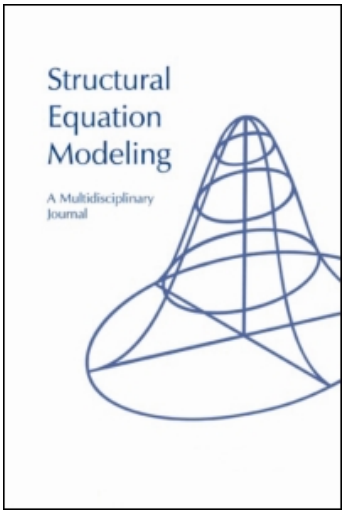


Structural Equation Modeling: A Multidisciplinary Journal
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653699

Equivalence and Differences Between Structural Equation Modeling and State-Space Modeling Techniques

Sy-Miin Chow [a]; Moon-ho R. Ho [b]; Ellen L. Hamaker [c];Conor V. Dolan [d]
[a] University of North Carolina at Chapel Hill, [b] Nanyang Technological University, [c] Utrecht University, [d] University of Amsterdam,

Online publication date: 19 April 2010

To cite this Article Chow, Sy-Miin , Ho, Moon-ho R. , Hamaker, Ellen L. andDolan, Conor V.(2010) 'Equivalence and Differences Between Structural Equation Modeling and State-Space Modeling Techniques', Structural Equation Modeling: A Multidisciplinary Journal, 17: 2, 303 — 332
To link to this Article: DOI: 10.1080/10705511003661553
URL: http://dx.doi.org/10.1080/10705511003661553

Structural Equation Modeling, 17:303–332, 2010

ISSN: 1070-5511 print/1532-8007 online
DOI: 10.1080/10705511003661553

Psychology Press
Taylor & Francis Group

Equivalence and Differences Between Structural Equation Modeling and State-Space Modeling Techniques

Sy-Miin Chow
University of North Carolina at Chapel Hill

Moon-ho R. Ho
Nanyang Technological University

Ellen L. Hamaker
Utrecht University

Conor V. Dolan
University of Amsterdam

State-space modeling techniques have been compared to structural equation modeling (SEM) techniques in various contexts but their unique strengths have often been overshadowed by their similarities to SEM. In this article, we provide a comprehensive discussion of these 2 approaches' similarities and differences through analytic comparisons and numerical simulations, with a focus on their use in representing intraindividual dynamics and interindividual differences. To demonstrate the respective strengths and weaknesses of the 2 approaches in representing these 2 aspects, we simulated data under (a) a cross-sectional common factor model, (b) a latent difference score model with random effects in intercept and slope, and (c) a bivariate dynamic factor analysis model with auto- and cross-regression parameters. Possible ways in which SEM and state-space modeling can be utilized as complementary tools in representing human developmental and other related processes are discussed.

The widespread popularity of structural equation modeling (SEM) software such as LISREL (Jöreskog & Sörbom, 1981) and Mx (Neale, Boker, Xie, & Maes, 1999) has helped spur increased complexity and sophistication in the types of research questions tested by social and behavioral scientists in the past few decades. On a positive note, tremendous growth has been observed in areas in which the use of SEM is deemed practical. Unfortunately, the

Correspondence should be addressed to Sy-Miin Chow, Department of Psychology, CB#3270 Davie Hall, University of North Carolina, Chapel Hill, NC 27599-3270, USA. E-mail: symiin@email.unc.edu

popularity of SEM has also limited researchers' interest in exploring models that cannot be readily implemented using SEM software.

Owing in part to the popularity of hierarchical linear modeling (Bryk & Raudenbush, 1987), there has been a recent revival of interest in collecting intensive repeated measures or time series data (e.g., using experience sampling and daily diary designs) among social and behavioral scientists. In the related area of longitudinal modeling, the specific formulation of SEM makes it ideally suited for handling panel data measured over relatively few measurement occasions but across many participants. However, because different measurement occasions for a single variable are typically input as different manifest variables in standard SEM practice, several estimation issues arise in SEM when the number of measurement occasions, T, exceeds the number of participants, N. Some alternatives have been developed to handle this kind of data, including using the block-Toeplitz method to reduce the dimension of the input matrix (Browne & Nesselroade, 2005; Hershberger, Corneal, & Molenaar, 1994; Molenaar, 1985), and using raw data maximum likelihood with special parameter constraints to ensure the positive definiteness of the model-implied covariance matrix (e.g., Hamaker, Dolan, & Molenaar, 2003). Each of these approaches brings with it some inherent modeling and estimation issues, some of which are discussed in this article.

In contrast to SEM, state-space modeling techniques are rooted in a very different modeling tradition. They were originally developed as single–subject time series estimation tools (e.g., Kalman, 1960), and have later been adapted for use with panel data (e.g., Oud & Jansen, 2000; Singer, 1998) and multiple-subject repeated measures data (e.g., Dolan, 2002). With the exceptions of the econometric literature and the work of a few others (e.g., Gregson, 1983; Heath, 2000; Ho, Shumway & Ombao, 2006; Molenaar & Raijmakers, 1998; Oud, van den Bereken, & Essers, 1990), state-space modeling techniques have received relatively little attention from most social and behavioral scientists despite their long history in engineering and economics. In the few instances where state-space techniques were introduced to a broader audience through their parallels with SEM (e.g., MacCallum & Ashby, 1986), the two approaches' striking similarities in the (more restrictive) cross-sectional case have led subsequent researchers to ignore their differences. As a result, many recent methodological developments within the state-space framework remain unfamiliar to most psychologists. This is an unfortunate (and perhaps, an unexpected) consequence because state-space techniques' unique appeal as a tool for handling time series data can help avoid some of the aforementioned SEM-related estimation issues.

In this article, we seek to clarify the differences between the two approaches while providing a synthesis of previous perspectives on the equivalence between the two (e.g., MacCallum & Ashby, 1986; Molenaar, 1985; Otter, 1986; Oud et al., 1990). We do so by using both analytical and simulation examples to illustrate their relative merits in addressing questions pertaining to intraindividual change and interindividual differences in intraindividual change. To conclude, we discuss several key methodological developments proposed in the state-space literature and suggest possible extensions of such techniques from an SEM perspective.

THE SEM FRAMEWORK

SEM is generally concerned with the modeling of variances and covariances and possibly the means of multivariate observations (Bollen, 1989). In this section we begin by presenting

the two equations that define the structural equation model, and the corresponding model-implied covariance and means structures. In addition, we summarize how maximum likelihood estimates of these parameters can be obtained. To differentiate between parameter matrices from the SEM framework and those from the state-space framework, we use the subscript s to identify elements associated with the SEM approach and the subscript k to denote elements within a state-space framework throughout this article.

Model Formulation

In an SEM framework, the relationships among latent and manifest variables are specified using a measurement model and a structural model. The measurement model is written as

$$y_i = \tau_s + \Lambda_s \eta_i + \epsilon_i, \tag{1}$$

where i is the subject index, y_i is a $p_s \times 1$ vector of manifest variables for person i, η_i is a $w_s \times 1$ vector of latent variables, and ϵ_i is a $p_s \times 1$ vector of measurement errors. All the parameters are held invariant across persons in the present context. These parameters include the vector of constant intercepts, τ_s, a $p_s \times w_s$ matrix of factor loadings, Λ_s, and Θ_s, the $p_s \times p_s$ covariance matrix of ϵ_i. In addition, it is assumed that the mean of ϵ_i across individuals is a $p_s \times 1$ vector of zeros. Because different measurement occasions of the same variable are typically incorporated into an SEM model by expanding the dimension of y_i (in contrast to the state-space formulation presented in next section, which is based explicitly on y_{it}), the number of manifest variables in an SEM model is denoted specifically as p_s and the corresponding number of latent variables is denoted as w_s.

The structural model specifies the relationships among latent variables as

$$\eta_i = \alpha_s + B_s \eta_i + \zeta_i, \tag{2}$$

where ζ_i is a $w_s \times 1$ vector of residuals. The parameters to be estimated include α_s, a $w_s \times 1$ vector of constant intercepts, B_s, a $w_s \times w_s$ matrix of regression coefficients and Ψ_s, a $w_s \times w_s$ covariance matrix of ζ_i.

In fitting longitudinal models to data with T measurement occasions, all T measurement occasions have to be incorporated into y_i and η_i. The sizes of all the associated parameter matrices or vectors will also change accordingly.[1]

The $p_s \times p_s$ population covariance matrix, Σ, is expressed as

$$\Sigma = \Lambda_s (I - B_s)^{-1} \Psi_s (I - B_s)^{-1'} \Lambda_s' + \Theta_s, \tag{3}$$

and the vector of structured population means, μ, is given by

$$\mu = \tau_s + \Lambda_s (I - B_s)^{-1} \alpha_s. \tag{4}$$

[1] Equation 1 is sometimes further partitioned into two equations to express the measurement model separately for the dependent (i.e., endogenous) and independent (i.e., exogenous) latent variables. We use a simpler expression here to better highlight the parallels between SEM models and state-space models. Equations 1 and 2 can also be combined into a single equation involving latent and manifest variables (using the reticular action model or RAM notations; McArdle & McDonald, 1984) but this simplified form is not considered here.

Parameter Estimation in SEM

In most SEM programs, a likelihood estimation procedure based on the assumption of multivariate normality is used to estimate model parameters. This procedure has the advantage that the observed covariance matrix and means are sufficient statistics so that raw data are not needed for estimation purposes (Bollen, 1989). In addition, many SEM programs also include a raw data likelihood function, primarily for handling missing data (Enders, 2001; Graham, Hofer, & MacKinnon, 1996; Schafer & Graham, 2002). This raw data likelihood function has been used to analyze repeated measures data when the number of time points exceeds the number of participants (Hamaker et al., 2003). As we show later, the similarities and differences between SEM and state-space modeling can also be highlighted more directly using the raw data likelihood function.

When the raw data maximum likelihood, also known as full information maximum likelihood (FIML; Enders, 2001; Graham et al., 1996), approach is used with data from N subjects, the log-likelihood function of a structural equation model can be obtained by summing the log-likelihood function contributed by each individual's data vector as (Bollen, 1989; Enders, 2001)

$$LL_{RML}(\theta_s) = \frac{1}{2}\sum_{i=1}^{N}[-p_s \log(2\pi) - \log|\Sigma| - (y_i - \mu)'\Sigma^{-1}(y_i - \mu)], \tag{5}$$

where p_s, the number of complete manifest variables, can become person–dependent if different subjects have different patterns of missingness,[2] Σ and μ are the population covariance matrix and the structured means vector expressed in Equations 3 and 4, respectively. Maximum likelihood estimates of the parameters are obtained by maximizing the function $LL_{RML}(\theta_s)$ with respect to the parameters of interest.

Under the assumption of multivariate normality, the log-likelihood function in Equation 5 can be rewritten into a log-likelihood ratio function expressed as (Bollen, 1989)

$$F_{ML}(\theta_s) = \frac{1}{2}[\log|\Sigma| + tr(S\Sigma^{-1}) - \log|S| - p] + \frac{1}{2}[(m - \mu)'\Sigma^{-1}(m - \mu)], \tag{6}$$

where S is the $p_s \times p_s$ observed covariance matrix and m is the corresponding $p_s \times 1$ vector of observed means of the manifest variables. Equation 6 represents the difference between the log-likelihood function of the model under study, and the log-likelihood function of a perfectly saturated model (i.e., a model in which $\Sigma = S$ and $\mu = m$). The value in Equation 6 multiplied by $-2 * (N - 1)$ yields a chi-square statistic for evaluating model fit (Bollen, 1989).

THE STATE-SPACE MODELING FRAMEWORK

The state-space framework encompasses at its root the Kalman filter (KF) and the related Kalman smoothers, which have been used in econometrics and engineering to track changes in dynamic systems and to implement diverse time series models. Simply speaking, the KF can be regarded as a factor score—or more broadly, a "latent variable"—estimation procedure. One

[2]In this case, a filter matrix needs to be used to reduce Σ and μ to the appropriate dimensions for each person.

common way of estimating parameters in the state-space framework requires the execution of two steps: the KF is first run to obtain the state estimates conditional on a set of parameters; by-products of the Kalman filter are then substituted into a likelihood function to be optimized by repeated iterations between these two steps until some predetermined convergence criteria are met, in which case maximum likelihood parameter estimates are obtained (Shumway, 2000). We begin by presenting the state-space model, followed by the KF algorithm and the parameter estimation procedure.

