	NA	NA	NA	NA
	0.9999	NA	NA	NA	NA	NA
19	0.5	1.000	0.97	0.842	1.008	1.042
	0.9	0.999	0.930	0.867	1.032	0.859
	0.99	1.031	1.031	1.001	1.215	1.213
	0.999	1.706	2.715	2.154	0.356	1.156
	0.9999	NA	NA	NA	NA	NA
49	0.5	1.000	1.000	0.947	1.002	1.056
	0.9	1.000	0.997	0.959	1.008	1.036
	0.99	1.000	1.000	1.001	1.076	1.075
	0.999	1.158	1.773	1.810	1.297	3.632
	0.9999	NA	NA	NA	NA	NA

Table 1B. Approximation by Tauchen's method: RMSE

N	ρ	$\frac{\lg(1-\hat{\rho}_1)}{\lg(1-\rho_1)}$	$\frac{\lg(1-\hat{\rho}_2)}{\lg(1-\rho_2)}$	$\frac{\hat{\alpha}}{\alpha}$	$\frac{\hat{\sigma}_{x_1}}{\sigma_{x_1}}$	$\frac{\hat{\sigma}_{x_2}}{\sigma_{x_2}}$
9	0.5	0.010	0.082	0.216	0.015	0.111
	0.9	0.017	0.144	0.165	0.065	0.264
	0.99	0.272	0.281	0.006	0.221	0.228
	0.999	NA	NA	NA	NA	NA
	0.9999	NA	NA	NA	NA	NA
19	0.5	0.004	0.026	0.158	0.008	0.043
	0.9	0.003	0.070	0.133	0.032	0.141
	0.99	0.032	0.032	0.003	0.215	0.213
	0.999	0.715	1.724	1.155	0.662	0.527
	0.9999	NA	NA	NA	NA	NA
49	0.5	0.003	0.003	0.053	0.003	0.056
	0.9	0.002	0.006	0.041	0.009	0.037
	0.99	0.005	0.004	0.003	0.077	0.075
	0.999	0.159	0.773	0.812	0.304	2.632
	0.9999	NA	NA	NA	NA	NA

Notes. Table 1A displays the mean of the estimated parameters of the data generated by Tauchen's method relative to their corresponding targets. Table 1B displays the RMSE of the estimated parameters relative to their true values. lg denotes logarithm with base 10. NA denotes the cases where the method cannot generate any data. See the text for details.

Table 2A. Approximated AR(1) process: Mean

N	ρ	Tauch.			T-H			T-H-F			A-C			Rouwn.		
		$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$
9	0.5	0.997	1.028	0.976	1.000	1.000	1.000	1.000	1.000	1.000	0.942	0.976	0.773	1.000	1.000	0.917
	0.9	0.993	1.1043	0.948	0.944	0.928	0.832	0.998	0.994	0.962	0.909	0.976	0.773	1.000	1.000	0.917
	0.99	1.432	1.286	0.875	0.622	0.398	0.623	1.220	0.906	0.729	0.799	0.976	0.773	1.000	1.000	0.916
	0.999	NA	NA	NA	0.426	0.130	0.601	NA	NA	NA	0.701	0.976	0.773	1.000	1.000	0.915
	0.9999	NA	NA	NA	0.320	0.041	0.599	NA	NA	NA	0.651	0.976	0.774	1.000	1.001	0.918
19	0.5	0.996	1.003	0.973	1.000	1.000	1.000	1.000	1.000	1.000	0.977	0.991	0.875	1.000	1.000	0.963
	0.9	0.995	1.016	0.960	0.998	0.997	0.981	1.000	1.000	1.000	0.960	0.991	0.875	1.000	1.000	0.963
	0.99	1.000	1.153	0.900	0.777	0.585	0.660	1.027	1.000	0.940	0.899	0.991	0.875	1.000	1.000	0.963
	0.999	1.723	1.261	0.913	0.543	0.200	0.609	1.897	0.652	1.480	0.790	0.991	0.875	1.000	1.001	0.964
	0.9999	NA	NA	NA	0.409	0.063	0.605	NA	NA	NA	0.718	0.990	0.875	0.998	0.993	0.964
49	0.5	0.996	0.998	0.973	1.000	1.000	1.000	1.000	1.000	1.000	0.993	0.997	0.945	1.000	1.000	0.986
	0.9	0.995	0.997	0.962	1.000	1.000	1.000	1.000	1.000	1.000	0.988	0.997	0.945	1.000	1.000	0.986
	0.99	0.994	1.018	0.942	0.917	0.822	0.753	1.000	1.000	0.999	0.964	0.997	0.945	0.997	0.994	0.986
	0.999	1.012	1.172	0.873	0.669	0.316	0.621	1.225	1.044	0.985	0.895	0.997	0.945	1.000	0.999	0.984
	0.9999	1.569	0.409	5.986	0.507	0.102	0.608	NA	NA	NA	0.806	1.007	0.946	1.001	1.005	0.986

Notes. Table 2A compares the accuracy of different approximation methods for an independent AR(1) process in terms of the mean of estimated parameters relative to their true values. Tauch. is Tauchen's (1986) method, T-H is Tauchen and Hussey's (1991) method, T-H-F is Flodén's alternative of Tauchen and Hussey's (1991) method, A-C is Adda and Cooper's (2003) method and Rouwn. is Rouwenhorst's (1995) method. lg denotes logarithm with base 10. NA denotes the cases where the corresponding method cannot generate any data.

