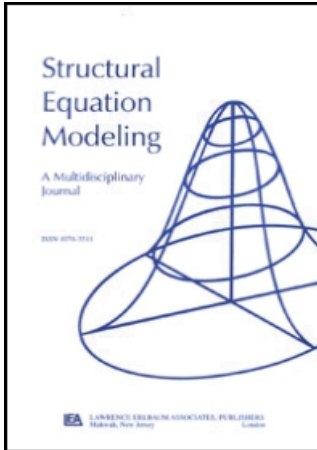


This article was downloaded by:[Zhang, Zhiyong]
On: 19 July 2008
Access Details: [subscription number 795149962]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Structural Equation Modeling: A Multidisciplinary Journal

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t775653699>

Comparisons of Four Methods for Estimating a Dynamic Factor Model

Zhiyong Zhang ^a; Ellen L. Hamaker ^b; John R. Nesselroade ^a

^a Department of Psychology, The University of Virginia,

^b Department of Methods and Statistics, The Utrecht University, Netherlands

Online Publication Date: 01 July 2008

To cite this Article: Zhang, Zhiyong, Hamaker, Ellen L. and Nesselroade, John R. (2008) 'Comparisons of Four Methods for Estimating a Dynamic Factor Model', Structural Equation Modeling: A Multidisciplinary Journal, 15:3, 377 — 402

To link to this article: DOI: 10.1080/10705510802154281

URL: <http://dx.doi.org/10.1080/10705510802154281>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Comparisons of Four Methods for Estimating a Dynamic Factor Model

Zhiyong Zhang

*Department of Psychology
The University of Virginia*

Ellen L. Hamaker

*Department of Methods and Statistics
The Utrecht University, Netherlands*

John R. Nesselroade

*Department of Psychology
The University of Virginia*

Four methods for estimating a dynamic factor model, the direct autoregressive factor score (DAFS) model, are evaluated and compared. The first method estimates the DAFS model using a Kalman filter algorithm based on its state space model representation. The second one employs the maximum likelihood estimation method based on the construction of a block-Toeplitz covariance matrix in the structural equation modeling framework. The third method is built in the Bayesian framework and implemented using Gibbs sampling. The fourth is the least squares method, which also employs the block-Toeplitz matrix. All 4 methods are implemented in currently available software. The simulation study shows that all 4 methods reach appropriate parameter estimates with comparable precision. Differences among the 4 estimation methods and related software are discussed.

As generalizations of the P-technique factor model (Cattell, Cattell, & Rhymer, 1947), the dynamic factor models (DFMs) can represent process and change more effectively than their precursor (Browne & Nesselroade, 2005; Engle &

Correspondence should be addressed to Zhiyong Zhang, Department of Psychology, P. O. Box 400400, University of Virginia, Charlottesville, VA 22904-4400. E-mail: zz5m@virginia.edu

Watson, 1981; Geweke & Singleton, 1981; McArdle, 1982; Molenaar, 1985; Nesselroade, McArdle, Aggen, & Meyers, 2001). DFMs not only model the structure between the latent factors and observed variables, but also characterize time-dependent relationships among factors or between factors and observed variables. Nesselroade et al. (2001) labeled proposed DFMs the direct autoregressive factor score (DAFS) model, which is characterized by the autoregressive structure of factors (Engle & Watson, 1981; McArdle, 1982; Molenaar, 1985), and the white noise factor score (WNFS) model, which is characterized by direct lagged effects of factors on the observed variables (Geweke & Singleton, 1981; Molenaar, 1985) and examined the differences and similarities of these models. Browne and Nesselroade (2005) also referred to these models as the process model and the shock model, respectively. These two sets of models can be transformed to each other (Browne & Nesselroade, 2005; Molenaar, 1985; Nesselroade et al., 2001). However, these potentially valuable models have not been widely applied partly because fitting them is not as straightforward as fitting the common factor analysis model, which can be done with most available general-purpose computer packages.

Compared to the DAFS model, estimation methods for the WNFS model have been well established and evaluated. Molenaar (1985) employed a maximum likelihood estimation (MLE) method based on the block-Toeplitz matrix in the structural equation modeling (SEM) framework. Molenaar and Nesselroade (1998) referred to the estimation method as pseudo-ML because it violates the assumption that the vector-valued observations are independent.¹ They also compared this method with the asymptotically distribution-free estimation method by a Monte Carlo study. Recently, Justiniano (2004) proposed a Bayesian estimation procedure to get the parameter estimates for a WNFS model. Discussion of the estimation of the WNFS model can also be found in Wood and Brown (1994), Hershberger (1998), Nesselroade et al. (2001), and Nesselroade and Molenaar (2002).

The estimation methods of the DAFS model have not been evaluated as much as those for the WNFS model. Engle and Watson (1981) showed that the DAFS model is a special case of a state space model and can be estimated by maximum likelihood methods using a Kalman filter algorithm. The maximum likelihood method based on the block-Toeplitz covariance matrix has also been used to estimate more complex DAFS models (Nesselroade et al., 2001). Recently, Bayesian methods have become more widely used for estimating the DAFS models in terms of its state space representation (Kim & Nelson, 1999, 2001; West, 2000). Also, Browne and Zhang (2006) recently proposed a least squares estimation method and provided free software for implementation (Browne &

¹We refer to it as maximum likelihood estimation throughout this paper bearing in mind the limitation.

Zhang, 2005). However, the performance of these four methods has not been evaluated, especially in terms of comparisons.

In this study, we evaluate and compare the four methods given for estimating the DAFS model. First, we show the connection between a state space model and the DAFS model and how to employ a Kalman filter algorithm to estimate the DAFS model. Second, we explain the underlying ideas of constructing a block-Toeplitz covariance matrix to estimate the DAFS model using MLE in SEM software. Third, we examine Bayesian estimation of the DAFS model using Gibbs sampling comprehensively. Fourth, the main ideas of the least squares estimation method are summarized. The performance of these four methods is compared by simulation studies. All four methods are implemented using currently available software and programs. We begin by defining the DAFS model.

THE DAFS MODEL

One form of the DAFS model is

$$\mathbf{y}_t = \mathbf{\Lambda} \mathbf{f}_t + \mathbf{e}_t \quad (1)$$

$$\mathbf{f}_t = \sum_{l=1}^L \mathbf{B}_l \mathbf{f}_{t-l} + \mathbf{v}_t \quad (2)$$

where \mathbf{y}_t is a $p \times 1$ vector of observed variables measured at time t ($t = 1, 2, \dots, T$), $\mathbf{\Lambda}$ is a $p \times q$ factor loading matrix, \mathbf{f}_t is a $q \times 1$ vector of common factors at time t , \mathbf{e}_t is a $p \times 1$ vector with measurement errors following a multivariate normal distribution with mean zeros and $p \times p$ covariance matrix \mathbf{Q} , \mathbf{f}_{t-l} ($t = 1, 2, \dots, T; l = 0, 1, 2, \dots, L$) is a $q \times 1$ vector of factors l occasions prior to occasion t , \mathbf{B}_l is a $q \times q$ matrix containing the autoregressive and cross-regressive coefficients, and \mathbf{v}_t is a $q \times 1$ random shock vector following a multivariate normal distribution with mean zeros and $q \times q$ covariance matrix \mathbf{D} .

For the DAFS model, $\mathbf{\Lambda}$, \mathbf{B}_l , \mathbf{Q} , and \mathbf{D} are usually specified to be invariant over time. The common factors have an autoregressive structure and a possible cross-regressive structure. The measurement errors \mathbf{e}_t may have autoregressive structures peculiar to themselves. Although the factors at any given time have no direct influence on future values of observed variables, they have an indirect influence as they influence future values of the factors, which, in turn, influence the observed variables concurrently.

To illustrate, a typical DAFS model is portrayed in Figure 1. There are six observed variables and two factors in this model. The first three observed variables load on the first factor and the other three observed variables load on the second factor. The factors have a one-lag autoregressive and cross-regressive structure, which means the first factor at the current time has a direct influence on the first and second factors at the next time, and so does the second factor.

Harvey, 1989). For example, the commonly used ARMA and ARIMA models (e.g., Hamilton, 1994) can be rewritten in a state space format. As shown later, the DAFS model can also be rewritten as a state space model. Actually, the state space model form (see Equations 3 and 4) is identical to the one-lag DAFS model given the time-invariant parameters. Thus, the well-defined MLE method using the Kalman filter algorithm (e.g., Durban & Koopman, 2001) for state space models can be used to estimate the DAFS model.