Model Formulation

Analogous to the specifications used in SEM programs, state-space models are used to represent the latent states of a system in terms of an observation equation and a transition equation. The observation equation in state-space form is expressed as

$$y_{it} = \tau_k + \Lambda_k \eta_{it} + \epsilon_{it}. \tag{7}$$

where y_{it} is a vector of dimension p_k with observations at time t of subject i, η_{it} is a $w_k \times 1$ vector of latent variables at time t, measured by $p_k \times 1$ manifest variables at time t, Λ_k is a $p_k \times w_k$ factor loading matrix, τ_k is a $p_k \times 1$ vector of intercepts, and ϵ_{it} is a p_k-variate vector of measurement errors for person i at time t. The subscript k (as opposed to s, as in the SEM model discussed earlier) is used to indicate elements in a state-space model.

The transition equation depicting the dynamics of the latent variables is expressed as

$$\eta_{it} = \alpha_k + B_k \eta_{i,t-1} + \zeta_{it}, \tag{8}$$

where α_k is $w_k \times 1$ vector of constants, B_k is a $w_k \times w_k$ transition matrix capturing the dynamics of the latent state variables, and ζ_{it} is a $w_k \times 1$ vector of residuals or dynamic noise, which is assumed to be uncorrelated with the measurement errors in Equation 7.

In standard state-space notations, all parameters (Λ_k, τ_k, B_k, α_k, and the covariance matrices of ζ_{it} and ϵ_{it}) are typically allowed to be time-varying. We only consider herein state-space models with time-invariant parameters. In some applications, additional exogenous input vectors are included in Equations 7 and 8 to capture the effects of other external variables on y_{it} and η_t, but this alternative representation form (see, e.g., Gelb, 1974) is not considered here. In sum, the linear state-space model is a very general way of representing the relations of a set of latent and manifest variables (e.g., in the forms of linear stochastic difference equations and a variety of time series models; Hamilton, 1994).

State Estimation Through the Kalman Filter

Closely associated with state-space models is a set of estimation algorithms collectively known as the KF—an online estimation procedure used to predict current or future states (e.g., factor scores) given information up to the current time point by minimizing prediction errors in the least squares sense (Zarchan & Musoff, 2000). The KF has been utilized in a number of areas in social sciences. For instance, Heath (2000) used the KF to detect sudden changes in the parameters of autoregressive moving average (ARMA) models. In the econometric and

psychometric literature, the KF is often used as a building block for estimating parameters within the state-space framework (e.g., Kim & Nelson, 1999; Molenaar & Raijmakers, 1998). Equally notable is the KF's practicality as a tool for estimating factor scores in longitudinal models (Oud et al., 1990).

The KF is first initiated by setting the state vector and its associated covariance matrix at time $t = 0$ (i.e., $\eta_{i,0|0}$ and $P_{0|0}$, respectively) for all persons to some user-specified initial conditions. The notation $\eta_{i,t|t}$ is used to denote the conditional expectation of person i's latent state vector at time t given manifest observations from time t (i.e., $\eta_{it|t} = E(\eta_{it}|Y_{it})$), and $P_{t|t}$ represents its associated covariance matrix. If no prior information is available at time $t = 0$, we can fix $\eta_{i,0|0}$ to be a $w_k \times 1$ vector of zeros and $P_{0|0}$ to be a diffuse diagonal matrix as (Harvey, 1989; Zarchan & Musoff, 2000)

$$P_{0|0} = \kappa I, \tag{9}$$

where κ is a large positive scalar and I is a $w_k \times w_k$ identity matrix. We see that as κ approaches ∞, $P_{0|0}^{-1}$ approaches a null matrix (Harvey, 1989). Therefore, this specification is uninformative in the sense that it does not provide any precise information regarding the initial state of a system.

Further, the measurement errors, ϵ_{it}, and the residuals, ζ_{it} (from Equations 7 and 8), respectively, are assumed to be multivariate normally distributed with covariance matrices Θ_k and Ψ_k as

$$\epsilon_{it} \sim N(0, \Theta_k), \text{and}$$

$$\zeta_{it} \sim N(0, \Psi_k).$$

Initial guesses on the parameters are then specified and used to obtain person i's state estimates across all measurement occasions. Specifically, the goal here is to obtain estimates for $\eta_{i,t|t}$ and $P_{t|t}$ for all time points but this is done in two phases. The first phase is to predict person i's state at time t ($\eta_{i,t|t-1}$) and the associated covariance matrix ($P_{t|t-1}$) based on the manifest observations from time $t-1$ as

$$\eta_{i,t|t-1} = \alpha_k + B_k \eta_{i,t-1|t-1}, \text{and} \tag{10}$$

$$P_{t|t-1} = B_k P_{t-1|t-1} B_k' + \Psi_k. \tag{11}$$

Notice that the preceding state estimates are based only on observations up to time $t-1$. Using the observations at time t, these estimates are then updated to yield $\eta_{i,t|t}$ and $P_{t|t}$, defined as

$$\eta_{i,t|t} = \eta_{i,t|t-1} + K_{t|t}[y_{it} - (\Lambda_k \eta_{i,t|t-1} + \tau_k)], \tag{12}$$

$$P_{t|t} = [I - K_{t|t}\Lambda_k] P_{t|t-1}, \tag{13}$$

where $K_{t|t} = P_{t|t-1}\Lambda_k'[\Lambda_k P_{t|t-1}\Lambda_k{}' + \Theta_k]^{-1}$, usually known as the Kalman gain function, determines how heavily the discrepancy between predicted and actual measurements is weighted in updating the current estimate of η_{it}. A more detailed explanation pertaining to how the gain function is derived can be found in Zarchan and Musoff (2000). The difference $[y_{it} - (\Lambda_k \eta_{i,t|t-1} + \tau_k)]$ shown in Equation 12 is often termed *innovation*, as it represents the difference

between the predicted measurement, $y_{i,t|t-1} = \Lambda_k \eta_{i,t|t-1} + \tau_k$, and the actual measurement y_{it}, or in other words, the new information brought in by the observations at time t. This term is also known as the one-step-ahead prediction error (e.g., Durbin & Koopman, 2001). In short, model-implied state and covariance estimates are derived in the prediction phase. Information is then incorporated in the updating phase, which, in general, improves the precision estimate for the state from the prediction phase. Thus, the covariace matrix at the updating phase, $P_{t|t}$, is smaller than its counterpart at the predicting phase, $P_{t|t-1}$, as evidenced in Equation 13.

Parameter Estimation by Prediction Error Decomposition

As the KF is cycling through the estimation across N persons and T time points, several by-products from the KF are available at each time point t from each person i. These by-products can be substituted into a log-likelihood function and through optimizing this function with respect to the model parameters we can obtain maximum likelihood estimates of these parameters. As mentioned earlier, the innovation for person i at time t is

$$\begin{aligned} e_{i,t|t-1} &= y_{it} - y_{i,t|t-1} \\ &= y_{it} - (\Lambda_k \eta_{i,t|t-1} + \tau_k), \end{aligned} \tag{14}$$

because it has zero mean, $E[e_{i,t|t-1}] = 0$. The covariance matrix of $e_{i,t|t-1}$, denoted later as G_t, is equal to

$$\begin{aligned} E[(e_{i,t|t-1})(e_{i,t|t-1})'] &= G_t \\ &= \Lambda_k P_{t|t-1} \Lambda_k' + \Theta_k \text{ for all } i. \end{aligned} \tag{15}$$

Now the log-likelihood function can be written as a function of the individual innovation vector, $e_{i,t|t-1}$, and its associated covariance matrix, G_t, yielding

$$LL_{KF}(\theta_k) = \frac{1}{2} \sum_{i=1}^{N} \sum_{t=1}^{T} [-p_k \log(2\pi) - log|G_t| - (e'_{i,t|t-1}) G_t^{-1} (e_{i,t|t-1})], \tag{16}$$

where p_k is the number of manifest variables, which can be person-dependent in the presence of missing data. Equation 16 is known as the prediction error decomposition (PED) function and maximizing this function with respect to the parameters in Λ_k, τ_k, B_k, α_k, Ψ_k, and Θ_k results in maximum likelihood estimates of these parameters (Goodwin & Payne, 1977; Harvey, 1989; Ljung & Söderström, 1983; Shumway, 2000). In addition, when the parameters are constrained to be invariant across persons, the resultant model captures the pooled dynamics in the sample as a whole. This is also the case with most SEM longitudinal models. Alternatively, if sufficient data points are available from each person, the parameters can also be allowed to vary across persons to examine intraindividual dynamics more thoroughly.

In sum, by first setting the parameters to some fixed initial values, each person's data are subjected to the KF algorithm and individual state estimates are thus obtained (from $i = 1, \ldots N$) by treating the parameters as fixed values. State estimates from N individuals are then treated as known values and are used to estimate the parameters using the PED function. The new parameter estimates are then treated as fixed values again to derive new state estimates. The procedures are repeated until convergence criteria are met.

ANALYTIC SIMILARITIES AND DIFFERENCES BETWEEN SEM AND STATE-SPACE MODELING

In most areas in psychology in which longitudinal studies prevail, panel designs involving relatively large numbers of participants and variables but relatively few measurement occasions remain the dominant sampling design. In fact, such discipline-specific preferences are so deeply entrenched in contemporary sociobehavioral research that they have helped shape the ways researchers conceive state-space models. MacCallum and Ashby (1986), for instance, noted that state-space modeling is a special case of SEM because multiple measurement occasions can simply be incorporated into a structural equation model by increasing the dimension of y_i and η_i. Their conclusion diverged from the perspectives of other researchers, who showed that the SEM model can be obtained as a constrained special case of the state-space model (e.g., Otter, 1986).

Indeed, SEM and the state-space model can be shown to be special cases of one another depending on the kinds of constraints a researcher imposes on each. In our opinion, the key difference between SEM and state-space model does not lie in whether one is a special case of another, but rather in the different modeling paradigms they exemplify. Whereas both approaches can be used to represent *pooled intraindividual* dynamics, the primary strengths of SEM are in capturing simultaneous structural relations among latent variables and possible *interindividual* differences in such relationships. State-space models, in contrast, are well suited as a tool for representing more complex *intraindividual* dynamics, particularly when $T > N$. Here, we provide more details regarding the similarities and differences between the two approaches with respect to state estimation and parameter estimation.

State Estimation

The state estimates yielded by the KF are estimates of the latent variables in a state-space model. This includes factor scores as a special case. In psychology, factor score estimation is one of the most commonly encountered applications of state estimation. In fact, the equivalence between the KF and other cross–sectional factor score estimators has already been discussed by several researchers. These factor score estimators are either available directly in most SEM programs, or can be readily implemented using elements from an SEM model. Here, we provide only a brief summary of these previous analytic results.

In one of the key papers on this topic, Oud, van den Bereken, and Essers (1990) showed that when $B_k = \mathbf{0}$, the KF is analytically identical to the regression method of estimating factor scores (see Lawley & Maxwell, 1971). Furthermore, the KF yields identical factor scores as another commonly used factor score estimator, the Bartlett estimator, when $B_k = \mathbf{0}$ and $P_{t|t-1}^{-1}$ approaches 0. The second scenario occurs when Ψ approaches ∞, or when a diffuse prior is used (as defined in Equation 9). In other words, when $B_k = \mathbf{0}$, if no prior information is available to help define the state or factor score covariance matrix a priori, results yielded by the KF are identical to results obtained from using the Barlett estimator.

Dolan and Molenaar (1991) considered the alternative scenario of a quasi-Markov model wherein B_k is not a null matrix. They showed that when $T > 1$, the regression method yields more accurate factor score estimates because to estimate factor scores at time t, it incorporates information from all time points (i.e., from time $t = 1, \ldots T$), rather than information from *up*

to time t as in the KF. The regression method in addition was shown to yield identical results as a KF-related smoothing algorithm known as the fixed interval smoother (for details see Dolan & Molenaar, 1991). Some of these relationships are verified in our simulations reported later.

