



Published in final edited form as:

*Multivariate Behav Res.* 2019 ; 54(2): 246–263. doi:10.1080/00273171.2018.1519406.

## A Limited Information Estimator for Dynamic Factor Models

**Zachary F. Fisher,**

University of North Carolina at Chapel Hill

**Kenneth A. Bollen,**

University of North Carolina at Chapel Hill

**Kathleen M. Gates**

University of North Carolina at Chapel Hill

### Abstract

Structural Equation Modeling (SEM) is an increasingly popular method for examining multivariate time series data. As in cross-sectional data analysis, structural misspecification of time series models is inevitable, and further complicated by the fact that errors occur in both the time series and measurement components of the model. In this paper we introduce a new limited information estimator and local fit diagnostic for dynamic factor models within the SEM framework. We demonstrate the implementation of this estimator and examine its performance under both correct and incorrect model specifications via a small simulation study. The estimates from this estimator are compared to those from the most common system-wide estimators and are found to be more robust to the structural misspecifications considered.

### Introduction

Cross-sectional statistical methods are often ill-equipped to assess psychological theories which describe phenomena as processes occurring within individuals. For this reason applied researchers are increasingly looking to supplement traditional nomothetic approaches with those capable of capturing the time-dependent variability inherent to many psychological processes. Of these methods dynamic factor models (DFMs) are one of the most well-developed for studying multivariate change at the individual level. Although we can trace the use of DFMs in psychological research back to P-technique factor analysis (Cattell, Cattell, & Rhymer, 1947) and the initial reactions to it (Anderson, 1963; Cattell, 1963; Holtzman, 1963), the past two decades have seen increases in both the application of DFMs to substantive research questions, as well as methodological development efforts (Chow & Zhang, 2013; Molenaar, 2017).

A number of estimation methods have been proposed for obtaining the parameters of linear DFMs. The most common family of estimators is based on maximum likelihood (ML) principles. Of the approaches employed in the psychological literature for estimating DFMs the application of ML estimation procedures to autocovariance matrices arranged in block-Toeplitz form is well-studied. Here, although ML routines are used the estimates themselves have been labeled “pseudo-maximum likelihood” (pseudo-ML; Molenaar & Nesselroade, 1998) as sample lagged covariance matrices do not follow a Wishart distribution. Within the

structural equation modeling (SEM) framework, methods for obtaining true ML estimates based on (1) a raw data ML function (Hamaker, Dolan, & Molenaar, 2003; Singer, 2010; Voelkle, Oud, von Oertzen, & Lindenberger, 2012) where the observed-data log-likelihood is maximized directly and (2) the Expectation Maximization (EM) algorithm for maximizing the expected complete-data log-likelihood (Lee, 2010) have also been established.

A least squares family of estimators has also been considered by Browne and Zhang (2005, 2007) who developed an OLS estimation method based on autocorrelation matrices for obtaining point estimates of DFM parameters and their associated standard errors (G. Zhang & Browne, 2010; G. Zhang, Browne, Ong, & Chow, 2014) and by Molenaar and Nesselroade (1998) who implemented the asymptotically distribution-free (ADF) method of Browne (1984). In terms of input matrices it should be noted that both the OLS and ADF approaches also utilize the block-Toeplitz form. Bayesian approaches using Gibbs sampling for both continuous (Justiano, 2004; Z. Zhang, Hamaker, & Nesselroade, 2008) and categorical (Z. Zhang & Nesselroade, 2007) indicators have also been studied. Finally, within the State Space modeling framework ML estimates can be obtained using the Kalman filter and smoother (KF; Kalman, 1960). Chow, Ho, Hamaker, and Dolan (2010) provide a comprehensive review of state-space approaches and their relation to SEM.

We seek to further complement this long tradition of SEM-based work on dynamic factor models by demonstrating how to specify and estimate dynamic factor models using Bollen's (1996; 2001) Model Implied Instrumental Variable (MIIV) technique combined with a Two Stage Least Squares (2SLS) estimator. Among the advantages of the MIIV-2SLS are: (1) it is a limited information estimator and thus applicable to many models that cannot be estimated with the system-wide approaches discussed earlier, such as when the number of observed variables is greater than the number of available timepoints; (2) it is noniterative and hence computationally efficient; (3) it places less restrictive distributional assumptions on the errors or observed variables, (4) in the domain of limited information estimators it shares many of the properties of a system-wide ML estimator (e.g., consistency, asymptotic normality), (5) compared to system-wide estimators it is more robust to structural misspecification, and (6) it provides local (equation-level) diagnostics.

After providing a brief review of dynamic factor models, we will introduce the MIIV-2SLS approach in the context of dynamic factor models. We will review methods for obtaining ML and pseudo-ML estimates of the parameters of DFMs. Finally, we will examine the relative performance of the MIIV-2SLS, pseudo-ML, and KF estimators in the context of dynamic factor models. We will base these comparisons on simulation conditions designed to mimic properties encountered by substantive researchers, including structural misspecification.