The Kalman filter is an algorithm for estimating the states of a system from a sequence of noisy observations (Chatfield, 2004). The algorithm together with MLE can be used to estimate the model parameters. The Kalman filter is based on formulating the time series model in state space format, which consists of two equations. First, the measurement equation, also known as the observation equation, relates the observations to the factor scores (or latent states); that is

$$\mathbf{y}_t = \mathbf{S}_t \mathbf{a}_t + \mathbf{d}_t + \mathbf{e}_t \quad (3)$$

where \mathbf{y}_t is a $p \times 1$ vector with observations at occasion t , \mathbf{S}_t is a $p \times k$ matrix with factor loadings, \mathbf{a}_t is a $k \times 1$ vector with latent states, and \mathbf{d}_t is a $p \times 1$ vector of intercepts. The $p \times 1$ vector \mathbf{e}_t contains the noise, which is assumed to be multivariate normally distributed with mean zero and covariance matrix \mathbf{Q}_t .

Second, the transition equation, also known as the state equation, is

$$\mathbf{a}_t = \mathbf{H}_t \mathbf{a}_{t-1} + \mathbf{c}_t + \mathbf{z}_t, \quad (4)$$

where \mathbf{a}_t is as already defined, \mathbf{H}_t is a $k \times k$ matrix with structural coefficients, and \mathbf{c}_t is a $k \times 1$ vector with constants. The $k \times 1$ vector of residuals or innovations, \mathbf{z}_t , is multivariate normally distributed with mean zero and covariance matrix \mathbf{G}_t .

Here, we limit our presentation to the case of time-invariant model parameters so that all the time indexes for the model matrices can be dropped. In addition, we assume the data are centered so that \mathbf{d} and \mathbf{c} can be omitted from the preceding equations. Although the true state \mathbf{a}_t cannot be known, it is estimated based on all observations up to occasion t ; that is, $E[\mathbf{a}_t | \mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_1] = \mathbf{a}_{t|t}$. Associated with this state estimate is an error covariance that is a measure of the estimated accuracy; that is, $\mathbf{P}_{t|t} = E[(\mathbf{a}_t - \mathbf{a}_{t|t})(\mathbf{a}_t - \mathbf{a}_{t|t})']$. To initiate the Kalman filter, one needs to specify $\mathbf{a}_{0|0}$, that is, the state estimation at occasion zero (before observations began), and its associated covariance matrix $\mathbf{P}_{0|0}$. Usually, fairly large variance and negligent covariance values are inserted in $\mathbf{P}_{0|0}$ to reduce the influence of the choice of the initial states. Based on these initial matrices, the equations that make up the Kalman filter are used recursively. First, the state \mathbf{a}_t is estimated based on the state estimation of the previous state $\mathbf{a}_{t-1|t-1}$ through

$$\mathbf{a}_{t|t-1} = \mathbf{H} \mathbf{a}_{t-1|t-1}. \quad (5)$$

Along with this state estimate there is an estimate of the error covariance

$$\mathbf{P}_{t|t-1} = \mathbf{H}\mathbf{P}_{t-1|t-1}\mathbf{H}' + \mathbf{G}. \quad (6)$$

Equations 5 and 6 are known as the prediction equations.

In the next step, the prediction of \mathbf{a}_t is updated in light of the observation \mathbf{y}_t . To this end, first \mathbf{y}_t is predicted from $\mathbf{a}_{t|t-1}$ through

$$\mathbf{y}_{t|t-1} = \mathbf{S}\mathbf{a}_{t|t-1}. \quad (7)$$

This prediction is used to determine the one-step-ahead prediction error along with its covariance matrix; that is,

$$\mathbf{e}_{t|t-1} = \mathbf{S}(\mathbf{a}_t - \mathbf{a}_{t|t-1}) + \mathbf{e}_t \quad (8)$$

$$\mathbf{F}_t = \mathbf{S}\mathbf{P}_{t|t-1}\mathbf{S}' + \mathbf{Q}. \quad (9)$$

Durban and Koopman (2001) showed that, by assuming multivariate normality, the estimate of \mathbf{a}_t can be updated through

$$\mathbf{a}_{t|t} = \mathbf{a}_{t|t-1} + \mathbf{J}_t\mathbf{F}_t^{-1}\mathbf{e}_{t|t-1}, \quad (10)$$

where $\mathbf{J}_t\mathbf{F}_t^{-1}$ is known as the gain matrix or the Kalman gain matrix (Chatfield, 2004). The matrix \mathbf{J}_t represents the covariance between the true state \mathbf{a}_t and the observation \mathbf{y}_t conditional on the information available up to occasion $t-1$. Because $E[\mathbf{a}_t|\mathbf{y}_{t-1}] = \mathbf{a}_{t|t-1}$ and $E[\mathbf{y}_t|\mathbf{y}_{t-1}] = \mathbf{S}\mathbf{a}_{t|t-1}$, \mathbf{J}_t can be written as

$$\mathbf{J}_t = \mathbf{P}_{t|t-1}\mathbf{S}'. \quad (11)$$

In addition, the covariance matrix of the updated state $\mathbf{a}_{t|t}$ is

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{J}_t\mathbf{F}_t^{-1}\mathbf{J}_t'. \quad (12)$$

Equations 10 and 12 are known as the update equations. By recursively using Equations 5 to 12 one can estimate the states for occasions 1 to T. Note that the Kalman filter procedure based on this prediction error decomposition method is not a full information method because only observations up to time t are used when estimating $\mathbf{a}_{t|t}$. An additional state smoothing step can be used to utilize all the information in the data (Durban & Koopman, 2001).

To obtain maximum likelihood estimates for the model parameters in \mathbf{S} , \mathbf{Q} , \mathbf{H} , and \mathbf{G} , one needs to formulate the likelihood function $L(\mathbf{y})$. Often, the following log-likelihood is used (Durban & Koopman, 2001),

$$\log L(\mathbf{y}) = \sum_{t=1}^T \log \text{Prob}(\mathbf{y}_t|\mathbf{Y}_{t-1}), \quad (13)$$

where $\mathbf{Y}_t = \{\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_1\}$ and $\text{Prob}(\mathbf{y}_1 | \mathbf{Y}_0) = \text{Prob}(\mathbf{y}_1)$. Because $\text{Prob}(\mathbf{y}_t | \mathbf{Y}_{t-1}) = \text{Prob}(\mathbf{y}_t | \mathbf{y}_{t-1}) = \text{Prob}(\mathbf{e}_{t|t-1})$, where $\mathbf{e}_{t|t-1}$ and its covariance matrix \mathbf{F}_t are by-products of the Kalman filter (see Equations 8 and 9), we can write

$$\log L = -\frac{p * T}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{F}_t| - \frac{1}{2} \sum_{t=1}^T \mathbf{e}'_{t|t-1} \mathbf{F}_t^{-1} \mathbf{e}_{t|t-1}. \quad (14)$$

Any DAFS model can be rewritten as its state space representation to fit the Kalman filter algorithm. For the model in Equations 1 and 2, its state space representation is

$$\begin{aligned} \mathbf{y}_t &= \mathbf{S} \mathbf{a}_t + \mathbf{e}_t \\ \mathbf{y}_t &= (\mathbf{A} \quad \mathbf{0} \quad \mathbf{0} \quad \dots \quad \mathbf{0}) \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_{t-1} \\ \mathbf{f}_{t-2} \\ \vdots \\ \mathbf{f}_{t-L+1} \end{pmatrix} + \mathbf{e}_t \\ \mathbf{a}_t &= \mathbf{H} \mathbf{a}_{t-1} + \mathbf{z}_t \\ \begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_{t-1} \\ \mathbf{f}_{t-2} \\ \vdots \\ \mathbf{f}_{t-L+1} \end{pmatrix} &= \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \dots & \mathbf{B}_{L-1} & \mathbf{B}_L \\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{f}_{t-1} \\ \mathbf{f}_{t-2} \\ \mathbf{f}_{t-3} \\ \vdots \\ \mathbf{f}_{t-L} \end{pmatrix} + \begin{pmatrix} \mathbf{v}_t \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \end{aligned} \quad (15)$$

where $\mathbf{0}$ in \mathbf{S} is a $p \times q$ zero matrix, $\mathbf{0}$ in \mathbf{H} is a $q \times q$ zero matrix, \mathbf{I} in \mathbf{H} is a $q \times q$ identity matrix, and $\mathbf{0}$ in \mathbf{z}_t is a $q \times 1$ zero vector. The covariance matrix of \mathbf{e}_t is \mathbf{Q} . The covariance matrix of \mathbf{z}_t is

$$\mathbf{G} = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where $\mathbf{0}$ is a $q \times q$ zero matrix.

After we rewrite the DAFS model in its state space format, we can estimate its parameters using the Kalman filter procedure given previously.