Parameter Estimation

Commonalities between the SEM and the state-space approach in regard to parameter estimation have also been addressed previously. Otter (1986) was among the first researchers who pointed out the equivalence between SEM and the state-space model in this respect when $T = 1$. However, he did not discuss the direct linkage between the raw data maximum likelihood function in SEM and the PED function used in estimating state-space models. In the Appendix, we show the analytic equivalence between the two functions when B_s and B_k are both null matrices. This arises naturally in cross-sectional models in which $T = 1$. Even when $T > 1$ and the number of manifest variables in SEM, p_s, is equal to $T^* p_k$ (p_k = the number of manifest variables in an equivalent state-space model), the log-likelihood functions used to obtain ML estimates in the two modeling contexts are still identical, as are the resultant parameter estimates, provided that B_s and B_k are both null matrices, and no further lagged relationships across measurement occasions are included in, for example, Λ_s, Ψ_s, or Θ_s.

When B_s and $B_k \neq \mathbf{0}$, the structural model in SEM and the transition equation in a state-space framework are generally not equivalent unless concurrent relationships among latent variables are included in the latter through specialized parameter constraints. Even with the appropriate constraints, the impact of the diffuse initial covariance matrix (i.e., $P_{0|0}$) used to initialize the KF is inherently embedded in the PED function through Equations 10 through 15, except when $B_k = 0$. Thus, unless the same initial state estimates and covariance matrix are also incorporated as part of an SEM model, the parameter estimates yielded from the two approaches would not be identical in general. Such constraints are best illustrated using concrete examples and we do so within the context of three modeling examples. To facilitate further comparisons between the SEM and the state-space frameworks, the modeling equations and likelihood functions assumed in the two approaches are gathered in Table 1.

ILLUSTRATIVE EXAMPLES AND RESULTS

We provide three simulation examples to illustrate the similarities and differences between state-space modeling and SEM in cross-sectional and longitudinal settings. The three simulation examples considered include (a) a cross-sectional common factor model, (b) a univariate latent difference score model, and (c) a bivariate cross-lag dynamic factor analysis model. Starting with the $T = 1$ model in Example 1, we hope to clarify the common misconception that state-space techniques can only be used to handle repeated measurement data from a single subject. The second example is a difference equation model involving B_s and B_k that are not null matrices. We show the constraints that have to be imposed to yield identical parameter estimates using the two approaches, including estimates for fixed and random effects. The last example is used to highlight the strengths of the state-space framework in handling data with $T > N$. All structural equation models were fitted using LISREL and the state-space models were either fitted using mkfm6, a FORTRAN program developed by Dolan (2002), or *SsfPack*,

TABLE 1
Modeling Equations and Log-Likelihood Functions Used in the Structural Equation Modeling and the State-Space Frameworks

	Equations	*No. in Text*
Structural Equation Modeling		
Measurement	$y_i = \tau_s + \Lambda_s \eta_i + \epsilon_i$	Eq. 1
Structural	$\eta_i = \alpha_s + B_s \eta_i + \zeta_i$	Eq. 2
Log-likelihood	$LL_{RML}(\theta_s) =$	
	$\frac{1}{2}\sum_{i=1}^{N}[-p_{si}\log(2\pi) - \log\lvert\Sigma\rvert - (y_i - \mu)'\Sigma^{-1}(y_i - \mu)]$	Eq. 5
	$\mu = \tau_s + \Lambda_s(I - B_s)^{-1}\alpha_s$	Eq. 4
	$\Sigma = \Lambda_s(I - B_s)^{-1}\Psi_s(I - B_s)^{-1'}\Lambda_s' + \Theta_s$	Eq. 3
State-space		
Observation	$y_{it} = \tau_k + \Lambda_k \eta_{it} + \epsilon_{it}$	Eq. 7
Transition	$\eta_{it} = \alpha_k + B_k \eta_{i,t-1} + \zeta_{it}$	Eq. 8
Log-likelihood	$LL_{KF}(\theta_k) =$	
	$\frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}[-p_k\log(2\pi) - log\lvert G_t\rvert - (e_{i,t\lvert t-1}')G_t^{-1}(e_{i,t\lvert t-1})]$	Eq. 16
	$e_{i,t\lvert t-1} = y_{it} - (\Lambda_k \eta_{i,t\lvert t-1} + \tau_k)$	Eq. 14
	$G_t = \Lambda_k P_{t\lvert t-1}\Lambda_k' + \Theta_k$	Eq. 15

a suite of C routines for implementing state-space modeling techniques (Koopman, Shephard, & Doornik, 1999).[3]

Example 1: Common Factor Model (Large N, $T = 1$)

In this example, we consider a common factor model with $N = 200$, $T = 1$, four manifest variables ($p_s = p_k = 4$), and one common factor ($w_s = w_k = 1$). In SEM, the measurement model is written as

$$\begin{bmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \\ y_{4i} \end{bmatrix} = \begin{bmatrix} \tau_{y1} \\ \tau_{y2} \\ \tau_{y3} \\ \tau_{y4} \end{bmatrix} + \begin{bmatrix} 1 \\ \lambda_{21} \\ \lambda_{31} \\ \lambda_{41} \end{bmatrix} \begin{bmatrix} f_{1i} \end{bmatrix} + \begin{bmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \\ u_{4i} \end{bmatrix} \tag{17}$$

$$y_i = \tau_s + \Lambda_s \eta_i + \epsilon_i,$$

where y_{ji} represents person i's manifest score on item j and λ_{j1} represents the loading of the jth item on the latent factor f_{1i} with the loading of the first item fixed at unity. η_i in this model only consists of one latent variable, f_{1i}, and the vector ϵ_i carries person i's uniquenesses associated with the four items (i.e., $u_{1i}, \ldots u_{4i}$) with a diagonal covariance matrix $E(\epsilon_i \epsilon_i') = \Theta$. The structural model in SEM is written as

$$\begin{bmatrix} f_{1i} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} f_{1i} \end{bmatrix} + \begin{bmatrix} z_{1i} \end{bmatrix}, \tag{18}$$

$$\eta_i = B_s \eta_i + \zeta_i,$$

[3] *SsfPack* is one of the many statistical packages implemented using Ox, an object-oriented matrix programming language with a comprehensive mathematical and statistical function library (Doornik, 1998). Mkfm6 can be downloaded from Conor Dolan's Web site at http://users.fmg.uva.nl/cdolan/, whereas Ox and *SSfPack* can be downloaded from http://www.doornik.com/download.html.

TABLE 2
True Parameters Used to Simulate a One-Factor Model and Parameters Recovered by Using a Structural Equation Modeling versus a State-Space Approach With $N = 200$, $T = 1$, and $p_s = p_k = 4$

Matrices	*True Parameters*	*Structural Equation Modeling Estimates (SE)*	*State-Space Estimates (SE)*
$\Lambda_s = \Lambda_k$	$\begin{bmatrix} 1 \\ 1.2 \\ .7 \\ .8 \end{bmatrix}$	$\begin{bmatrix} =1 \\ 1.2823(.0864) \\ .8227(.0700) \\ .8417(.0877) \end{bmatrix}$	$\begin{bmatrix} =1 \\ 1.2823(.0854) \\ .8227(.0697) \\ .8417(.0876) \end{bmatrix}$
$E(\epsilon_i \epsilon_i') = \Theta_s = \Theta_k$	$\begin{bmatrix} 5 \\ 4 \\ 6 \\ 10 \end{bmatrix}$	$\begin{bmatrix} 5.3627(.7151) \\ 2.1878(.7762) \\ 6.0199(.6934) \\ 11.4892(1.2303) \end{bmatrix}$	$\begin{bmatrix} 5.3627(.7037) \\ 2.1879(.7698) \\ 6.0199(.6978) \\ 11.4891(1.2264) \end{bmatrix}$
$\tau_s = \tau_k$	$\begin{bmatrix} 3 \\ 4 \\ 5 \\ 7 \end{bmatrix}$	$\begin{bmatrix} 2.7074(.2868) \\ 3.6790(.3193) \\ 4.9759(.2602) \\ 6.6474(.3113) \end{bmatrix}$	$\begin{bmatrix} 2.7074(.2861) \\ 3.6790(.3185) \\ 4.9759(.2595) \\ 6.6474(.3105) \end{bmatrix}$
$\Psi_s = \Psi_k$	$[10]$	$[11.0045(1.6204)]$	$[11.0045(1.6121)]$

where z_{1i} is the disturbance term associated with f_{1i} with variance–covariance matrix $E(\zeta_i^2) = \Psi$. In any cross-sectional model (i.e., $T = 1$), $B_s = B_k = \mathbf{0}$, $y_i = y_{it}$, $\eta_i = \eta_{it}$, $\epsilon_i = \epsilon_{it}$, $\zeta_i = \zeta_{it}$, $\tau_s = \tau_k$, $\Lambda_s = \Lambda_k$ and $\alpha_i = \alpha_{it}$. Consequently, the log-likelihood functions based on SEM and state-space modeling are also identical (see Table 1).

We simulated data using this model and obtained parameter estimates by optimizing $LL_{RML}(\theta_s)$ and $LL_{KF}(\theta_k)$ using LISREL and mkfm6, respectively. The true parameters used to simulate the data, together with the parameter estimates obtained from LISREL and mkfm6, are shown in Table 2. Results indicated that estimates from the two programs are identical, with minor rounding differences in some elements of the Θ matrix (the differences are less than or around .0001).[4] Because mkfm6, the Fortran program used for model fitting in this example, was written to use the observed information matrix to compute standard error estimates, whereas LISREL uses the expected information matrix to derive standard error estimates (Azzelini, 1996), very minor differences were observed between the standard error estimates given by the two programs, but these differences were to be expected. Factor score estimates obtained using the regression method and the KF were also identical (correlation between the two was 1.0). The factor score estimates obtained using the two approaches are shown in Table 3 for the first three and last three subjects together with their true factor scores. In sum, by using a cross-sectional common factor model, we demonstrated that the KF combined with the PED function yielded identical parameter and factor score estimates as LISREL and the regression estimator when $T = 1$ and $B_k = B_s = \mathbf{0}$.

[4]Note that because there is only one replication, some systematic biases were present in the parameter estimates in both cases.

TABLE 3
Factor Score Estimates Obtained by Using the Regression Method and the Kalman Filter for the First Three and Last Three Subjects

Subjects	*True Factor Scores*	*Regression Estimates*	*Kalman Filter Estimates*
1	−4.753	−4.437	−4.437
2	−1.188	−1.413	−1.413
3	−7.847	−5.582	−5.582
198	−0.107	−1.252	−1.252
199	2.559	2.060	2.060
200	−1.800	−0.577	−0.577

Example 2: A Univariate Latent Difference Score Model (Large *N*, Small *T*)

In this example, we consider a difference equation model, termed the dual change score model, proposed by McArdle and Hamagami (2001). This model is an example where B_k and B_s are not null matrices. Using this model, we show how the initial state vector and its covariance matrix can be used to specify fixed and random effects for an intercept and a slope component.

A latent difference component that represents the change from time t to $t + 1$ is expressed as

$$\eta_{i,t+1} = \eta_{it} + \Delta\eta_{i,t+1}, \tag{19}$$

where η_{it} is a latent variable of interest characterizing person i at t and $\Delta\eta_{i,t+1}$ is a latent difference component capturing person i's change from time t to $t + 1$.

McArdle and Hamagami (2001) proposed a dual change score representation to further express the latent difference score as

$$\Delta\eta_{it} = \eta_{si} + \beta\eta_{i,t-1}, \tag{20}$$

where η_{si} is a constant term that determines person i's average magnitude of change in η per unit of time, the parameter β is a person-invariant "self-feedback" parameter that governs the magnitude of change proportionate to the previous level of η at $t - 1$.