Table 2B. Approximated AR(1) process: RMSE

N	ρ	Tauch.			T-H			T-H-F			A-C			Rouwn.		
		$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$	$\frac{\lg(1-\hat{\rho})}{\lg(1-\rho)}$	$\frac{\hat{\sigma}}{\sigma}$	$\frac{\hat{\kappa}}{\kappa}$
9	0.5	0.003	0.028	0.024	0.001	0.000	0.000	0.001	0.000	0.001	0.058	0.024	0.227	0.001	0.000	0.083
	0.9	0.007	0.104	0.052	0.056	0.072	0.167	0.003	0.006	0.038	0.091	0.024	0.227	0.001	0.001	0.083
	0.99	0.432	0.286	0.125	0.378	0.602	0.378	0.219	0.094	0.281	0.201	0.024	0.227	0.001	0.002	0.084
	0.999	NA	NA	NA	0.574	0.870	0.399	NA	NA	NA	0.299	0.024	0.227	0.002	0.007	0.085
	0.9999	NA	NA	NA	0.680	0.959	0.401	NA	NA	NA	0.349	0.024	0.226	0.004	0.020	0.086
19	0.5	0.004	0.003	0.027	0.001	0.000	0.000	0.001	0.000	0.001	0.023	0.009	0.125	0.001	0.000	0.037
	0.9	0.005	0.016	0.040	0.002	0.003	0.019	0.001	0.001	0.001	0.038	0.009	0.125	0.001	0.001	0.037
	0.99	0.001	0.153	0.100	0.222	0.414	0.340	0.027	0.003	0.060	0.101	0.009	0.125	0.001	0.002	0.037
	0.999	0.724	0.273	0.164	0.457	0.801	0.391	0.900	0.371	2.464	0.209	0.009	0.125	0.002	0.007	0.037
	0.9999	NA	NA	NA	0.591	0.936	0.396	NA	NA	NA	0.282	0.011	0.124	0.005	0.024	0.047
49	0.5	0.004	0.002	0.027	0.001	0.000	0.001	0.001	0.000	0.001	0.007	0.003	0.055	0.001	0.000	0.014
	0.9	0.005	0.003	0.037	0.001	0.001	0.001	0.000	0.001	0.001	0.012	0.003	0.055	0.001	0.001	0.014
	0.99	0.006	0.018	0.058	0.083	0.178	0.247	0.001	0.002	0.004	0.036	0.003	0.055	0.001	0.002	0.014
	0.999	0.012	0.172	0.127	0.330	0.684	0.379	0.226	0.084	0.087	0.105	0.005	0.055	0.002	0.008	0.019
	0.9999	0.575	0.614	11.97	0.493	0.898	0.392	NA	NA	NA	0.194	0.009	0.055	0.004	0.020	0.037

Notes. Table 2B displays the RMSE of the estimated parameters relative to their true values. Tauch. is Tauchen's (1986) method, T-H is Tauchen and Hussey's (1991) method, T-H-F is Flodén's alternative of Tauchen and Hussey's (1991) method, A-C is Adda and Cooper's (2003) method and Rouwn. is Rouwenhorst's (1995) method. lg denotes logarithm with base 10. NA denotes the cases where the corresponding method cannot generate any data.

Table 3A. Approximation by the new method: Mean

N	ρ	$\frac{\lg(1-\hat{\rho}_1)}{\lg(1-\rho_1)}$	$\frac{\lg(1-\hat{\rho}_2)}{\lg(1-\rho_2)}$	$\frac{\hat{\alpha}}{\alpha}$	$\frac{\hat{\sigma}_{x_1}}{\sigma_{x_1}}$	$\frac{\hat{\sigma}_{x_2}}{\sigma_{x_2}}$
	0.5	1.001	0.999	1.000	1.000	0.998
	0.9	1.000	0.999	1.000	1.000	0.998
9	0.99	0.998	0.999	0.999	0.996	0.997
	0.999	1.000	0.999	1.001	1.002	0.999
	0.9999	0.996	0.999	0.992	0.984	0.997
	0.5	0.999	0.999	1.000	1.000	0.999
	0.9	1.000	1.000	1.000	1.000	1.000
19	0.99	0.999	0.999	0.999	0.998	0.998
	0.999	1.000	1.000	0.995	1.001	0.999
	0.9999	1.002	1.000	0.989	1.011	1.001
	0.5	1.000	1.001	1.000	1.000	1.003
	0.9	1.000	1.000	1.000	1.000	1.000
49	0.99	0.999	0.999	0.998	0.998	0.997
	0.999	1.003	1.001	0.992	1.010	1.022
	0.9999	1.007	1.001	0.920	1.037	0.999