MLE USING THE BLOCK-TOEPLITZ
COVARIANCE MATRIX

Molenaar (1985) proposed an MLE method to estimate the DFM parameters using block-Toeplitz lagged covariance matrices. This method is very flexible because it can be implemented in available SEM software. Nesselroade, McArdle, Aggen, and Meyers (2001) employed the same strategy to obtain the estimations for both the WNFS model and the DAFS model in LISREL (Jöreskog, Sörbom, Du Toit, & Du Toit, 2001). However, Nesselroade et al. (2001) only gave a brief outline of this method. We elaborate this method based on a general DAFS model.

In terms of structural equation models, the lag L DAFS model can be expressed as

$$\tilde{\mathbf{y}}_t = \tilde{\mathbf{\Lambda}} \tilde{\mathbf{f}}_t + \tilde{\mathbf{e}}_t, \quad (16)$$

$$\tilde{\mathbf{f}}_t = \tilde{\mathbf{B}} \tilde{\mathbf{f}}_t + \tilde{\mathbf{v}}_t. \quad (17)$$

where

$$\tilde{\mathbf{y}}_t = \begin{pmatrix} \mathbf{y}_{t-L} \\ \vdots \\ \mathbf{y}_{t-1} \\ \mathbf{y}_t \end{pmatrix}, \quad \tilde{\mathbf{\Lambda}} = \begin{pmatrix} \mathbf{\Lambda} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{\Lambda} \end{pmatrix}, \quad \tilde{\mathbf{f}}_t = \begin{pmatrix} \mathbf{f}_{t-L} \\ \vdots \\ \mathbf{f}_{t-1} \\ \mathbf{f}_t \end{pmatrix},$$

$$\tilde{\mathbf{e}}_t = \begin{pmatrix} \mathbf{e}_{t-L} \\ \vdots \\ \mathbf{e}_{t-1} \\ \mathbf{e}_t \end{pmatrix}, \quad \tilde{\mathbf{B}} = \begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_L & \cdots & \mathbf{B}_1 & \mathbf{0} \end{pmatrix}, \quad \text{and} \quad \tilde{\mathbf{v}}_t = \begin{pmatrix} \mathbf{f}_{t-L} \\ \vdots \\ \mathbf{f}_{t-1} \\ \mathbf{v}_t \end{pmatrix}$$

with $t = L + 1, \dots, T$.

The population or the model indicated covariance matrix for this model is

$$\Sigma = COV(\tilde{\mathbf{y}}_t) = \tilde{\mathbf{\Lambda}} (\mathbb{I} - \tilde{\mathbf{B}})^{-1} \tilde{\mathbf{\Omega}} [(\mathbb{I} - \tilde{\mathbf{B}})^{-1}]' \tilde{\mathbf{\Lambda}}'$$

$$= \begin{pmatrix} \mathbf{\Lambda} \mathbf{\Omega}_0 \mathbf{\Lambda}' + \mathbf{Q} & & & \\ \vdots & \ddots & & \\ \mathbf{\Lambda} \mathbf{\Omega}_{L-1} \mathbf{\Lambda}' & \cdots & \mathbf{\Lambda} \mathbf{\Omega}_0 \mathbf{\Lambda}' + \mathbf{Q} & \\ \mathbf{\Lambda} \sum_{i=1}^L (\mathbf{B}_{L-i+1} \mathbf{\Omega}_{|L-i|}) \mathbf{\Lambda}' & \cdots & \mathbf{\Lambda} \sum_{i=1}^L (\mathbf{B}_{L-i+1} \mathbf{\Omega}'_{|L-i|}) \mathbf{\Lambda}' & \\ & & \mathbf{\Lambda} (\sum_{i=1}^L \sum_{j=1}^L \mathbf{B}_i \mathbf{\Omega}_{|i-j|} \mathbf{B}'_j + \mathbf{D}) \mathbf{\Lambda}' + \mathbf{Q} & \end{pmatrix}, \quad (18)$$

where $\mathbf{\Omega}_i, i = 0, \dots, L-1$ is the autocovariance matrix of the common factors at lag i .

Given the covariance stationarity assumption of the observed variables,

$$\mathbf{\Omega}_0 = \sum_{i=1}^L \sum_{j=1}^L \mathbf{B}_i \mathbf{\Omega}_{|i-j|} \mathbf{B}'_j + \mathbf{D}, \quad (19)$$

and

$$\mathbf{\Omega}_i = \sum_{l=1}^L \mathbf{B}_{L-l+1} \mathbf{\Omega}'_{|L-l-i+1|}, \quad (20)$$

with $i = 1, \dots, L-1$.

Then Equation 18 can be simplified to

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Lambda} \mathbf{\Omega}_0 \mathbf{\Lambda}' + \mathbf{Q} & & & & \\ & \vdots & \ddots & & \\ & \mathbf{\Lambda} \mathbf{\Omega}_{L-1} \mathbf{\Lambda}' & \cdots & \mathbf{\Lambda} \mathbf{\Omega}_0 \mathbf{\Lambda}' + \mathbf{Q} & \\ \mathbf{\Lambda} \sum_{l=1}^L (\mathbf{B}_{L-l+1} \mathbf{\Omega}_{|l-1|}) \mathbf{\Lambda}' & \cdots & \mathbf{\Lambda} \mathbf{\Omega}_1 \mathbf{\Lambda}' & \mathbf{\Lambda} \mathbf{\Omega}_0 \mathbf{\Lambda}' + \mathbf{Q} & \end{pmatrix}. \quad (21)$$

The sample variance-covariance matrix is

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}(0) & & & \\ \vdots & \ddots & & \\ \mathbf{S}(L-1) & \cdots & \mathbf{S}(0) & \\ \mathbf{S}(L) & \cdots & \mathbf{S}(1) & \mathbf{S}(0) \end{pmatrix}, \quad (22)$$

where $\mathbf{S}(0)$ and $\mathbf{S}(l)$ are the lag 0 and lag l ($l = 1, \dots, L$) $p \times p$ covariance matrices that can be calculated directly from the observed data. Note that $\mathbf{S}(l)$ is usually not a symmetric matrix because given any two variables X and Y , X lagging Y and Y lagging X are typically different. \mathbf{S} is called the block-Toeplitz covariance matrix and it is symmetric.

The parameters in Equations 16 and 17 can then be estimated using the maximum likelihood method via SEM software such as LISREL (Jöreskog et al., 2001) and Mplus (Muthén & Muthén, 2001) by minimizing the following function:

$$F_{ML} = \log |\mathbf{\Sigma}| + \text{tr}(\mathbf{S}\mathbf{\Sigma}^{-1}) - \log |\mathbf{S}| - (L+1)p. \quad (23)$$

In using the lagged observations $\tilde{\mathbf{y}}_t$ to calculate the observed covariance matrix, we are violating the observation independency assumption (Bollen, 1989) underlying the maximum likelihood method. However, Hamaker, Dolan, and Molenaar (2002) showed that for the univariate case, the estimates for the

autoregressive model were still consistent. The simulation study from Molenaar and Nesselroade (1998) also showed that this method gave reasonable estimates for WNFS models. Similar results were found in our simulation presented later on the DAFS models.

BAYESIAN ESTIMATION USING GIBBS SAMPLING

Although there has been no specific application to the DAFS model, Gibbs sampling has been used for estimating both linear and nonlinear state space models (Carlin, Polson, & Stoffer, 1992; Carter & Kohn, 1994). In this article, we employed the same strategy to construct the Bayesian estimation procedure for the DAFS model. Because any DAFS model can be reformed as a one-lag DAFS model using the exact same formulation as in Equation 15, we only illustrate the Gibbs sampling procedure for the one-lag DAFS model for the sake of simplifying the display of equations. The following two-step procedure can be used to sample factor scores and model parameters.

1. Conditional on the model parameters $\Lambda, \mathbf{B}, \mathbf{Q}, \mathbf{D}$, and the observed data, generate the factor scores \mathbf{f}_t ,
2. Conditional on the factor scores \mathbf{f}_t and the observed data, generate $\Lambda, \mathbf{B}, \mathbf{Q}$, and \mathbf{D} .

Sample \mathbf{f}_t

For the first step, conditional on the $\Lambda, \mathbf{B}, \mathbf{Q}, \mathbf{D}$, and the observed data, the method used in Carlin et al. (1992) can be implemented for sampling the factor scores. Let the prior distribution of \mathbf{f}_0 be

$$\mathbf{f}_0 \sim MN(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0),$$

where MN represents the multivariate normal distribution. The posterior distribution of \mathbf{f}_t is

$$\mathbf{f}_t | \mathbf{f}_{j \neq t}, \mathbf{f}_0, \Lambda, \mathbf{B}, \mathbf{Q}, \mathbf{D} \sim MN(\boldsymbol{\Sigma}_t \cdot \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), \quad (24)$$

with

$$\boldsymbol{\Sigma}_t^{-1} = \mathbf{D}^{-1} + \Lambda' \mathbf{Q}^{-1} \Lambda + \mathbf{B}' \mathbf{D}^{-1} \mathbf{B}$$

and

$$\boldsymbol{\mu}_t' = \mathbf{f}'_{t-1} \mathbf{B} \mathbf{D}^{-1} + \mathbf{y}'_t \mathbf{Q}^{-1} \Lambda + \mathbf{f}'_{t+1} \mathbf{D}^{-1} \mathbf{B}.$$

From the distribution in Equation 24, we can sample \mathbf{f}_t .