The latent variable, η_{it}, is in turn identified using a single indicator, y_{it}, yielding

$$y_{it} = \eta_{it} + e_{it}, \tag{21}$$

where e_{it} represents the measurement error associated with person i's observed score at time t. Incorporating $\Delta\eta_{t+1}$ from Equation 20 into Equation 19 yields

$$\eta_{i,t+1} = \eta_{it} + \eta_{si} + \beta\eta_{it} \tag{22}$$

$$= \eta_{si} + (1 + \beta)\eta_{it}, \tag{23}$$

Following the modeling conventions used in growth curve models, η_{i0}, the intercept term of person i at time 0, can be reexpressed as a function of a group average, $\mu_{\eta 0}$, and an individualized deviation from this group average, $\zeta_{\eta 0 i}$:

$$\eta_{0i} = \mu_{\eta 0} + \zeta_{\eta 0 i}, \tag{24}$$

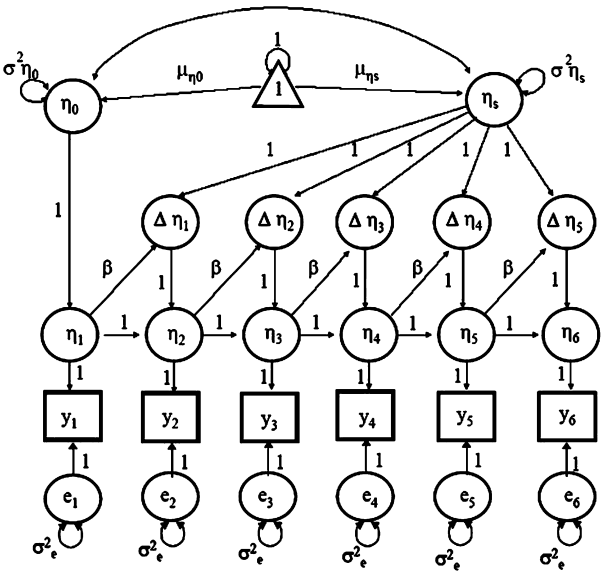


FIGURE 1 McArdle and Hamagami's (2001) dual change score model. Note that the triangle represents a unit vector used to specify means or intercepts in a model (see McArdle & Epstein, 1987).

Similarly, the constant change term for person i, η_{si}, can also be written as a function of a group average, $\mu_{\eta s}$, and an individualized deviation from this group average, $\zeta_{\eta si}$, written as

$$\eta_{si} = \mu_{\eta s} + \zeta_{\eta si}. \tag{25}$$

Essentially, $\mu_{\eta 0}$ and the associated interindividual differences (i.e., $E(\zeta^2_{\eta 0}) = \sigma^2_{\eta 0}$) are conceptually identical to the average intercept term and interindividual differences thereof in a growth curve modeling context (McArdle & Epstein, 1987; Meredith & Tisak, 1990; Rao, 1958; Tucker, 1958). Likewise, $\mu_{\eta s}$ and interindividual differences thereof (i.e., $E(\zeta^2_{\eta s}) = \sigma^2_{\eta s}$) are analogous to the latent slope in growth curve modeling and its associated variance, except that the slope of the trajectory is now further determined by additional relationships as expressed in Equation 20. A path diagram representation of the full model is shown in Figure 1. Collectively, this model allows for simple random effects within a difference equation model. In doing so, it helps reveal interindividual differences in intraindividual change through three sources of individual differences, namely (a) interindividual differences in intercept, (b) interindividual differences in average slope, and (c) measurement errors. Such interindividual differences aspects have interesting theoretical implications in the study of human development (e.g., McArdle et al., 2004).[5]

We reexpressed the dual change score model in a more compact form to better highlight the equivalence of this model in an SEM versus a state-space modeling framework. The resultant

[5]Bivariate extensions of this univariate model with coupling between two constructs have also been presented by McArdle and Hamagami (2001). Although the bivariate model is not considered here, results from this article are readily generalizable to other equivalent latent difference equation models.

structural model in SEM is written as

$$\begin{bmatrix} \eta_{i1} \\ \eta_{i2} \\ \eta_{i3} \\ \vdots \\ \eta_{i,T-1} \\ \eta_{iT} \\ \eta_{si} \end{bmatrix} = \begin{bmatrix} \mu_{\eta 0} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \mu_{\eta s} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1+\beta & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 1+\beta & 0 & \dots & 0 & 0 & 1 \\ \vdots & \dots & 1+\beta & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 & 1 \\ 0 & 0 & 0 & \dots & 1+\beta & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_{i1} \\ \eta_{i2} \\ \eta_{i3} \\ \vdots \\ \eta_{i,T-1} \\ \eta_{iT} \\ \eta_{si} \end{bmatrix} + \begin{bmatrix} \zeta_{0i} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \zeta_{si} \end{bmatrix} \tag{26}$$

$$\eta_i = \alpha_s + B_s \eta_i + \zeta_i \tag{27}$$

and the corresponding measurement model is written as

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{i,T-1} \\ y_{iT} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \dots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \eta_{i1} \\ \eta_{i2} \\ \eta_{i3} \\ \vdots \\ \eta_{i,T-1} \\ \eta_{iT} \\ \eta_{si} \end{bmatrix} + \begin{bmatrix} e_{i1} \\ e_{i2} \\ e_{i3} \\ \vdots \\ e_{i,T-1} \\ e_{iT} \end{bmatrix}, \tag{28}$$

$$y_i = \Lambda_s \eta_i + \epsilon_i \tag{29}$$

where Equation 26 is constructed based on Equation 23 directly without the intermediate latent difference score component, $\Delta\eta_{it}$. This specific formulation yields identical estimates as the model shown in the path diagram in Figure 1, although the estimate for β has to be obtained by subtracting 1 from the regression coefficient of $(1 + \beta)$ in B_s, or by specifying β as a constrained parameter.

An equivalent transition equation in the state-space approach can be formulated as

$$\begin{bmatrix} \eta_{it} \\ \eta_{sit} \end{bmatrix} = \begin{bmatrix} 1+\beta & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_{i,t-1} \\ \eta_{si,t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{30}$$

$$\eta_{it} = B_k \eta_{it} + \zeta_{it} \tag{31}$$

where η_{it} is person i's true score at time t and η_{sit} is person i's average slope, η_{si}, at time t, which has to be explicitly constrained to be time-invariant in the state-space formulation. The corresponding observation equation is written as

$$y_{it} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \eta_{it} \\ \eta_{sit} \end{bmatrix} + e_{it} \tag{32}$$

$$y_{it} = \Lambda_k \eta_{it} + \epsilon_{it}. \tag{33}$$

Note that the latent state at time 1, η_{i1}, does not show lagged influence from η_{i0} in the SEM model in Equation 26. To build an equivalent state-space model, the conditional means and

TABLE 4
True Parameters Used to Simulate the Dual Change Score Model and Parameters Recovered by LISREL and Ox with $N = 200$ and $T = 10$

Matrices	*True Parameters*	*Prediction Error Decomposition Estimates (SE)*	*LISREL Estimates (SE)*
μ_{η_0}	3.00	3.0087 (.1036)	3.0087 (.1039)
μ_{η_s}	1.00	1.0574 (.0777)	1.0574 (.0777)
σ_e^2	.50	.5098 (.0180)	.5098 (.0181)
β	−.30	−.2945 (.0102)	−.2945 (.0102)
$\sigma_{\eta_0}^2$	2.00	1.8395 (.2174)	1.8395 (.2180)
$\sigma_{\eta_s}^2$	1.00	.9550 (.1056)	.9550 (.1056)
σ_{η_0,η_s}	−.30	−.2316 (.1057)	−.2316 (.1059)

covariance matrix of the two state variables at time $t = 1$ have to be explicitly specified as

$$\eta_{i1|1} = \begin{bmatrix} \mu_{\eta_0} \\ \mu_{\eta_s} \end{bmatrix} \text{ for all } i \text{ and} \tag{34}$$

$$P_{i1|1} = \begin{bmatrix} \sigma_{\eta_0}^2 & \sigma_{\eta_0,\eta_s} \\ \sigma_{\eta_0,\eta_s} & \sigma_{\eta_s}^2 \end{bmatrix} \text{ for all } i. \tag{35}$$

Thus, as shown in Equations 34 and 35, fixed effects of the intercept and slope are contained in the initial state vector at time 1. Any covariate that can help account for additional between-person deviations in intercept and slope can also be incorporated into Equation 34. The random effects representing between-person deviations in intercept, slope, and their covariance ($\sigma_{\eta_0}^2$, $\sigma_{\eta_s}^2$, and σ_{η_0,η_s} respectively) now appear in the initial state covariance matrix at time 1. Subsequent state and measurement predictions can then be obtained based on Equations 30 and 32 starting from $t = 2$.

The parameters $\mu_{\eta 0}$, $\mu_{\eta s}$, $\sigma_{\eta 0}^2$, $\sigma_{\eta s}^2$, $\sigma_{\eta 0,\eta s}$, σ_e^2, and β were estimated by means of the PED. Results from fitting the SEM and the state-space model to simulated data with $T = 10$ and $N = 400$ using LISREL and Ox are summarized in Table 4. As expected, the parameter estimates yielded by the two approaches were identical. We thus showed that the initial state vector and its associated covariance matrix can be used to accommodate simple random effects such as the ones assumed in the dual change score model. These parameters can in turn be estimated by using the PED function.[6]

Both approaches can be used to fit the same difference equation model, but what exactly does the state-space approach add? In our view, the state-space approach complements what SEM lacks in at least two important ways. For one, SEM is not ideally suited for fitting longitudinal

[6]Another option is to fit a reformulated version of the dual change score model as explicated by McArdle and Hamagami (2001). Using this approach, the dual change score model can be expressed as an oblique two-factor model with the two factors representing latent level (i.e., η_0) and slope (i.e., η_s). The appropriate parameter constraints are then imposed on the factor loadings of η_s on y_t to obtain β. Thus, the resultant model eliminates all the latent change scores, $\Delta\eta_t$, and the true score components, η_t, from the path diagram in Figure 2. We did not consider this approach for this illustration.

models to data with $T > N$; the state-space approach, in contrast, makes it easy to handle such data. For another, even when $N > T$ but T is still relatively large (e.g., $N = 400$ and $T = 20$), the state-space approach still offers a convenient way to fit more complicated extensions of the dual change score model. Such extensions include having multiple indicators for η_{it}, accommodating for within-person deviations in dynamics in the form of process noise for η_{it}, and allowing for irregularly spaced data. In our opinion, such strengths are where the major appeal of the state-space approach resides, because model specification and the kinds of invariance constraints that a researcher has to impose in fitting more complex dynamic models within an SEM framework quickly become cumbersome as T increases. This is demonstrated in our next example in which one prototypical time series model is fitted to intensive repeated measures data.

Example 3: Autoregressive Dynamic Factor Model (Large T, $N = 1$)