Table 3B. Approximation by the new method: RMSE

N	ρ	$\frac{\lg(1-\hat{\rho}_1)}{\lg(1-\rho_1)}$	$\frac{\lg(1-\hat{\rho}_2)}{\lg(1-\rho_2)}$	$\frac{\hat{\alpha}}{\alpha}$	$\frac{\hat{\sigma}_{x_1}}{\sigma_{x_1}}$	$\frac{\hat{\sigma}_{x_2}}{\sigma_{x_2}}$
	0.5	0.004	0.004	0.004	0.001	0.009
	0.9	0.003	0.004	0.003	0.003	0.008
9	0.99	0.005	0.005	0.003	0.011	0.010
	0.999	0.008	0.004	0.012	0.027	0.012
	0.9999	0.021	0.004	0.054	0.082	0.030
	0.5	0.004	0.004	0.004	0.001	0.009
	0.9	0.003	0.004	0.003	0.003	0.009
19	0.99	0.005	0.004	0.003	0.011	0.01
	0.999	0.008	0.003	0.014	0.026	0.009
	0.9999	0.017	0.004	0.061	0.078	0.024
	0.5	0.004	0.004	0.004	0.001	0.010
	0.9	0.003	0.004	0.003	0.003	0.009
49	0.99	0.005	0.005	0.004	0.012	0.011
	0.999	0.011	0.004	0.015	0.039	0.010
	0.9999	0.025	0.004	0.116	0.123	0.013

Notes. Table 3A displays the mean of the estimated parameters of the data generated by the new method relative to their targets. Table 3B displays the RMSE of the estimated parameters relative to their true values. lg denotes logarithm with base 10. See the text for details.

Table 4. Symmetric shocks

Ratios	Tauchen	New	True
	Mean		
$\text{std}(\hat{x}_1^2)/\text{std}(\hat{x}_2^2)$	0.9774	1.0009	1
$\text{std}(\hat{x}_1^3)/\text{std}(\hat{x}_2^3)$	0.9471	1.0006	1
$\text{std}(\hat{x}_1^4)/\text{std}(\hat{x}_2^4)$	0.8544	1.0005	1
	RMSE		
$\text{std}(\hat{x}_1^2)/\text{std}(\hat{x}_2^2)$	0.0239	0.0084	0
$\text{std}(\hat{x}_1^3)/\text{std}(\hat{x}_2^3)$	0.0595	0.0095	0
$\text{std}(\hat{x}_1^4)/\text{std}(\hat{x}_2^4)$	0.1465	0.0116	0

Notes. Table 4 shows the simulation results based on the example considered in Section 2.2.3. See the text for details.

Table 5. Results from value function iteration

$\overline{N}$	$\text{cv}(r_t)$		$\text{corr}(r_t, r_{t+1})$	
	Tauchen	New	Tauchen	New
$\rho_1 = 0.5$, $\rho_2 = 0.7$, $\gamma = 0.9$				
5	0.0040	0.0040	0.7399	0.7623
9	0.0044	0.0043	0.7582	0.7635
19	0.0044	0.0043	0.7629	0.7633
29	0.0044	0.0044	0.7630	0.7639
49	0.0044	0.0044	0.7634	0.7639
$\rho_1 = 0.5$, $\rho_2 = 0.7$, $\gamma = -0.9$				
5	0.0070	0.0072	0.6309	0.6483
9	0.0076	0.0076	0.6492	0.6519
19	0.0077	0.0076	0.6529	0.6529
29	0.0077	0.0076	0.6529	0.6533
49	0.0077	0.0077	0.6532	0.6539
$\rho_1 = 0.99$, $\rho_2 = 0.97$, $\gamma = 0.9$				
5	0.1063	0.0395	0.9998	0.9890
9	0.0952	0.0411	0.9971	0.9895
19	0.0613	0.0418	0.9916	0.9896
29	0.0518	0.0419	0.9904	0.9896
49	0.0459	0.0422	0.9895	0.9897
$\rho_1 = 0.99$, $\rho_2 = 0.97$, $\gamma = -0.9$				
5	0.3166	0.1131	0.9998	0.9823
9	0.2571	0.1117	0.9961	0.9813
19	0.1518	0.1124	0.9839	0.9812
29	0.1315	0.1126	0.9816	0.9812
49	0.1205	0.1130	0.9810	0.9814

Notes. Table 5 shows the results from the value function iteration where $\text{cv}(r_t)$ and $\text{corr}(r_t, r_{t+1})$ are the volatility and the serial correlation of the vacancy filling rate respectively.