Sample Λ , B , Q , and D

Conditional on \mathbf{f}_t , Equations 1 and 2 are independent. For Equation 1, because the uniqueness factors are independent, it can be viewed as a group of independent multiple regression models that have the form of

$$y_{it} = \lambda_i \mathbf{f}_t + e_{it},$$

where $i = 1, 2, \dots, p$, λ_i is the i th row in the factor loading matrix Λ and the variance of e_{it} is q_i , the i th diagonal element of Q .

Let the prior distribution of λ_i^t be

$$\lambda_i^t \sim MN(\lambda_{i0}^t, \Sigma_{i0}).$$

Conditional on the other parameters and the observed data, the posterior distribution of λ_i^t is

$$\lambda_i^t | \mathbf{f}_t, \mathbf{y}_t, q_i \sim MN(\lambda_{i1}^t, \Sigma_{i1}), \quad (25)$$

with

$$\lambda_{i1}^t = (\Sigma_{i0}^{-1} + q_i^{-1} \mathbf{X}'\mathbf{X})^{-1} (\Sigma_{i0}^{-1} \lambda_{i0}^t + q_i^{-1} \mathbf{X}'\mathbf{y}_t)$$

and

$$\Sigma_{i1} = (\Sigma_{i0}^{-1} + q_i^{-1} \mathbf{X}'\mathbf{X})^{-1},$$

where \mathbf{X} is a $q \times T$ matrix of factor scores and \mathbf{y}_i is a $T \times 1$ vector with the observations for the i th observed variable.

Given the prior distribution of q_i ,

$$q_i \sim IG\left(\frac{v_0}{2}, \frac{\delta_0}{2}\right),$$

the posterior of q_i is

$$q_i | \lambda_i, \mathbf{f}_t, \mathbf{y}_t \sim IG\left(\frac{v_1}{2}, \frac{\delta_1}{2}\right), \quad (26)$$

with

$$v_1 = v_0 + T,$$

$$\delta_1 = \delta_0 + \sum_{t=1}^T (y_{it} - \lambda_i \mathbf{f}_t)^2,$$

where IG represents the inverse Γ distribution.

For Equation 2, if the shocks or residuals \mathbf{v}_t are uncorrelated, the same procedure given earlier can be used to get the posterior distribution of \mathbf{B} and \mathbf{D} . Usually, \mathbf{v}_t are correlated. We then obtain the conditional distribution based on the correlated shocks following Kadiyala and Karlsson (1997).

Let the prior distribution of \mathbf{B} be

$$\text{vec}(\mathbf{B}) \sim MN(\text{vec}(\mathbf{B}_0), \mathbf{D} \otimes \boldsymbol{\Sigma}_{\mathbf{B}_0}),$$

where the $\text{vec}(\mathbf{B})$ means column vectorization of the matrix \mathbf{B} . Then the posterior distribution of \mathbf{B} is

$$\text{vec}(\mathbf{B}) | \mathbf{f}_t, \mathbf{y}_t, \mathbf{D} \sim MN(\text{vec}(\mathbf{B}_1), \mathbf{D} \otimes \boldsymbol{\Sigma}_{\mathbf{B}_1}) \quad (27)$$

with

$$\mathbf{B}_1 = \boldsymbol{\Sigma}_{\mathbf{B}_1} (\boldsymbol{\Sigma}_{\mathbf{B}_0}^{-1} \mathbf{B}_0 + \mathbf{X}'_{t-1} \mathbf{X}_{t-1} \hat{\mathbf{B}})$$

and

$$\boldsymbol{\Sigma}_{\mathbf{B}_1} = (\boldsymbol{\Sigma}_{\mathbf{B}_0}^{-1} + \mathbf{X}'_{t-1} \mathbf{X}_{t-1})^{-1},$$

where $\hat{\mathbf{B}}$ are the ordinary least squares estimates and \mathbf{X}_{t-1} is a $q \times (T-1)$ matrix of factor scores.

Let the variance-covariance matrix \mathbf{D} have a inverse Wishart prior distribution

$$\mathbf{D} \sim IWish(v_0, \mathbf{D}_0).$$

The posterior distribution of \mathbf{D} is then expressed as

$$\mathbf{D} | \mathbf{f}_t, \mathbf{y}_t, \mathbf{B} \sim IWish(v_1, \mathbf{D}_1) \quad (28)$$

with

$$v_1 = v_0 + T - 1$$

and

$$\begin{aligned} \mathbf{D}_1 = & \hat{\mathbf{B}}' \mathbf{X}'_{t-1} \mathbf{X}_{t-1} \hat{\mathbf{B}} + \mathbf{B}'_0 \boldsymbol{\Sigma}_{\mathbf{B}_0}^{-1} \mathbf{B}_0 + \mathbf{D}_0 \\ & + (\mathbf{X}_t - \mathbf{X}_{t-1} \hat{\mathbf{B}})' (\mathbf{X}_t - \mathbf{X}_{t-1} \hat{\mathbf{B}}) - \mathbf{B}'_1 (\boldsymbol{\Sigma}_{\mathbf{B}_0}^{-1} + \mathbf{X}'_{t-1} \mathbf{X}_{t-1}) \mathbf{B}_1. \end{aligned}$$

After obtaining all the conditional posterior distributions, the following procedure can be used to generate a sequence for each parameter:

1. Sample \mathbf{f}_t from $\mathbf{f}_t | \mathbf{f}_{j \neq t}, \mathbf{f}_0, \boldsymbol{\Lambda}, \mathbf{B}, \mathbf{Q}, \mathbf{D}$, $t = 1, \dots, T$.
2. Sample \mathbf{B} from $\text{vec}(\mathbf{B}) | \mathbf{f}_t, \mathbf{y}_t, \mathbf{D}$.

3. Sample \mathbf{D} from $\mathbf{D}|\mathbf{f}_t, \mathbf{y}_t, \mathbf{B}$.
4. Sample λ_i from $\lambda_i'|\mathbf{f}_t, \mathbf{y}_t, \mathbf{Q}$, $i = 1, \dots, p$.
5. Sample q_i from $q_i|\mathbf{f}_t, \mathbf{y}_t, \lambda_i$, $i = 1, \dots, p$.
6. Repeat this process to get a sequence of data for each estimated parameter.

The statistical inferences are carried out based on the generated sequence. For example, a point estimation of the parameter can be represented by the mean of the sequence for that parameter. The confidence interval can be constructed using the percentiles of the sequence.

LEAST SQUARES ESTIMATION METHOD

The least squares method has been discussed comprehensively by Browne and Zhang (2005, 2006). Thus, we do not repeat the technical details and only summarize the main ideas here. In the derivation of the least squares estimation method, the model used by Browne and Zhang (2005, 2006; see also Browne & Nesselroade, 2005) is

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{f}_t + \mathbf{e}_t \quad (29)$$

$$\mathbf{f}_t = \sum_{l=1}^L \mathbf{B}_l \mathbf{f}_{t-l} + \mathbf{v}_t + \sum_{j=1}^d \mathbf{A}_j \mathbf{v}_{t-j} \quad (30)$$

where $\boldsymbol{\mu}$ is a mean vector, d is the total number of moving average terms, \mathbf{A}_j are the moving average weight matrices, and the other terms have the same definition as those in the Equations 1 and 2. Clearly, the DAFS model discussed in this study can be viewed as a special case of this model.

The least squares estimation is based on the lagged correlation matrix. The theoretically lagged correlation matrix $\boldsymbol{\Sigma}_l = (\sigma_{lij})_{p \times p}$ at lag l can be obtained from the model in Equations 29 and 30. The empirical lagged block-Toeplitz correlation matrix $\mathbf{S}_l = (s_{lij})_{p \times p}$ at lag l can be calculated from the observed data. The parameter estimates in this model can then be obtained by minimizing the following discrepancy function:

$$F = \sum_{l=0}^L \sum_{i=1}^p \sum_{j=1}^p (\sigma_{lij} - s_{lij})^2. \quad (31)$$

In this method, the observed covariance matrix \mathbf{S}_0 and lagged covariance matrices \mathbf{S}_l ($l = 1, \dots, L$) are used only once.

SIMULATION STUDIES

In the previous sections, four estimation methods were discussed more or less from a theoretical perspective. In this section, the four methods are compared at a more concrete level via simulation studies. To facilitate applications of the DAFS model, all four estimation methods are implemented using currently available software.