P-technique model (Cattell, Cattell, & Rhymer, 1947) is one variation of the common factor model often used to identify intraperson dynamics manifested by a single individual on a set of manifest variables. As such, the P-technique model and its extensions have been used to represent patterns of intraindividual variability in different contexts (see, e.g., Jones & Nesselroade, 1990). Several researchers have proposed and applied extensions of the conventional P-technique model to incorporate lagged relationships among the latent factors and one's manifest variables of interest (e.g., Molenaar, 1985). The lagged relationships among factors and manifest variables are represented in the form of auto- and cross-regressions among factors (Chow, Nesselroade, Shifren, & McArdle, 2004; Nesselroade, McArdle, Aggen, & Meyers, 2002). The observation equation and transition equation are written respectively as

$$y_{it} = \Lambda_k \eta_{it} + \zeta_{it} \text{ and} \tag{36}$$

$$\eta_{it} = A_1 \eta_{i,t-1} + A_2 \eta_{i,t-2} + \ldots + A_r \eta_{i,t-s} + \zeta_{it}, \tag{37}$$

where y_{it} is a p_k-variate time series of manifest variables ($t = 1, 2, \ldots, T$) for person i, Λ_k is a $p_k \times w_k$ matrix of concurrent (i.e., lag 0) factor loadings, $\eta_{i,t-q}$ is a w_k-variate time series of factor scores q measurement occasions ago ($t = 1, 2, \ldots, T; q = 0, 1, 2, \ldots, r$) for person i, u_{it} is a p_k-variate time series of uniquenesses that can have autocorrelational but no cross-correlational structure for person i, and ζ_{it} is a disturbance term for person i representing parts of the factor scores that are common to all p_k manifest variables, but do not manifest auto- or cross-regression relationships over time. A_q is a weight matrix that contains the auto- and cross-regression weights among the latent factors such that the element a_{qjk} from the jth row and kth column represents the influence of the kth factor from q occasions earlier on the current value of the jth factor. We consider herein a dynamic model with two factors and two lags (i.e., with lag 0 and lag 1). In the state-space framework, the observation equation with three indicators for each factor is written as

$$\begin{bmatrix} y_{1it} \\ y_{2it} \\ y_{3it} \\ y_{4it} \\ y_{5it} \\ y_{6it} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & 1 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \end{bmatrix} \begin{bmatrix} f_{1it} \\ f_{2it} \end{bmatrix} + \begin{bmatrix} u_{1it} \\ u_{2it} \\ u_{3it} \\ u_{4it} \\ u_{5it} \\ u_{6it} \end{bmatrix}, \tag{38}$$

where λ_{jk} is the loading of the jth item on the kth factor with λ_{11} and λ_{32} fixed at unity for identification purposes. f_{kit} denotes factor k for person i at time t, and u_{jit} is the uniqueness associated with person i on item j, whose variance is denoted as $\sigma^2_{u_j}$. Θ_k is constrained to be a diagonal matrix with elements $\Theta_k = \text{diag}[\sigma^2_{u_1}\ \sigma^2_{u_2}\ \sigma^2_{u_3}\ \sigma^2_{u_4}\ \sigma^2_{u_5}\ \sigma^2_{u_6}]$. The corresponding transition equation is expressed as

$$\begin{bmatrix} f_{1it} \\ f_{2it} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} f_{1i,t-1} \\ f_{2i,t-1} \end{bmatrix} + \begin{bmatrix} z_{1it} \\ z_{2it} \end{bmatrix}, \tag{39}$$

where b_{jk} is the lag–1 regression weight associated with the influence of factor k from time $t-1$ on factor j at time t and z_{jit} is the process noise or within-person deviation in factor score j that cannot be accounted for by the vector autoregression process of order 1 (i.e., VAR(1) process). Its variance is denoted as $\sigma^2_{z_j}$ and the covariance between the two process noise terms is represented as $\sigma_{z_{12}}$.

The state-space model depicted in Equations 38 and 39 can be fitted directly to time series data by means of the KF and the PED. However, if other conventional SEM approaches are used to fit this model to data with $T > N$, the resultant data matrix would be of size $N \times Tp_k$. In this case, the covariance matrix would be singular and thus not amenable for maximum likelihood estimation using Equation 6. There are at least two ways of handling data such as these in SEM packages. First,the raw data likelihood function in Equation 5 can be used to circumvent this problem, with the appropriate constraints imposed on the structure of Σ to ensure its positive definiteness. This was the approach presented by Hamaker et al. (2003). Still, even when used in conjunction with the flexible programming features offered by, for example, Mx, the model specifications involved are relatively complex, and this approach remains computationally inefficient given a large number of measurement occasions because this method requires the inversion of a $Tp_k \times Tp_k$ covariance matrix at each iteration.

Alternatively, we can use the block–Toeplitz approach, which is a commonly adopted SEM-based approach for fitting dynamic models to intensive repeated measures data (Boker & Bisconti, 2006; Chow et al., 2004; Molenaar, 1985; Nesselroade et al., 2002; Wood & Brown, 1994). This approach requires the computation of a block Toeplitz matrix that serves as the observed covariance matrix in SEM software. A block Toeplitz matrix consists of the lagged covariance matrices combined such that in the main block diagonal we have the lag zero covariance matrix, and in the off-diagonal block we have the lag one covariance matrix, and so on. The number of lags that is included in the block Toeplitz matrix is to some extent an arbitrary choice, but it also depends on the model one wants to fit (Molenaar, 1985; Wood & Brown, 1994). A pictorial representation of a block Toeplitz matrix with lag zero and lag one covariance matrices is provided in Figure 2. Clearly, there are many redundant elements in this matrix, and failing to take this into account results in an overestimation of the actual degrees of freedom. To avoid this we freed all the shaded elements in the block Toeplitz approach to exhaust the false degrees of freedom. Another issue to keep in mind is that the statistical independence assumption in SEM is violated in this case because of the temporal dependence in the time series data. As a result, this approach only yields pseudo-maximum likelihood estimates (cf. Molenaar & Nesselroade, 1998). It is this latter SEM approach that we will compare against PED-based estimates in the following simulations.

To highlight the utility of the state-space approach in comparison to SEM for handling data where $T >$ N, we ran two sets of Monte Carlo simulations, one with $N = 1$ and $T = 200$,

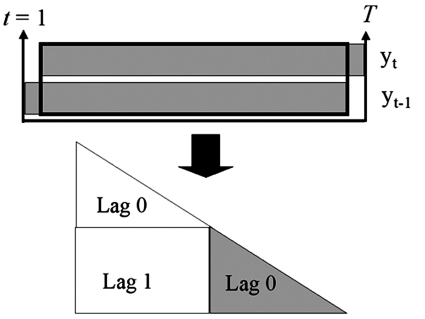


FIGURE 2 The process of creating a block Toeplitz matrix with 0 and 1 lags using univariate data. The lag 0 matrix shown here is just the usual concurrent covariance matrix. To create a matrix with lag 1 variances and covariances, the data are lagged by one time point to create two arrays of data representing y_t and y_{t-1}. The asymmetric covariance matrix computed using these two data blocks gives the lag 1 covariance matrix, which is then stacked together with two lag 0 covariance matrices to create a symmetric super matrix.

and another one with $N = 10$ and $T = 200$. We chose a moderately large T to compare the two approaches under relatively optimal model-fitting scenarios. A total of 500 Monte Carlo replications was run under each condition. Furthermore, we ran an additional 5,000 Monte Carlo replications to obtain two indexes of empirical standard deviations. Specifically, the first index is simply the standard deviation of the PED estimates of each parameter across 5,000 Monte Carlo runs and it is computed as

$$SD_{MC_1} = \sqrt{\sum_{k=1}^{K}(\hat{\theta}_{ak} - \bar{\hat{\theta}}_a)^2/K}, \tag{40}$$

where K is the total number of Monte Carlo replications ($K = 5{,}000$ in this simulation), $\hat{\theta}_{ak}$ is the estimate for parameter a in the kth Monte Carlo replication, and $\bar{\hat{\theta}}_a$ is the average of parameter a over K Monte Carlo runs. The block Toeplitz approach and the state-space approach would each yield a different Monte Carlo standard deviation (i.e., SD_{MC_1}); we only used the one based on the PED.[7] To have a more objective basis for comparing the standard error estimates yielded by the two approaches, we also computed a second index of empirical standard deviation, expressed as

$$SD_{MC_2} = \sqrt{RMSE_a^2 - (\bar{\hat{\theta}}_a - \theta_a)^2}, \tag{41}$$

where $\bar{\hat{\theta}}_a$ is, again, the average of parameter a over K Monte Carlo runs, θ_a is the true value of parameter a, and $RMSE_a$ is the root mean squared error (RMSE) associated with the estimates for parameter a, defined as

$$RMSE_a = \sqrt{\sum_{k=1}^{K}(\hat{\theta}_{ak} - \theta_a)^2/K}. \tag{42}$$

[7]This is because parameter estimates yielded from the PED are known to be maximum likelihood estimates (Shumway, 2000) but estimates obtained using the block Toeplitz method in general are not (Hamaker, Dolan, & Molenaar, 2002).

Specifically, this second index of empirical standard deviation is computed based on the discrepancies between the Monte Carlo estimates and the true parameter values. These empirical standard deviations are taken as the "true" standard errors of the model parameters against which the standard error estimates obtained through the PED and the block Toeplitz approach were compared. PED-based parameter estimates were obtained using mkfm6, whereas the block Toeplitz-based parameter estimates were obtained from LISREL.

Simulation results based on fitting the autoregressive dynamic factor analysis model to (a) a block Toeplitz matrix computed from data with $N = 1$ and $T = 200$, and (b) a block Toeplitz matrix computed from data with $N = 10$ and $T = 200$ are summarized in Tables 5 and 6, respectively. All parameter estimates yielded by the two approaches were different, as expected. When $N = 1$, the parameter estimates obtained from the PED were in general unbiased (see box plots of selected parameter estimates in Figure 3a). These estimates are, by definition, maximum likelihood estimates (see Shumway & Stoffer, 2004) and were generally more accurate than those yielded by using the block Toeplitz approach. Using the latter approach, estimates for the factor loading matrix, Λ_s, and the measurement error covariance matrix, Θ_s, were unbiased (see Figure 3a). Biases were observed among parameters that appeared in the process noise covariance matrix, Ψ_s and the transition matrix, B_s (see also the median estimates for all parameters in Table 5). In particular, the parameter estimates $\hat{\sigma}_{z12}$, $\hat{\sigma}^2_{z_2}$, $\hat{b}_{12}$, and $\hat{b}_{21}$ based on the block Toeplitz approach were characterized by substantially larger RMSEs (see Equation 42;

TABLE 5
Parameter Estimates From the Autoregressive Dynamic Factor Analysis Model Across 500 Monte Carlo Runs with $T = 200$ and $N = 1$

Matrices	*True Values* $(SD_{MC_1}/SD_{MC_{2P}}/SD_{MC_{2S}})$[a]	*Median PED Estimates (SE)*	*Median SEM Estimates (SE)*
$\begin{bmatrix}\lambda_{21}\\ \lambda_{31}\\ \lambda_{42}\\ \lambda_{52}\end{bmatrix}$	$\begin{bmatrix}1.2(.083/.083/.092)\\ .8(.072/.070/.072)\\ .9(.055/.058/.058)\\ 1.1(.048/.050/.053)\end{bmatrix}$	$\begin{bmatrix}1.197(.087)\\ .797(.077)\\ .900(.058)\\ 1.100(.050)\end{bmatrix}$	$\begin{bmatrix}1.199(.080)\\ .799(.064)\\ .900(.045)\\ 1.101(.042)\end{bmatrix}$
$\begin{bmatrix}\sigma^2_{u_1}\\ \sigma^2_{u_2}\\ \sigma^2_{u_3}\\ \sigma^2_{u_4}\\ \sigma^2_{u_5}\\ \sigma^2_{u_6}\end{bmatrix}$	$\begin{bmatrix}1.0(.175/.160/.177)\\ 1.2(.237/.223/.250)\\ 2.0(.226/.230/.238)\\ .8(.119/.121/.131)\\ 1.5(.171/.166/.177)\\ .4(.110/.106/.120)\end{bmatrix}$	$\begin{bmatrix}1.001(.168)\\ 1.166(.226)\\ 1.975(.225)\\ .794(.116)\\ 1.483(169)\\ .399(.107)\end{bmatrix}$	$\begin{bmatrix}1.007(.180)\\ 1.184(.200)\\ 1.983(.200)\\ .792(.118)\\ 1.502(.155)\\ .397(.119)\end{bmatrix}$
$\begin{bmatrix}\sigma^2_{z_1}\\ \sigma_{z_{12}}\\ \sigma^2_{z_2}\end{bmatrix}$	$\begin{bmatrix}2.0(.296/.287/.289)\\ .6(.178/.179/.206)\\ 2.0(.255/.270/.315)\end{bmatrix}$	$\begin{bmatrix}1.966(.296)\\ .597(.175)\\ 1.973(.259)\end{bmatrix}$	$\begin{bmatrix}1.917(.261)\\ .713(.319)\\ 2.335(.243)\end{bmatrix}$
$\begin{bmatrix}b_{11}\\ b_{12}\\ b_{21}\\ b_{22}\end{bmatrix}$	$\begin{bmatrix}.5(.073/.074/.077)\\ -.1(.060/.063/.068)\\ -.3(.073/.073/.084)\\ .6(.059/.060/.070)\end{bmatrix}$	$\begin{bmatrix}.482(.071)\\ -.097(.059)\\ -.304(.070)\\ .588(.058)\end{bmatrix}$	$\begin{bmatrix}.545(.102)\\ -.291(.088)\\ -.193(.126)\\ .638(.077)\end{bmatrix}$

Note. PED = prediction error decomposition; SEM = structural equation modeling.