Simulation Design

To generate the data, the DAFS model portrayed in Figure 1 was employed. Two major influence factors were varied in the simulation: the length of the time series and the amount of the measurement errors. For the length of the time series, three conditions were used, $T = 50, 100, \text{ and } 200$. Four conditions were manipulated for the measurement errors. The percentage of the measurement error variance to the variance of the observed variables was 9%, 23%, 33%, and 50%, which corresponded to the ratios of the variance of measurement errors to latent factors, .1, .3, .5, and 1. These two sets of conditions were fully crossed, resulting in 12 cells in the simulation of data.

In matrix notation, the DAFS model true parameter values in Figure 1 are given as follows. The factor loading matrix $\Lambda = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}'$. The measurement error covariance matrix had four conditions, $\mathbf{Q}_1 = \text{diag}(.1, .1, .1, .1, .1, .1)$, $\mathbf{Q}_2 = \text{diag}(.3, .3, .3, .3, .3, .3)$, $\mathbf{Q}_3 = \text{diag}(.5, .5, .5, .5, .5, .5)$, and $\mathbf{Q}_4 = \text{diag}(1, 1, 1, 1, 1, 1)$, corresponding to the four error conditions already defined. $\text{Diag}()$ implies a diagonal matrix. The autoregressive matrix $\mathbf{B}_1 = \begin{pmatrix} .8 & 0 \\ 0 & .8 \end{pmatrix}$ and the covariance matrix for the random shock \mathbf{v}_t was $\mathbf{D} = \begin{pmatrix} .36 & .18 \\ .18 & .36 \end{pmatrix}$. Here \mathbf{B}_1 and \mathbf{D} were chosen to ensure that the factors were stationary time series and the variance of the factors was 1 for convenience without losing generality. Furthermore, although the cross-regressive parameters were set to 0, these parameters were estimated as an unknown parameter in the later section.

Given the predefined parameter values, the following procedure was followed to generate the data:

1. Set the length of data T and the vector $(f_{10}, f_{20})' = (0, 0)'$.
2. Set the iteration number $t = 1$.
3. Generate the vector $(v_{1t}, v_{2t})'$ from the bivariate normal distribution $MN(\mathbf{0}, \mathbf{D})$.
4. Calculate $(f_{1t}, f_{2t})'$ using Equation 2.

5. Generate $\mathbf{e}_{it}, i = 1, \dots, 6$ from the multivariate normal distribution $MN(\mathbf{0}, \mathbf{Q})$.
6. Calculate $\mathbf{y}_{it}, i = 1, \dots, 6$ using Equation 1.
7. Set $t = t + 1$ and return to 3.
8. Repeat until $t > T + 1000$.
9. Save the data point from 1001 to $T + 1000$.

For each $T = 50, 100,$ and $200,$ and measurement covariance matrix $\mathbf{Q} = \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3,$ and $\mathbf{Q}_4,$ 100 samples or replications were simulated using this procedure.

To analyze the simulated data, the four estimation methods were implemented in currently available software or programs. First, the data were analyzed by MLE using the Kalman filter (KF). To use this method, the DAFS was written as its state space representation for the simulated model by letting $\mathbf{S} = \mathbf{A}, \mathbf{a}_t = \mathbf{f}_t, \mathbf{z}_t = \mathbf{v}_t, \mathbf{Q} = \mathbf{Q},$ and $\mathbf{G} = \mathbf{D}.$ The free Fortran program MKFM6 written by Dolan (2005) was used to implement the estimation procedure.² Second, the data were analyzed by MLE using a block-Toeplitz lagged covariance matrix (BT). The block-Toeplitz matrix was calculated by DFA (IDHRM, 2004) and analyzed in *Mplus* (Muthén & Muthén, 2001). Third, the Bayesian method (BE) was used to estimate the model and implemented in the free software WinBUGS (Spiegelhalter, Thomas, Best, & Lunn, 2003). Finally, the least squares method (LS) was employed and implemented in the free software DyFA (Browne & Zhang, 2005).³

The first and third methods are based on the raw data and the second is based on the lagged covariance matrix. To identify the model, the variance of the random shocks were fixed at the true values. The parameter estimates can be compared with the true parameter values directly. The fourth method is based on the lagged correlation matrix. The program DyFA automatically calculates the lagged correlation matrix for the raw data and fixes the factor variance at lag 0 to be 1. Thus, the parameter estimates were rescaled for the aim of comparison. The following scaling formulas were used to rescale the parameter estimates from DyFA,⁴

$$\mathbf{A} \rightarrow \mathbf{E}\mathbf{A}\mathbf{F}$$

$$\mathbf{B}_1 \rightarrow \mathbf{F}^{-1}\mathbf{B}_1\mathbf{F}$$

$$\mathbf{Q} \rightarrow \mathbf{E}\mathbf{Q}\mathbf{E}$$

$$\mathbf{D} \rightarrow \mathbf{F}^{-1}\mathbf{D}\mathbf{F}^{-1},$$

²We are grateful to Conor V. Dolan, University of Amsterdam, for providing the simulation program.

³All the simulation and estimation codes can be obtained from <http://dfa.psychstat.org>.

⁴The authors are grateful to Michael W. Browne for the rescaling formulas.

where $\mathbf{E} = \text{Diag}^{1/2}(\mathbf{S})$ in which \mathbf{S} is the observed covariance matrix at lag 0 and $\mathbf{F} = \text{Diag}^{1/2}(d_{ii}/w_{ii})$ in which d_{ii} is the i th diagonal element of the shock covariance matrix \mathbf{D} and w_{ii} is the i th diagonal element of the desired matrix \mathbf{W} after rescaling.

Simulation Results

For each of the four estimation methods, we present the results from the simulation design involving the 12 conditions. The means of the parameter estimates from 100 replications for the four estimation methods are summarized in Tables 1 through 4. Visual inspection revealed that all four methods obtained close and accurate parameter estimates, compared to the true population values used to generate the data. Longer time series and smaller measurement errors led to more accurate parameter estimates.

To compare the four methods, we first calculated the absolute error (AE; Casella & Berger, 2001) or bias for each parameter. The AE for the i th parameter was calculated by $AE_i = |E(\theta_{ij}) - TP_i|$ with θ_{ij} representing the parameter estimate of the j th replication and TP_i representing the true value of the i th parameter to generate the data. AE measured the unbiasedness or accuracy of an estimation method. Because of the involvement of multiple parameters in our model, we calculated the total AE $TAE = \sum_{i=1}^{n_p} AE_i$ with n_p denoting the total number of parameters in the model, which is 17 in this study. We then calculated the mean square errors (MSE; Casella & Berger, 2001) for each parameter, which is defined as $MSE_i = E[(\theta_{ij} - TP_i)^2] = \text{Var}(\theta_i) + b^2(\theta_i)$ with $b(\theta_i) = E(\theta_i) - TP_i$ denoting the bias. MSE incorporates two components, one measuring the variability or precision and the other measuring the bias or accuracy of an estimation method. A smaller MSE means a better estimation method. Similar to TAE, we also calculated total MSE $TMSE = \sum_{i=1}^{n_p} MSE_i$. The TAEs and TMSEs are given in Table 5.

Just as with the visual inspection, with longer time series and smaller measurement errors, the bias became smaller based on TAEs. With the longest time series ($T = 200$) and the smallest measurement errors ($\mathbf{Q} = \mathbf{Q}_1$), the model parameters were almost exactly recovered. The maximum AE for all parameters was less than or equal to .03. Even with the shortest time series ($T = 50$) and the largest measurement errors ($\mathbf{Q} = \mathbf{Q}_4$), the results are still quite acceptable for all methods. The maximum absolute deviance was only .21 and the absolute deviance was less than .1 for most of the parameters. Overall, we can conclude the following order of accuracy for the four methods, $KF > LS > BT > BE$ based on the values of TAE. However, there was no significant difference between KF and LS especially when the measurement errors were small (based on paired t test).

TABLE 1
Parameter Estimates from MKFM6 Using MLE Method via Kalman Filter

Labels of Model Parameters	True Population Parameter Values	Amount of Measurement Errors											
		0.1			0.3			0.5			1.0		
		Time Series Length			Time Series Length			Time Series Length			Time Series Length		
		50 ^a	100 ^a	200 ^a	50 ^b	100 ^a	200 ^a	50 ^c	100 ^d	200 ^e	50 ^f	100 ^g	200 ^e
λ_{11}	1	0.98	1.00	1.00	0.98	0.98	1.01	0.94	0.96	0.98	0.99	0.98	0.99
λ_{21}	1	0.99	1.00	1.00	0.98	0.99	1.01	0.95	0.97	0.98	0.96	0.95	0.97
λ_{31}	1	0.98	0.99	1.00	0.98	0.98	1.00	0.93	0.97	0.98	0.99	0.96	0.97
λ_{42}	1	0.97	0.99	0.99	0.95	0.96	0.97	0.93	0.97	0.99	0.97	0.99	0.98
λ_{52}	1	0.97	0.98	0.99	0.96	0.97	0.98	0.94	0.98	0.99	0.98	1.00	0.99
λ_{62}	1	0.97	0.98	0.98	0.96	0.98	0.99	0.91	0.96	0.99	1.00	1.01	0.99
q_1		0.10	0.10	0.10	0.27	0.29	0.29	0.50	0.50	0.51	0.97	0.99	0.99
q_2		0.10	0.10	0.10	0.29	0.29	0.29	0.47	0.49	0.49	0.97	0.97	0.98
q_3		0.10	0.10	0.10	0.29	0.30	0.30	0.49	0.50	0.50	0.97	1.00	1.00
q_4		0.10	0.10	0.10	0.29	0.30	0.30	0.48	0.49	0.49	0.93	0.97	0.99
q_5		0.10	0.10	0.10	0.29	0.29	0.31	0.49	0.51	0.51	0.93	1.00	1.00
q_6		0.10	0.10	0.10	0.29	0.30	0.30	0.53	0.51	0.51	0.93	0.97	0.98
b_{11}	0.8	0.73	0.77	0.78	0.72	0.77	0.79	0.74	0.78	0.80	0.69	0.76	0.78
b_{12}	0	0.04	0.02	0.00	0.02	-0.01	0.01	-0.01	-0.01	0.00	0.03	0.00	0.00
b_{21}	0	-0.03	-0.01	-0.01	0.00	-0.01	0.00	0.02	0.01	0.01	0.02	0.01	0.01
b_{22}	0.8	0.77	0.78	0.79	0.74	0.77	0.79	0.72	0.75	0.78	0.68	0.74	0.78
d_{12}	0.18	0.18	0.18	0.18	0.17	0.18	0.18	0.16	0.17	0.17	0.16	0.16	0.17

^aBased on 100 replications. ^bBased on 93 replications. ^cBased on 90 replications. ^dBased on 97 replications. ^eBased on 99 replications. ^fBased on 81 replications. ^gBased on 91 replications.

TABLE 2
Parameter Estimates From *Mplus* Using Maximum Likelihood Estimation Method Based on Block-Toeplitz Lagged Covariance Matrix

Labels of Model Parameters	True Population Parameter Values	Amount of Measurement Errors											
		0.1			0.3			0.5			1.0		
		Time Series Length			Time Series Length			Time Series Length			Time Series Length		
		50 ^a	100 ^a	200 ^a	50 ^a	100 ^a	200 ^a	50 ^a	100 ^a	200 ^a	50 ^b	100 ^c	200 ^a
λ_{11}	1	0.98	0.99	0.99	0.98	0.97	1.00	0.90	0.95	0.97	1.03	0.96	0.98
λ_{21}	1	0.99	1.00	1.00	0.97	0.98	1.00	0.91	0.96	0.97	1.04	0.96	0.97
λ_{31}	1	0.98	0.99	1.00	0.97	0.96	0.99	0.91	0.97	0.97	1.05	0.94	0.96
λ_{42}	1	0.97	0.99	0.99	0.95	0.96	0.96	0.95	0.98	0.99	1.05	0.99	0.97
λ_{52}	1	0.97	0.98	0.99	0.96	0.96	0.96	0.97	1.01	1.00	1.00	0.97	0.97
λ_{62}	1	0.97	0.98	0.98	0.97	0.98	0.97	0.94	0.98	1.00	1.05	1.00	0.98
q_1		0.09	0.10	0.10	0.27	0.28	0.29	0.48	0.50	0.50	0.94	0.97	0.97
q_2		0.09	0.10	0.10	0.29	0.29	0.29	0.46	0.49	0.49	0.91	0.94	0.96
q_3		0.09	0.10	0.10	0.28	0.29	0.30	0.47	0.48	0.49	0.92	0.98	0.98
q_4		0.09	0.10	0.10	0.28	0.29	0.30	0.46	0.49	0.49	0.82	0.93	0.97
q_5		0.10	0.10	0.10	0.29	0.29	0.30	0.47	0.50	0.50	0.92	0.98	0.98
q_6		0.09	0.10	0.10	0.28	0.29	0.30	0.51	0.50	0.50	0.88	0.94	0.98
b_{11}	0.8	0.68	0.75	0.77	0.70	0.76	0.78	0.71	0.76	0.78	0.66	0.73	0.77
b_{12}	0	0.05	0.02	0.00	0.01	-0.01	0.00	-0.07	-0.03	0.00	0.00	0.01	0.00
b_{21}	0	-0.02	-0.01	-0.01	0.00	-0.01	0.00	-0.01	0.00	0.01	-0.02	-0.01	-0.01
b_{22}	0.8	0.73	0.76	0.78	0.71	0.75	0.78	0.67	0.73	0.77	0.63	0.73	0.78
d_{12}	0.18	0.19	0.18	0.19	0.19	0.19	0.19	0.20	0.18	0.18	0.19	0.19	0.18

^aBased on 100 replications. ^bBased on 88 replications. ^cBased on 99 replications.

TABLE 3
Parameter Estimates From WinBUGS Using Bayesian Estimation Method

Labels of Model Parameters	True Population Parameter Values	Amount of Measurement Errors											
		0.1			0.3			0.5			1.0		
		Time Series Length			Time Series Length			Time Series Length			Time Series Length		
		50 ^a	100 ^a	200 ^a	50 ^a	100 ^a	200 ^a	50 ^b	100 ^b	200 ^a	50 ^c	100 ^b	200 ^a
λ_{11}	1	1.07	1.04	1.02	1.10	1.04	1.03	1.08	1.03	1.01	1.21	1.07	1.05
λ_{21}	1	1.08	1.05	1.02	1.10	1.05	1.03	1.10	1.05	1.02	1.19	1.06	1.03
λ_{31}	1	1.07	1.04	1.02	1.10	1.04	1.03	1.08	1.05	1.02	1.21	1.07	1.03
λ_{42}	1	1.06	1.04	1.01	1.07	1.02	1.00	1.10	1.05	1.03	1.18	1.09	1.04
λ_{52}	1	1.06	1.03	1.01	1.08	1.02	1.00	1.12	1.07	1.03	1.16	1.09	1.04
λ_{62}	1	1.06	1.02	1.01	1.09	1.03	1.01	1.08	1.05	1.03	1.21	1.10	1.05
q_1		0.10	0.10	0.10	0.29	0.30	0.30	0.53	0.52	0.51	1.03	1.02	1.00
q_2		0.10	0.10	0.10	0.32	0.30	0.29	0.49	0.51	0.50	1.01	1.00	0.99
q_3		0.11	0.10	0.10	0.31	0.31	0.30	0.52	0.51	0.50	1.00	1.02	1.01
q_4		0.11	0.10	0.11	0.31	0.31	0.30	0.50	0.51	0.50	0.96	0.98	1.00
q_5		0.11	0.10	0.10	0.32	0.30	0.31	0.52	0.53	0.52	1.01	1.02	1.00
q_6		0.10	0.10	0.10	0.31	0.31	0.31	0.56	0.53	0.52	0.96	0.98	1.00
b_{11}	0.8	0.71	0.76	0.78	0.67	0.75	0.78	0.65	0.75	0.78	0.56	0.68	0.76
b_{12}	0	0.04	0.02	0.00	0.04	0.00	0.01	0.03	0.00	0.01	0.06	0.04	0.01
b_{21}	0	-0.02	-0.01	0.00	-0.01	0.00	0.01	-0.01	0.00	0.01	0.03	0.00	0.01
b_{22}	0.8	0.74	0.77	0.79	0.69	0.75	0.78	0.63	0.72	0.77	0.57	0.68	0.75
d_{12}	0.18	0.17	0.18	0.18	0.16	0.17	0.18	0.14	0.16	0.17	0.12	0.14	0.16

^aBased on 100 replications. ^bBased on 99 replications. ^cBased on 93 replications.