[a] SD_{MC_1} = empirical standard deviation computed using Equation 40 and the PED estimates; $SD_{MC_{2P}}$ = empirical standard deviation computed using Equation 41 with the averages and root mean square errors (RMSEs) of the PED estimates; $SD_{MC_{2S}}$ = empirical standard deviation computed using Equation 41 with the averages and RMSEs of the block Toeplitz estimates.

TABLE 6
Parameter Estimates From the Autoregressive Factor Analysis Model Across 500 Monte Carlo Runs with $T = 200$ and $N = 10$

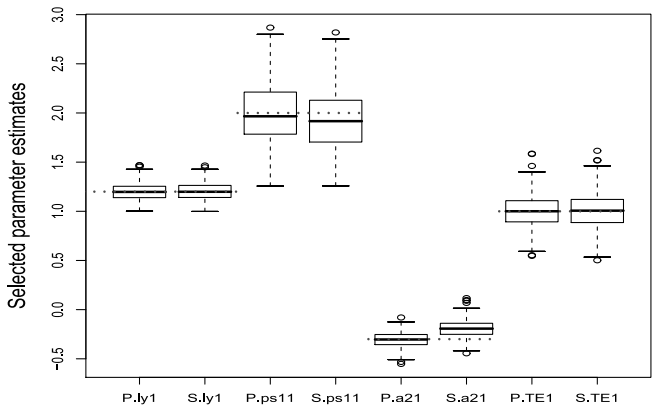
Matrices	*True Values* ($SD_{MC_1}/SD_{MC_{2P}}/SD_{MC_{2S}}$)[a]	*Median PED Estimates (SE)*	*Median SEM Estimates (SE)*
$\begin{bmatrix}\lambda_{21}\\ \lambda_{31}\\ \lambda_{42}\\ \lambda_{52}\end{bmatrix}$	$\begin{bmatrix}1.2(.025/.026/.028)\\ .8(.023/.024/.024)\\ .9(.017/.017/.017)\\ 1.1(.015/.015/.016)\end{bmatrix}$	$\begin{bmatrix}1.1984(.028)\\ .7996(.024)\\ .9008(.018)\\ 1.1003(.016)\end{bmatrix}$	$\begin{bmatrix}1.1977(.026)\\ .8002(.020)\\ .9003(.014)\\ 1.1011(.014)\end{bmatrix}$
$\begin{bmatrix}\sigma^2_{u_1}\\ \sigma^2_{u_2}\\ \sigma^2_{u_3}\\ \sigma^2_{u_4}\\ \sigma^2_{u_5}\\ \sigma^2_{u_6}\end{bmatrix}$	$\begin{bmatrix}1.0(.053/.054/.058)\\ 1.2(.071/.071/.079)\\ 2.0(.071/.071/.072)\\ .8(.037/.035/.038)\\ 1.5(.054/.054/.055)\\ .4(.034/.032/.036)\end{bmatrix}$	$\begin{bmatrix}.997(.053)\\ 1.195(.072)\\ 1.991(.071)\\ .799(.037)\\ 1.498(.054)\\ .400(.034)\end{bmatrix}$	$\begin{bmatrix}.992(.057)\\ 1.193(.079))\\ 1.991(.063)\\ .807(.038)\\ 1.506(.050)\\ .396(.038)\end{bmatrix}$
$\begin{bmatrix}\sigma^2_{z_1}\\ \sigma_{z_{12}}\\ \sigma^2_{z_2}\end{bmatrix}$	$\begin{bmatrix}2.0(.092/.093/.097)\\ .6(.057/.059/.072)\\ 2.0(.083/.083/.088)\end{bmatrix}$	$\begin{bmatrix}1.999(.095)\\ .595(.056)\\ 2.000(.083)\end{bmatrix}$	$\begin{bmatrix}2.086(.083)\\ .630(.097)\\ 2.191(.080)\end{bmatrix}$
$\begin{bmatrix}b_{11}\\ b_{12}\\ b_{21}\\ b_{22}\end{bmatrix}$	$\begin{bmatrix}.5(.022/.022/.025)\\ -.1(.018/.018/.021)\\ -.3(.022/.022/.022)\\ 6(.019/.020/.019)\end{bmatrix}$	$\begin{bmatrix}.498(.022)\\ -.100(.019)\\ -.301(.022)\\ .598(.018)\end{bmatrix}$	$\begin{bmatrix}.519(.026)\\ -.138(.026)\\ -.367(.038)\\ .656(.028)\end{bmatrix}$

Note. PED = prediction error decomposition; SEM = structural equation modeling.
[a] SD_{MC_1} = empirical standard deviation computed using Equation 40 and the PED estimates; $SD_{MC_{2P}}$ = empirical standard deviation computed using Equation 41 with the averages and root mean square errors (RMSEs) of the PED estimates; $SD_{MC_{2S}}$ = empirical standard deviation computed using Equation 41 with the averages and RMSEs of the block Toeplitz estimates.

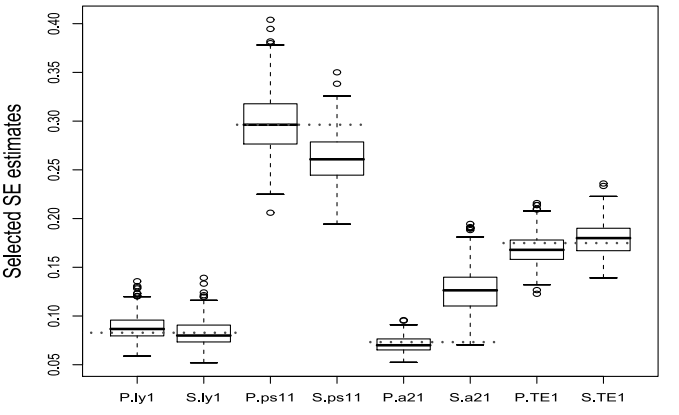
RMSEs = .24, .48, .20, and .14, respectively) than the corresponding PED estimates (RMSEs = .18, .27, .06, and .07, respectively). Particularly worth noting is that as a consequence of these higher biases, results from running Wald's tests on these parameters would have suggested that $\hat{b}_{21}$ was not significantly different from zero based on its standard error estimate (see Table 5). This would have led to the wrong conclusion that changes in f_2 were not related to f_1 but in fact, the cross-regressive relationships between f_1 and f_2 were reciprocal in nature. The PED estimates and their corresponding standard errors, in contrast, led to the correct statistical conclusion. RMSEs for other parameter estimates were comparable across the two approaches.

Similar results were observed when $T = 200$ and $N = 10$. Even though the parameter estimates yielded by both approaches were more accurate and showed less variability across Monte Carlo runs compared to the estimates from the previous condition (i.e., $T = 200$ and $N = 1$), block Toeplitz estimates of the structural parameters in Ψ_s and B_s still showed higher biases compared with the PED estimates (see Figure 4a and Table 6). RMSEs associated with the PED estimates for $\hat{\sigma}^2_{z_1}$, $\hat{\sigma}_{z_{12}}$, $\hat{\sigma}^2_{z_2}$, $\hat{b}_{11}$, $\hat{b}_{12}$, $\hat{b}_{21}$, and $\hat{b}_{22}$ were .09, .06, .08, .02, .02, .02, and .02. RMSEs for the corresponding block Toeplitz estimates were .13, .08, .21, .03, .04, .07, and .06. RMSEs for other parameter estimates were comparable across the two approaches.

The two indexes of empirical standard deviations computed using Equations 40 and 41 yielded values that were generally close to one another (see Tables 5 and 6). In some cases, the



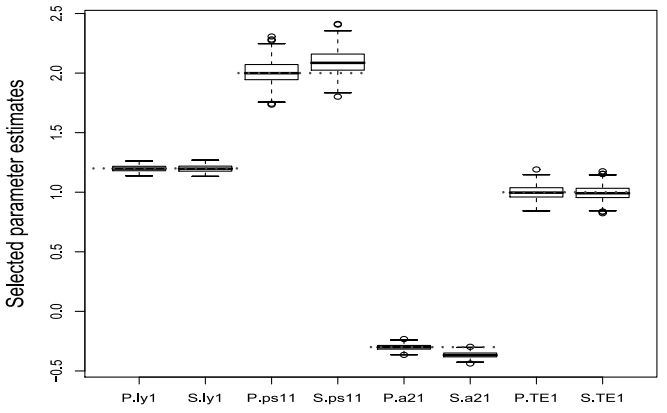
(a) $T = 200$, $N = 1$



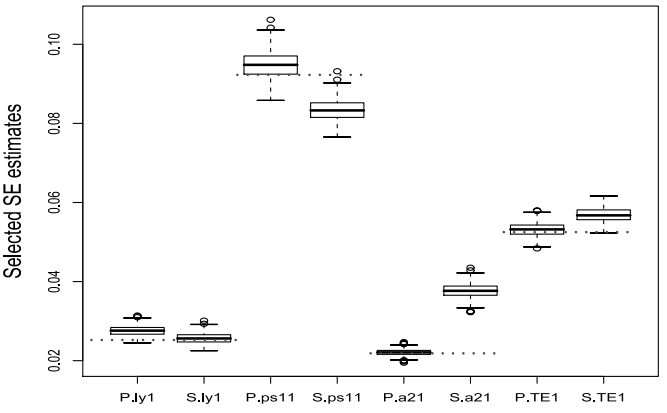
(b) $T = 200$, $N = 1$

FIGURE 3 (a) Selected parameter estimates obtained using the prediction error decomposition (PED) and structural equation modeling (SEM) and (b) their corresponding standard error estimates. Here, $N = 1$ and $T = 200$; P.ly1 $= \hat{\lambda}_{21}$ by PED; S.ly1 $= \hat{\lambda}_{21}$ by SEM; P.ps11 $= \hat{\psi}_{11}$ by PED; S.ps11 $= \hat{\psi}_{11}$ by SEM; P.a21 $= \hat{a}_{21}$ by PED; S.a21 $= \hat{a}_{21}$ by SEM; P.TE1 $= \hat{\theta}_{11}$ by PED; S.TE1 $= \hat{\theta}_{11}$ by SEM. In (a), the true parameter values are marked with thick dotted lines, and in (b) the true empirical standard deviations (i.e., SD_{MC_1}) are marked with thick dotted lines.

second empirical standard deviations computed using the block Toeplitz estimates (denoted as $SD_{MC_{2S}}$ in Tables 5 and 6) were slightly higher than the comparable empirical standard deviations computed using the PED estimates. This was due largely to the relatively large RMSEs associated with some of the block Toeplitz parameter estimates. As a whole, standard error estimates obtained from the PED-based observed information matrix gave close approximations to the true empirical standard errors across both simulation conditions (see Figures 3b and 4b; see also Tables 5 and 6). However, the block Toeplitz-based standard error estimates for Ψ_s and B_s consistently deviated from the true standard errors across both simulation conditions, particularly when $N = 1$. The relative discrepancies in standard error estimates [given by



(a) $T = 200$, $N = 10$



(b) $T = 200$, $N = 10$

FIGURE 4 (a) Selected parameter estimates obtained using the prediction error decomposition (PED) and structural equation modeling (SEM) and (b) their corresponding standard error estimates. Here, $N = 10$ and $T = 200$; P.ly1 $= \hat{\lambda}_{21}$ by PED; S.ly1 $= \hat{\lambda}_{21}$ by SEM; P.ps11 $= \hat{\psi}_{11}$ by PED; S.ps11 $= \hat{\psi}_{11}$ by SEM; P.a21 $= \hat{a}_{21}$ by PED; S.a21 $= \hat{a}_{21}$ by SEM; P.TE1 $= \hat{\theta}_{11}$ by PED; S.TE1 $= \hat{\theta}_{11}$ by SEM. In (a), the true parameter values are marked with thick dotted lines, and in (b) the true empirical standard deviations (i.e., SD_{MC_1}) are marked with thick dotted lines.

$(\hat{SE} - SD_{MC})/SD_{MC}$] were comparable across both simulation conditions, with a slight trend for reduced discrepancies with larger N.

The state-space approach has one other useful property that the block Toeplitz approach does not provide. That is, the KF can be readily used with the state-space model to obtain subject-specific longitudinal factor score estimates. In Figure 5, we plotted the true longitudinal factor scores of two "subjects" randomly drawn from the Monte Carlo simulations with $T = 200$ and $N = 10$, in conjunction with their estimated factor scores obtained by using the KF. The KF-based factor score estimates, generated by using the estimated parameters, approximated the true factor scores closely.

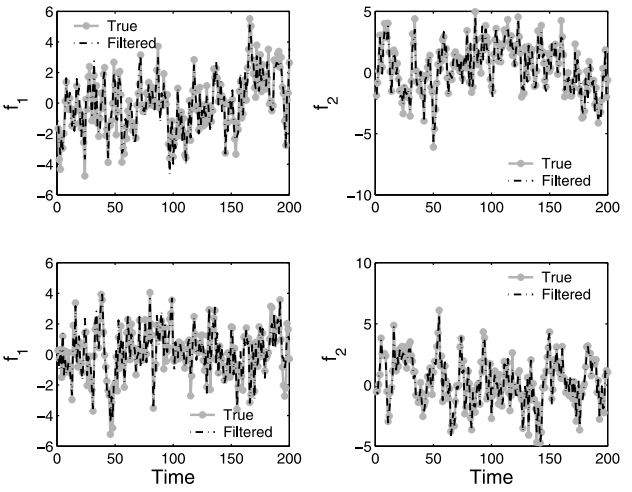


FIGURE 5 The true factor scores and estimated factor scores given by the Kalman factor from two randomly selected "subjects." True factor scores are represented in the four plots using thicker lines with asterisk, and the corresponding smoothed estimates are plotted using thin dotted-dashed lines.

In summary, the state-space approach yielded maximum likelihood parameter estimates that were unbiased. Parameters in Ψ_s and B_s obtained using the block Toeplitz method showed higher biases and consequently, led to an inaccurate statistic conclusion regarding the dynamics of the system when $N = 1$ and $T = 200$. Compared with the block Toeplitz method, the state-space approach has three additional useful properties. First, standard error estimates output from the PED-based observed information matrix can readily be used to assess the precision of the parameter estimates. Second, because the within-person time dependency in the data is accounted for explicitly in the KF as well as in the PED, the associated log-likelihood statistics can be used to compare the fit of nested models. This is not the case with the pseudo-maximum likelihood approach implemented using the block Toeplitz matrix. Third, unlike the latter, longitudinal factor score estimates can be obtained directly by running the KF or other related Kalman smoothers, and they provide a reasonable approximation to the true factor scores.

DISCUSSION

Implications of Simulation Results

SEM and the software associated with it have collectively formed one of the most popular modeling tools in social sciences. The general model discussed by Jöreskog (1974) and later popularized within the LISREL framework (Jöreskog & Sörbom, 1981) was proven to be general enough for fitting all instances of linear models. The state-space modeling approach and the KF, in contrast, have received relatively little attention from psychometricians despite their long history in other disciplines (e.g., econometrics and engineering). The two approaches, as we have explicated, have great commonalities but are nevertheless not merely replicates of

one another. More specifically, the unique features of state-space modeling techniques make them particularly conducive for representing intensive within-person changes. Likewise, SEM techniques are a much more convenient choice for representing interindividual differences in cross-sectional settings and interindividual differences in intraindividual change in panel studies. From a modeling standpoint, we see great promise in treating the two as complementary modeling tools, rather than labeling them as a replacement (or special case) of the other. In some cases, features of these two approaches have been combined to fit newer dynamic models (e.g., Oud & Jansen, 2000).

One key condition that is needed to establish the equivalence between SEM and state-space approaches is that the same initial condition specification has to be used in both settings. Specification of initial condition is indeed an issue that deserves further attention. Within the SEM tradition, it is not until recently that SEM researchers have begun to consider this issue at length (see, e.g., the work of Browne & Nesselroade, 2005; du Toit & Browne, 2001). In contrast, initial condition has been known to play an important role in influencing the results of filtering and smoothing, as well as parameter estimation, particularly when T is small (e.g., De Jong & Mazzi, 2001). Indeed, different options for handling unknown initial condition in state-space models have been proposed by various researchers over the years (e.g., Ausley & Kolm, 1985; De Jong, 1991; Harvey, 2001). Some of these approaches can be readily generalized to the SEM framework.

The latent difference score model presented by McArdle and Hamagami (2001) is a prototypical example of longitudinal models often considered in the human development and aging literature. The PED was used to estimate all parameters, including those associated with the random effects components. Analogous to the widely known growth curve models (Meredith & Tisak, 1990), disturbance terms (e.g., z_{1it}–z_{4it} in the autoregressive dynamic factor model in simulation Example 4) are omitted at the latent level in the original version of the dual change score model considered by McArdle and Hamagami (2001). As suggested by these authors, this does not necessarily have to be the case. Modeling uncertainties in manner of transitory or random shocks (Browne & Nesselroade, 2005) can be an interesting addition to the model from a theoretical perspective. In particular, if multiple indicators are used to identify a latent factor, the random shocks are a disturbance component that captures a certain amount of common variability among all indicators. However, the shocks do not, in themselves, show continuity over time (Chow et al., 2004). Such transitory shock terms can help distinguish stable intraindividual trends from labile common influences (i.e., process noise such as daily hassles) and influences that are unique to an indicator (i.e., measurement error). Multiple-indicator and stochastic extensions of the latent difference score model along this line can be readily accommodated within the state-space framework even with large T and irregularly spaced data. The ease of implementing multivariate measurement models in a state-space context when dynamic processes are involved might also help change the deeply rooted bias in favor of deterministic models.

The flip side of the story is that because all measurement occasions are typically incorporated simultaneously into an SEM model, freeing the constraint of time-invariant parameters in an SEM-based latent difference score model (e.g., the constant slope parameter and the self-feedback parameter) is relatively trivial. In the state-space formulation we considered in this context, the constant slope parameter, as part of the state vector, can be allowed to be time-varying in a straightforward manner. Allowing the self-feedback parameter to be time-varying, in contrast, would require a different type of model specification (e.g., by specifying $T = 1$

and allowing the model for that one time point to include information from all time points as in SEM, or by introducing the self-feedback parameter as a state variable, which makes the dynamic model nonlinear). Some specifications are more convenient for certain model-fitting purposes than others. This is a modeling choice that the researcher would have to make.

Our last example focused explicitly on the practicality of state-space techniques in handling data with $T > N$. In an earlier paper, Hamaker, Dolan, and Molenaar (2002) showed that the block Toeplitz method, when used with longitudinal data with $T > N$, produces moment estimates that are asymptotically close to maximum likelihood estimates for manifest autoregressive processes of lag q (i.e., $AR(q)$), as well as vector autoregressive (VAR) processes given correctly specified lag. The estimates associated with moving average (MA) models were in contrast, biased. Generalizations of such results to multiple-indicator VAR and MA processes have been addressed in other, more recent simulation studies (see also Z. Zhang, Hamaker, & Nesselroade, 2008). Here, we found that in a multiple-indicator VAR(1) model, estimates in the structural model showed more biases than estimates that appeared in the measurement model (e.g., factor loadings and the measurement error variances).

As a whole, the block Toeplitz method remains a practical method for estimating parameters in dynamic factor analysis models with VAR relations among factors. It also provides a foundation for developing newer modeling extensions such as exploratory dynamic factor analysis (e.g., using the free computer program DyFA; Browne & Zhang, 2005) and dynamic factor analysis with categorical data (e.g., G. Zhang, 2010). However, the extent of the biases in the structural parameters has to be further established in more elaborate simulation studies. This approach is also more limited for model comparison and standard error estimation purposes. On the contrary, the state-space procedures are well suited for fitting time series models and deriving standard errors and state estimates (e.g., factor scores), even in the presence of missing data (see Section 6.4 in Shumway & Stoffer, 2004).

Future Extensions and Conclusion

Due to disciplinary differences, social scientists have been relatively slow in adopting and propagating recent progress in the state-space modeling and the KF literature over the last decade. The KF technique is often used to update state estimates in real time for navigational purposes as data become available. This does not seem to be relevant to social science applications at first glance. However, with the advent of modern technology, many psychologists have begun to collect bio-signals, such as EEG, ECG, EMG, skin conductance, body temperature, respiration rate, and heart rate. Such data have been used to study the psychophysiological characteristics of emotion (Mauss, McCarter, Levenson, Wilhelm, & Gross, 2005) and neuropsychological disorders. The KF as an online factor score estimator can also provide a means of monitoring online changes in psychophysiological process in clinical settings (for an application of ECG online monitoring in a surgery through the KF procedure, see Daumer & Falk, 1998).

Some of the newer developments in the areas of nonlinear Kalman filtering, regime switching, and importance sampling techniques (Chow, Ferrer, & Nesselroade, 2007; Durbin & Koopman, 2001; Julier, Uhlmann, & Durrant-Whyte, 1995; Kim & Nelson, 1999) also offer new possibilities for modeling nonlinear non-Gaussian, as opposed to strictly linear, Gaussian dynamic processes. In addition, the random effects components we considered in the latent difference scores example can be formulated within a state-space framework relatively easily be-

cause the associated "parameters" are simply state variables (e.g., level and slope) in the model. In other cases, specification of random effects might not be as straightforward and the issue of evaluating high-dimensional integrals in fitting models with random effects still exist. Some alternative approaches for doing so have been proposed (e.g., Hamerle, Singer, & Nagl, 1993).

Just as social scientists accustomed to the modeling conventions of SEM can benefit from ideas inspired by state-space modeling techniques, researchers from a state-space modeling tradition can certainly benefit from the former as well. Some of the potential extensions have already been discussed by MacCallum and Ashby (1986), including the possible incorporation of SEM-based practices, such as hypothesis testing and evaluation of goodness of fit into a state-space modeling framework. Other potential extensions include, but are not limited to, alternative methods or designs for handling incomplete data (e.g., planned incompleteness; Graham et al., 1996; McArdle, 1994; McArdle & Woodcock, 1997), power estimation (e.g., MacCallum, Browne, & Sugawara, 1996; Satorra & Saris, 1985), and options for performing multilevel modeling. Currently, the flexibility of SEM software such as Mx allows time-varying parameters or states to be estimated as a multilevel model (i.e., by using the definition function in Mx). Most of the software available for performing KF algorithms is, in contrast, either in its early development stages, or is not written with social scientists as its target users (e.g., do not compute likelihood fit statistics and standard errors). The programs mkfm6 and Ox (and its state-space library *SSfpack*) are exceptions in the sense that they include many standard modeling options commonly used in the social sciences. These features are certainly helpful additions to other available KF software or toolboxes (e.g., the Matlab toolbox ReBEL; Van der Merwe, 2003).

Decades of methodological advances in social sciences have helped spur myriads of studies aimed more at identifying interindividual differences than studying within-person changes closely. Such advances did not come without a cost—other modeling options have arguably "suffered" from the popularity of interindividual differences analysis and SEM-based techniques. Although recent progress in hierarchical linear modeling (Bryk & Raudenbush, 1987) and multilevel modeling (Laird & Ware, 1982) has helped emphasize the need to account for heterogeneity across individuals, certain aspects of intraindividual change are still consistently overlooked. Much needed are thus studies and modeling tools that are suited for testing such hypotheses. Of course, our goal here is not to discount any one approach, nor to argue that intraindividual analysis should take precedence over interindividual analysis. However, state-space modeling, as other time series and dynamic modeling techniques, helps draw researchers' focus back to an idiographic modeling orientation, which has arguably lost some momentum since the advent of interindividual differences techniques. We hope to have put this message across through our last illustrative example.

ACKNOWLEDGMENTS

We thank John Nesselroade, Nilam Ram, Steve Boker, and a few other colleagues at the University of Notre Dame for their thoughtful comments on an earlier draft of this article. The Fortran program used for model fitting in this article can be downloaded from Conor V. Dolan's Web site at http://users.fmg.uva.nl/cdolan. Preparation of this article was supported in part by NSF grant BCS-0826844, a research fellowship from the Alexander von Humboldt Foundation, and NIA grant T32 AG20500-01.

REFERENCES

Ansley, C. F., & Kohn, R. (1985). Estimation, filtering and smoothing in state space models with incompletely specified initial conditions. *Annals of Statistics, 13,* 1286–1316.

Azzelini, A. (1996). *Statistical inference based on the likelihood.* London: Chapman & Hall.