TABLE 4
Parameter Estimates From DyFA Using the Least Squares Method

Labels of Model Parameters	True Population Parameter Values	Amount of Measurement Errors											
		0.1			0.3			0.5			1.0		
		Time Series Length			Time Series Length			Time Series Length			Time Series Length		
		50 ^a	100 ^a	200 ^a	50 ^a	100 ^a	200 ^a	50 ^a	100 ^a	200 ^a	50 ^a	100 ^a	200 ^a
λ_{11}	1	1.02	1.01	1.01	0.99	0.99	1.01	0.93	0.96	0.98	0.99	0.98	0.99
λ_{21}	1	1.03	1.02	1.01	0.99	0.99	1.01	0.94	0.98	0.99	1.00	0.97	0.97
λ_{31}	1	1.01	1.01	1.01	1.00	0.98	1.01	0.94	0.98	0.99	1.02	0.96	0.97
λ_{42}	1	1.01	1.01	1.00	0.99	0.99	0.98	0.97	1.00	1.01	1.01	0.99	0.98
λ_{52}	1	1.00	1.00	1.00	0.99	0.98	0.97	1.00	1.03	1.01	0.98	0.97	0.98
λ_{62}	1	1.01	1.00	0.99	0.99	1.00	0.99	0.96	1.00	1.00	1.03	0.99	0.99
q_1		0.09	0.10	0.09	0.28	0.29	0.29	0.49	0.50	0.51	0.96	0.98	0.99
q_2		0.10	0.10	0.10	0.30	0.30	0.29	0.46	0.49	0.49	0.93	0.96	0.98
q_3		0.10	0.10	0.10	0.27	0.29	0.30	0.49	0.49	0.49	0.93	0.99	0.99
q_4		0.10	0.10	0.10	0.28	0.29	0.30	0.47	0.49	0.49	0.88	0.94	0.99
q_5		0.10	0.10	0.10	0.29	0.29	0.30	0.48	0.50	0.50	0.93	1.00	0.98
q_6		0.09	0.10	0.10	0.29	0.30	0.30	0.52	0.51	0.51	0.89	0.96	0.99
b_{11}	0.8	0.67	0.74	0.77	0.69	0.76	0.77	0.70	0.76	0.78	0.66	0.73	0.77
b_{12}	0	0.05	0.02	0.01	0.01	-0.01	0.00	-0.05	-0.02	0.00	0.02	0.01	0.00
b_{21}	0	-0.02	-0.01	0.00	0.01	-0.01	0.00	0.00	0.01	0.01	-0.02	-0.03	0.00
b_{22}	0.8	0.72	0.76	0.78	0.69	0.74	0.78	0.65	0.73	0.77	0.65	0.73	0.78
d_{12}	0.18	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.18	0.18	0.18	0.18	0.18

^aBased on 100 replications.

TABLE 5
Comparisons of the Four Estimation Methods

		<i>Amount of Measurement Errors</i>											
		<i>0.1</i>			<i>0.3</i>			<i>0.5</i>			<i>1.0</i>		
		<i>Time Series Length</i>			<i>Time Series Length</i>			<i>Time Series Length</i>			<i>Time Series Length</i>		
		<i>50</i>	<i>100</i>	<i>200</i>	<i>50</i>	<i>100</i>	<i>200</i>	<i>50</i>	<i>100</i>	<i>200</i>	<i>50</i>	<i>100</i>	<i>200</i>
TAE	KF	0.30	0.15	0.09	0.42	0.25	0.14	0.70	0.34	0.19	0.68	0.36	0.22
	BT	0.44	0.20	0.13	0.53	0.37	0.21	0.92	0.35	0.19	1.18	0.59	0.38
	BE	0.65	0.34	0.16	0.93	0.34	0.21	1.11	0.54	0.25	1.90	0.88	0.38
	LS	0.40	0.20	0.12	0.37	0.25	0.18	0.71	0.29	0.18	0.91	0.49	0.26
TMSE	KF	0.17	0.07	0.04	0.29	0.15	0.10	0.62	0.35	0.11	3.92	3.31	0.21
	BT	0.22	0.08	0.04	0.36	0.17	0.08	0.88	0.29	0.14	1.81	0.75	0.33
	BE	0.21	0.08	0.04	0.39	0.16	0.07	0.64	0.26	0.11	1.60	0.61	0.23
	LS	0.23	0.09	0.04	0.37	0.18	0.08	0.71	0.30	0.14	1.63	0.76	0.32
MASE	KF	0.08	0.06	0.04	0.12	0.08	0.06	0.15	0.10	0.07	0.23	0.15	0.11
	BT	0.08	0.06	0.06	0.08	0.08	0.08	0.14	0.10	0.10	0.18	0.17	0.16
	BE	0.09	0.06	0.04	0.13	0.09	0.06	0.17	0.11	0.07	0.26	0.17	0.11
	LS ^a	0.10	0.06	0.05	0.14	0.10	0.07	0.19	0.13	0.09	0.30	0.20	0.13
TDE	KF	0.06	0.00	-0.01	0.11	0.12	0.22	0.46	0.46	0.10	0.38	0.24	0.01
	BT	0.19	0.04	-0.23	0.89	0.24	-0.27	1.21	0.37	-0.28	2.17	0.60	-0.45
	BE	-0.03	-0.05	0.01	-0.02	0.07	0.01	0.04	0.08	0.03	0.07	0.13	-0.01
	LS	—	—	—	—	—	—	—	—	—	—	—	—

Note. TAE = total absolute error; TMSE = total mean square error; MASE = mean average standard error; TDE = total difference error; KF = Kalman filter method; BT = block-Toeplitz based on maximum likelihood estimation; BE = Bayesian method; LS = least squares method.

^aBased on the standard deviations of the parameter estimates.

Furthermore, there were no big differences when the number of time data points was large ($T = 200$). BT and BE had larger bias than the other two methods. The bias for BT mainly came from the estimation of measurement error variances and the bias for BE mainly came from the estimation of factor loadings.

TMSE can be used as a measure of overall performance of parameter estimation. From TMSE, we found again that there were no large differences among the four estimation methods. However, the increase of measurement error resulted in a quick increase in TMSEs for all methods, especially for the KF method. Note that although the bias (TAEs) of BE was the largest among the four estimation methods, the TMSEs were comparable or smaller than those of the other methods, meaning that BE was more precise.

To compare the efficiency of the four methods, the average standard error (ASE) was calculated.⁵ Here, $ASE_i = 1/n_r \sum_{j=1}^{n_r} SE_{ij}$ with SE_{ij} representing the standard error for i th parameter from the j th replication and n_r denoting the total number of replications, which is 100 in this study. To summarize information, we calculated the mean ASE by $MASE = 1/n_p \sum_{i=1}^{n_p} ASE_i$. Estimation methods with higher efficiency will have smaller MASE values. The MASEs for different methods in each condition are given in Table 5.⁶ There was no noticeable difference for the four methods based on MASE, which means that the efficiencies were comparable to each other. However, it was very clear that short time series and large measurement errors resulted in less efficient parameter estimates (larger MASE).

To evaluate the consistency of the standard error estimation, the total different error (TDE) was constructed. The $TDE = \sum_{i=1}^{n_p} (SD_i - ASE_i)$, with SD_i denoting the standard deviation for i th parameter across all the replications. A TDE larger than zero implies the underestimation of the standard errors and a TDE less than zero implies the overestimation of the standard errors. The TDEs for different methods in each condition are given in Table 5. Overall, the BE method obtained consistent estimation of standard errors in all conditions. The KF method also obtained consistent estimation of standard errors when the measurement errors were small ($\mathbf{Q} = \mathbf{Q}_1$). However, when the measurement errors became large, the standard errors were underestimated. For the BT method, when the time series was short, the standard errors were underestimated but when the time series was long, the standard errors were overestimated.

⁵The ASE and TDE are created for the convenience of comparisons. We acknowledge they may have limitations. However, in terms of summarizing information and facilitating comparisons, they did work well here.

⁶Because DyFA does not provide standard error estimation at current stage, we used standard deviations to represent ASE.

DISCUSSION

Four methods for estimating the DAFS model—the MLE method using the Kalman filter, the MLE method based on lagged block-Toeplitz covariance matrix, the Bayesian method via Gibbs sampling, and the least squares method based on lagged correlation matrix—were summarized and compared. To numerically evaluate and compare the four methods, a simulation study with varying length of time series and amount of measurement errors was conducted. Although we focused only on the one-lag DAFS model for demonstration purposes, the same procedure works for more complex models because any DAFS model can be rewritten as a one-lag model. The simulation results showed all four methods yielded acceptable parameter estimates with comparable accuracy in almost all conditions.

Because the differences in the four estimation methods have been compared through the technical form and simulation study, the discussion only focuses on our experience with different software or programs in estimating DAFS models. MKFM6 is a free open-source program, which means interested researchers can even modify the program for their specific uses. Although MKFM6 is customized for the state space model, it can be used as a special program for estimating the DAFS model because we can rewrite any DAFS model using its state space representation. MKFM6 is relatively easy to use, given some understanding of the relationship between the DAFS model and the state space model. As shown earlier, both the estimates of parameters and associated standard errors were consistent. However, we found that the program may produce a negative log-likelihood value and unreasonable standard errors, even when the process converged and the parameter estimates were very good. Furthermore, when the time series was short and measurement error was large, MKFM6 may have problems reaching convergence (Table 1). Finally, although any DAFS model can be rewritten in its state space format theoretically, with the increase of lags, the matrix \mathbf{S} , \mathbf{H} , and \mathbf{G} will become very large, rendering the estimation procedure impractical.

After calculating the block-Toeplitz lagged covariance matrix by the program DFA, any SEM software can be used to obtain the parameter estimates. We employed *Mplus* because of its flexibility in specifying constraints of parameters to ensure covariance stationarity and its flexibility in performing simulations. *Mplus* is a good choice for estimating the DAFS model for those who are familiar with SEM. However, *Mplus* itself cannot calculate the lagged covariance matrix. We suggest using the DFA program to calculate the lagged covariance because it can also be used to investigate the similarity of more than one lagged covariance matrix and pool information of the participants with similar lagged covariance matrices when the time series is short (Nesselroade & Molenaar, 1999). However, because this method itself violates the assumption of independent observations,

the fit statistics from *Mplus* are not applicable. Furthermore, with more lags in the model, the constraints can become very complex. In this situation, *Mx*, a more flexible program regarding constraints, may be used (Neale, Boker, Xie, & Maes, 1999).

WinBUGS is also free software. Because it is a general program for Bayesian analysis, it demands the most programming work in estimating the DAFS model. Fortunately, the WinBUGS codes can be modified easily for different models. Furthermore, WinBUGS is the most flexible software for estimating more complex dynamic factor models, such as the categorical DAFS model and WNFS model (Zhang & Nesselroade, 2007) and nonstationary dynamic models. Finally, WinBUGS is also the only program that can incorporate the prior information into the model, although how to appropriately implement this awaits further investigation. However, the computation time of WinBUGS is long and increases quickly with the length of time series. For example, it took approximately 6 min when $T = 100$ and about 12 min when $T = 200$ to run one replication. In contrast, it only took the other programs only a few seconds to analyze one replication of the longest time series.

Among all four programs, DyFA, which is another free program, is the only one especially written for estimating DAFS models. Its input scripts are directly related to the model and it is the easiest program to use. Besides the estimation of the parameters reported in this study, DyFA also has many other distinctive and useful features, such as calculating the correlation matrix of endogenous factors and squared multiple correlations of factors at the current time point on previous time points, checking the stationarity, and conducting exploratory dynamic factor analysis with factor rotation. DyFA is highly recommended for beginners who would like to investigate the merits of dynamic factor models. The current version of DyFA does not calculate the standard errors of the parameter estimates, but an additional bootstrap method can be used to obtain them.⁷ However, DyFA is less flexible in specifying the models. For example, changing model constraints is more complex than in some of the other software we used.

Finally, we want to draw readers' attention to several important aspects of the DAFS model and estimation methods still requiring further investigation. First, missing data are often present in time series, especially in psychological research. Fortunately, all four estimation methods and programs can deal with missing data easily through multiple imputation (e.g., WinBUGS and MKFM6) or pairwise deletion (e.g., *Mplus* and DyFA). Second, in terms of model fitting, different fit indexes can be used. For MKFM6, the log-likelihood can be used to

⁷In personal communication with both Guanjian Zhang and Michael W. Browne, they suggested the use of bootstrap for obtaining standard errors. We hope to see this feature in the next version of DyFA.

measure model fit. For WinBUGS, deviance information criterion (Spiegelhalter, Best, Carlin, & Linde, 2002). For *Mplus* and DyFA, appropriate fit indexes still need to be developed. Third, in this study, only single-subject data were analyzed. However, multiple-subject analysis can be easily implemented for all four methods through multiple group analysis by treating each individual as a group. Furthermore, the similarities and differences can be evaluated through comparison of the estimated model parameters of each subject.

As one can see, the four estimation methods and software each have their own advantages and disadvantages. It is hard to say which method is the best because all methods can obtain comparable parameter estimates. Thus, the choice of the four methods is mainly based on the user's familiarity with each method and software.

REFERENCES

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of ARMA time series models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Psychometrics: A festschrift to Roderick P. McDonald* (pp. 415–452). Mahwah, NJ: Lawrence Erlbaum Associates.
- Browne, M. W., & Zhang, G. (2005). DyFA: Dynamic factor analysis of lagged correlation matrices, Version 2.00 [Computer software and manual]. Retrieved December 1, 2004 from <http://quantrm2.psy.ohio-state.edu/browne/>
- Browne, M. W., & Zhang, G. (2006). Developments in the factor analysis of individual time series. In R. Cudek & R. C. MacCallum (Eds.) *Factor analysis at 100: Historical developments and future directions* (pp. 265–291). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Carlin, B. P., Polson, N. G., & Stoffer, D. S. (1992). A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association*, *87*, 493–500.
- Carter, C. K., & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, *81*, 541–553.
- Casella, G., & Berger, R. L. (2001). *Statistical inference* (2nd ed.). Brooks/Cole.
- Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika*, *12*, 267–288.
- Chatfield, C. (2004). *The analysis of time series: An introduction* (6th ed.). London: Chapman & Hall.
- Dolan, C. V. (2005). *MKFM6: Multi-group, multi-subject stationary timeseries modeling based on the Kalman filter*. Retrieved December 1, 2004 from <http://users.fmg.uva.nl/cdolan/>
- Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods*. New York: Oxford University Press.
- Engle, R., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association*, *76*, 774–781.
- Geweke, J. F., & Singleton, K. J. (1981). Maximum likelihood “confirmatory” factor analysis of economic time series. *International Economic Review*, *22*, 37–54.
- Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2002). On the nature of SEM estimates of ARMA parameters. *Structural Equation Modeling*, *9*, 347–368.
- Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

- Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter series*. Cambridge, UK: Cambridge University Press.
- Hershberger, S. L. (1998). Dynamic factor analysis. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 217–249). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- IDHRM. (2004). *DFA: Dynamic factor analysis*. Retrieved December 1, 2004 from <http://www.people.virginia.edu/~zz5m>
- Jöreskog, K. G., Sörbom, D., Du Toit, S., & Du Toit, M. (2001). *LISREL 8: New statistical features*. Chicago: Scientific Software International.
- Justiniano, A. (2004). *Estimation and model selection in dynamic factor analysis*. Retrieved December 1, 2004 from www.princeton.edu/~jalejand
- Kadiyala, K. R., & Karlsson, S. (1997). Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics*, *12*, 99–132.
- Kim, C. J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge, MA: MIT Press.
- Kim, C. J., & Nelson, C. R. (2001). A Bayesian approach to testing for Markov-switching in univariate and dynamic factor models. *International Economic Review*, *42*, 989–1013.
- McArdle, J. J. (1982). *Structural equation modeling of an individual system: Preliminary results from "A case study in episodic alcoholism."* Unpublished manuscript, Department of Psychology, University of Denver, Denver, CO.
- Molenaar, P. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, *50*, 181–202.
- Molenaar, P., & Nesselroade, J. R. (1998). A comparison of pseudo-maximum likelihood and asymptotically distribution-free dynamic factor analysis parameter estimation in fitting covariance-structure models to block-Toeplitz matrices representing single-subject multivariate time-series. *Multivariate Behavioral Research*, *33*, 313–342.
- Muthén, L., & Muthén, B. (2001). *Mplus user's guide*. Los Angeles: Muthén and Muthén.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *MX: Statistical modeling* (5th ed.). Richmond: Virginia Commonwealth University.
- Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. (2001). Dynamic factor analysis models for multivariate time series analysis. In D. M. Moscovitz & S. L. Hershberger (Eds.) *Modeling individual variability with repeated measures data: Advances & techniques* (pp. 235–265). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time-series for dynamic factor analysis. In R. H. Hoyle (Ed.) *Statistical strategies for small sample research* (pp. 223–250). Newbury Park, CA: Sage.
- Nesselroade, J. R., & Molenaar, P. C. M. (2002). Applying dynamic factor analysis in behavioral and social science research. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 335–344). Newbury Park, CA: Sage.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. V. D. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*, 583–639.
- Spiegelhalter, D. J., Thomas, A., Best, N., & Lunn, D. (2003). *WinBUGS Manual version 1.4*. Cambridge, UK: MRC Biostatistics Unit, Institute of Public Health. Retrieved December 1, 2004 from <http://www.mrc-bsu.cam.ac.uk/bugs>
- West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics*, *18*, 338–357.
- Wood, P., and Brown, D. (1994). The study of intraindividual differences by means of dynamic factor models: Rationale, implementation, and interpretation. *Psychological Bulletin*, *116*, 166–186.
- Zhang, Z., & Nesselroade, J. R. (2007). Bayesian estimation of categorical dynamic factor models. *Multivariate Behavioral Research*, *42*, 729–756.