Boker, S. M., & Bisconti, T. L. (2006). Dynamical systems modeling in aging research. In C. S. Bergeman & S. M. Boker (Eds.), *Quantitative methodology in aging research* (pp. 185–229). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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.), *Contemporary psychometrics: A Festschrift for Roderick P. McDonald (multivariate applications book)* (pp. 415–452). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Browne, M. W., & Zhang, G. (2005). *DyFA: Dynamic factor analysis of lagged correlation matrices, version 2.03* [computer software and manual]. Available at: http://quantrm2.psy.ohio-state.edu/browne/

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101,* 147–158.

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.

Chow, S.-M., Ferrer, E., & Nesselroade, J. R. (2007). An unscented Kalman filter approach to the estimation of nonlinear dynamical systems models. *Multivariate Behavioral Research, 42,* 283–321.

Chow, S.-M., Nesselroade, J. R., Shifren, K., & McArdle, J. J. (2004). Dynamic structure of emotions among individuals with Parkinsons disease. *Structural Equation Modeling, 11,* 560–582.

Daumer, M., & Falk, M. (1998). On-line change-point detection (for state space models) using multi-process Kalman filters. *Linear Algebra and Its Applications, 284*(1–3), 125–135.

De Jong, P. (1991). The diffuse Kalman filter. *The Annals of Statistics, 19,* 1073–1083.

De Jong, P., & Mazzi, S. (2001). Modeling and smotohing unequally spaced sequence data. *Statistical Inference for Stochastic Processes, 4,* 53–71.

Dolan, C. V. (2002). *MKF provisional documentation* (Unpublished mkfm6 user manual), University of Amsterdam.

Dolan, C. V., & Molenaar, P. C. M. (1991). A note on the calculation of latent trajectories in the quasi Markov simplex model by means of regression method and the discrete Kalman filter. *Kwantitatieve Methoden, 38,* 29–44.

Doornik, J. A. (1998). *Object-oriented matrix programming using Ox 2.0*. London: Timberlake Consultants Press.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state-space methods*. New York: Oxford University Press.

du Toit, S. H. C., & Browne, M. W. (2001). The covariance structure of a vector ARMA time series. In R. Cudeck, S. H. C. du Toit, & D. Sorbom, *Structural equation modeling: Present and future* (pp. 279–314). Chicago, IL: Scientific Software International.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8,* 128–141.

Gelb, A. (1974). *Applied optimal estimation*. Cambridge, MA: MIT Press.

Goodwin, G. C., & Payne, R. L. (1977). *Dynamic system identification: Experiment design and data analysis*. New York: Academic Press.

Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research, 31,* 197–218.

Gregson, R. A. M. (1983). *Time series in psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2003). ARMA-based SEM when the number of time points *T* exceeds the number of cases *N*: Raw data maximum likelihood. *Structural Equation Modeling, 10,* 352–379.

Hamerle, A., Singer, H., & Nagl, W. (1993). Identification and estimation of continuous time dynamic systems with exogenous variables using panel data. *Econometric Theory, 9,* 283–295.

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*. Princeton, NJ: Princeton University Press.

Harvey, A. C. (2001). *Forecasting, structural time series models and the Kalman filter*. Cambridge, UK: Cambridge University Press.

Heath, R. A. (2000). *Nonlinear dynamics: Techniques and applications in psychology*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hershberger, S. L., Corneal, S. E., & Molenaar, P. C. M. (1994). A dynamic factor analysis of the emotional response patterns underlying stepdaughter–stepfather relationships. *Structural Equation Modeling, 2,* 31–52.

Ho, M-H. R., Shumway, R., & Ombao, H. (2006). State-space models for longitudinal data with applications in the biological and social sciences. In T. Walls & J. L. Shafer (Eds.), *Models for intensive longitudinal data* (pp. 148–175). New York: Oxford University Press.

Jones, C. J., & Nesselroade, J. R. (1990). Multivariate, replicated, single-subject designs and P-technique factor analysis: A selective review of the literature. *Experimental Aging Research, 16,* 171–183.

Jöreskog, K. G. (1974). Analyzing psychological data by structural analysis of covariance matrices. In D. H. Krantz, R. C. Atkinson, R. L. Duncan, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Vol. 2. Measurement, psychophysics and neural information processing* (pp. 1–56). San Francisco, CA: W. H. Freeman.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL: Analysis of linear structural relationships by maximum likelihood and least squares methods.* Chicago, IL: National Educational Resources.

Julier, S. J., Uhlmann, J. K., & Durrant-Whyte, H. F. (1995, June). A new approach for filtering nonlinear systems. In *Proceedings of the American Control Conference* (pp. 1628–1632). Seattle, WA: IEEE.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering, 82*(Series D), 35–45.

Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge, MA: MIT Press.

Koopman, S. J., Shephard, N., & Doornik, J. A. (1999). Statistical algorithms for models in state space using ssfpack 2.2. *Econometrics Journal, 2,* 113–166.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics, 38,* 963–974.

Lawley, D. N., & Maxwell, M. A. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworths.

Ljung, L., & Söderström, T. (1983). *Theory and practice of recursive identification*. Cambridge, MA: MIT Press.

MacCallum, R. C., & Ashby, G. F. (1986). Relationships between linear systems theory and covariance structure modeling. *Journal of Mathematical Psychology, 30*(1), 1–27.

MacCallum, R. C., Browne, M., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149.

Mauss, I., McCarter, L., Levenson, R., Wilhelm, F., & Gross, J. (2005). The tie that binds? Coherence among emotion experience, behavior and physiology. *Emotion, 5,* 175–190.

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29,* 409–454.

McArdle, J. J., & Epstein, D. B. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58*, 110–133.

McArdle, J. J., & Hamagami, F. (2001). Latent difference score structural models for linear dynamic analysis with incomplete longitudinal data. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 139–175). Washington, DC: American Psychological Association.

McArdle, J. J., & Woodcock, R. W. (1997). Expanding test-retest design to include developmental time-lag components. *Psychological Methods, 2,* 403–435.

McArdle, J. J., Hamagami, F., Jones, K., Jolesz, F., Kikinis, R., Spiro, A. I., et al. (2004). Structural modeling of dynamic changes in memory and brain structure using longitudinal data from the normative aging study. *The Journal of Gerentology: Psychological Studies, 59,* 294–304.

McArdle, J. J., & McDonald, R. P. (1984). Some algebraic properties of the reticular action model for moment structures. *British Journal of Mathematical and Statistical Psychology, 27,* 234–251.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55,* 107–122.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50*, 181–202.

Molenaar, P. C. M., & 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.

Molenaar, P. C. M., & Raijmakers, M. E. J. (1998). Fitting nonlinear dynamical models directly to observed time series. In K. M. Newell & P. C. M. Molenaar (Eds.), *Applications of nonlinear dynamics to developmental process modeling* (pp. 269–297). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *Mx: Statistical modeling* (5th ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.

Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. M. (2002). Alternative dynamic factor models for multivariate time-series analyses. In D. M. Moskowitz & S. L. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data: Advances and techniques* (pp. 235–265). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Otter, P. (1986). Dynamic structure systems under indirect observation: Indentifiability and estimation aspects from a system theoretic perspective. *Psychometrika, 51*, 415–428.

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65*, 199–215.

Oud, J. H. L., van den Bercken, J. H., & Essers, R. J. (1990). Longitudinal factor score estimation using the Kalman filter. *Applied Psychological Measurement, 14*, 395–418.

Rao, C. R. (1958). Some statistical methods for the comparison of growth curves. *Biometrics, 14,* 1–17.

Satorra, A., & Saris, W. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*, 83–90.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147–177.

Shumway, R. H. (2000). Dynamic mixed models for irregularly observed time series. *Resenhas—Reviews of the Institute of Mathematics and Statistics, 4,* 433–456.

Shumway, R. H., & Stoffer, D. S. (2004). *Time series analysis and its applications*. New York: Springer-Verlag.

Singer, H. (1998). Continuous panel models with time dependent parameters. *Journal of Mathematical Sociology, 23*, 77–98.

Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika, 23*, 19–23.

Van der Merwe, R. (2003). *ReBEL: Recursive Baysian Estimation Library* (ReBEL online documentation), Oregon Health & Science University.

Wood, P., & Brown, D. (1994). The study of intraindividual differences by means of dynamic factor models: Rationale, implementation, and interpretation. *Psychological Bulletin, 116*, 166–186.

Zarchan, P., & Musoff, H. (2000). *Fundamentals of Kalman filtering: A practical approach.* Reston, VA: American Institute of Aeronautics and Astronautics, Inc.

Zhang, G. (2010). Dynamic factor analysis with ordinal variables. In S.-M. Chow, E. Ferrer, & F. Hsieh, *Notre Dame series on quantitative methodology (vol. 4). Statistical methods for modeling human dynamics: An interdisciplinary dialogue.* New York: Taylor & Francis.

Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (2008). Comparison of four methods for estimating a dynamic factor model. *Structural Equation Modeling, 15,* 377–402.

APPENDIX
ANALYTIC EQUIVALENCE BETWEEN SEM AND THE STATE-SPACE FRAMEWORK

When $T = 1$, all the model elements in SEM and the state-space framework are identical. As a result, the raw data likelihood function in Equation 5 and the PED function in Equation 16 are also identical. When $T > 1$, the measurement vector in SEM, y_i, is generally a $p_s \times 1$ or equivalently, a $Tp_k \times 1$ vector. In other words, y_{it} can be viewed as partitions of y_i into a total of T vectors of dimension p_k. The same relationship holds between w_s and w_k, α_s and α_k, and τ_s and τ_k. In addition, Λ_k, Ψ_k, and Θ_k are also submatrices of Λ_s, Ψ_s, and Θ_s. In this case, under the conditions that (a) $B_k = B_s = \mathbf{0}$, and (b) no further lagged relationships across measurement occasions are included in Λ_s, Ψ_s, or Θ_s, the relationships between different

components of $LL_{RML}(\theta_s)$ and $LL_{KF}(\theta_k)$ in Equations 5 and 16 can be shown to be equal. Specifically, each of the three terms in Equation 5 can be reexpressed into the corresponding three terms in Equation 16.

From the first term in Equation 5:

$$\frac{1}{2}\sum_{i=1}^{N} -p_i \log(2\pi) = \frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T} -p_{it} \log(2\pi), \text{ and}$$

from the second term in Equation 5:

$$\log|\Sigma_i| = \log|\Lambda_s \Psi_s \Lambda_s' + \Theta_s| = \log\left(\prod_{t=1}^{T} |\Lambda_k \Psi_k \Lambda_k' + \Theta_k|\right) = \sum_{t=1}^{T} \log|G_t|.$$

Furthermore, when B_s and B_k are both null matrices, the dynamic equation in the state-space model (Equation 8) reduces to

$$\eta_{i,t|t-1} = \alpha_k \text{ and}$$

From Equation 7, we can obtain:

$$y_{i,t|t-1} = \tau_k + \Lambda_k \eta_{i,t|t-1} = \tau_k + \Lambda_k \alpha_k.$$

Thus,

$$e_{i,t|t-1} = y_{it} - y_{it|t-1}, = y_{it} - (\tau_k + \Lambda_k \alpha_k).$$

In the context of SEM,

$$y_i - \mu_i = y_i - (\tau_s + \Lambda_s \alpha_s). \tag{A1}$$

Based on Equations A1 and A2,

$$(y_i - \mu)' \Sigma_i^{-1} (y_i - \mu) = [y_i - (\tau_s + \Lambda_s \alpha_s)]' [\Lambda_s \Psi_s \Lambda_s' + \Theta_s]^{-1} [y_i - (\tau_s + \Lambda_s \alpha_s)]$$

$$= \sum_{t=1}^{T} [(e_{i,t|t-1}') G_t^{-1} (e_{i,t|t-1})]$$

Thus, under these constraints, the raw data maximum likelihood function in Equation 5 and the PED function in Equation 16 are identical. More complicated scenarios arise when B_s and B_k are not null matrices. In this case, B_s can accommodate concurrent regression parameters among latent variables but B_k can only do so with special reparameterization of the state-space model. The structure of B_s that is needed to establish such equivalence is demonstrated in Examples 2 and 3.