## Dynamic Factor Models

Broadly, DFMs are used to model multivariate time series data when latent or unobserved processes are hypothesized to drive change in a set of measured variables. Under differing labels two classes of DFMs have emerged in the psychological literature. These classes have been alternatively referred to as Process Factor Analysis (PFA) and Shock Factor Analysis (SFA) models. In PFA models the change process itself occurs at the latent level - latent

states are hypothesized to both manifest and cause change, while for SFA models the common factors are conceptualized as random shocks that induce changes in an observed process differentially depending on their lag. Here the process is thought to occur at the manifest variable level. Defined in this manner for some finite lag each model brings unique advantages to the study of individual change and a thorough review of their similarities and differences can be found in Browne and Nesselrode (2005).

Although we do find the distinction between these two classes valuable in that each provide a clear and distinct model of change, we are also dubious of reifying specific classes of models, which are themselves approximations of the process of interest. Equally as likely is that a given change process occurs solely at the latent or observed variable level is that it occurs at both, and these distinctions will rarely be clear from the data alone. For these reasons we define the DFM broadly in a manner that encompasses both model classes. Although we will not consider all variations of the DFM model allowed by this specification, we will introduce and assess the MIIV-2SLS estimator in situations where the data generating process contains components from each model.

### Dynamic Factor Model Specification

With minor modifications we use the notation of Molenaar (2017) to specify a general form for the dynamic factor model:

$$\mathbf{y}_t = \boldsymbol{\alpha}_y + \sum_{u=0}^s \boldsymbol{\Lambda}_u \boldsymbol{\eta}_{t-u} + \boldsymbol{\varepsilon}_t \quad (1)$$

$$\boldsymbol{\eta}_t = \boldsymbol{\alpha}_\eta + \sum_{u=1}^p \boldsymbol{\Phi}_u \boldsymbol{\eta}_{t-u} + \sum_{u=1}^q \boldsymbol{\Theta}_u \boldsymbol{\zeta}_{t-u} + \boldsymbol{\zeta}_t \quad (2)$$

where (1) describes the observed variable or measurement model and (2) describes a vector autoregressive moving average (VARMA) latent time series. In the measurement model  $\mathbf{y}_t$  is a  $k \times 1$  vector of observations at time  $t$ ,  $\boldsymbol{\alpha}_y$  is a  $k \times 1$  vector of intercept terms,  $\boldsymbol{\Lambda}_u$  is a sequences of  $k \times m$  factor loading matrices up to order  $s$ ,  $\boldsymbol{\eta}_t$  is an  $m \times 1$  vector of latent factors at time  $t$ , and  $\boldsymbol{\varepsilon}_t$  is a  $k \times 1$  vector of unique factors at time  $t$ , and  $\text{Cov}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Xi}$ . For the latent variable time series  $\boldsymbol{\eta}_t$ ,  $\boldsymbol{\alpha}_\eta$  is a  $m \times 1$  vector of intercept terms allowing the latent time series to have a non-zero mean,  $\boldsymbol{\Phi}_u$  is a series of  $m \times m$  matrices up to order  $p$  containing the autoregressive and cross-regressive weights,  $\boldsymbol{\Theta}_u$  is a series of  $m \times m$  matrices up to order  $q$  containing the moving average weights, and  $\boldsymbol{\zeta}(t)$  is a  $m \times 1$  vector of random shocks or innovations with  $\text{Cov}(\boldsymbol{\zeta}_t) = \boldsymbol{\Psi}$ . We also assume

$\text{Cov}(\boldsymbol{\varepsilon}_{t+\ell}, \boldsymbol{\varepsilon}_t) = 0$ ,  $\text{Cov}(\boldsymbol{\varepsilon}_{t+\ell}, \boldsymbol{\zeta}_t) = 0$ , and  $\text{Cov}(\boldsymbol{\zeta}'_{t+\ell}, \boldsymbol{\zeta}_t) = 0$  for lag  $\ell > 0$ . In addition we assume  $\text{Cov}(\boldsymbol{\eta}_t, \boldsymbol{\zeta}'_t) = 0$  for all  $t$ . We can condense this notation to DFM( $p, q, s$ ) indicating a dynamic factor model of autoregressive order  $p$ , moving average order  $q$  and containing lagged factor loadings up to order  $s$ .

## ML-Based Estimation of Standard Dynamic Factor Models

Standard dynamic factor analysis typically refers to the application of DFMs to weakly stationary time series measured at equidistant intervals. A process can be said to be weakly stationary if it satisfies the following three conditions: (1) the mean of the process is independent of  $t$ , (2) the variance of the process is independent of  $t$ , and (3) the autocovariance of the process at two points in the series,  $x_t$  and  $x_{t-\ell}$  depends only on their difference, or lag,  $\ell$  (Harvey, 1981). Although DFMs have been extended to handle nonstationary time series (see Chow & Zhang, 2013; Chow, Zu, Shifren, & Zhang, 2011; Molenaar, 1994; Molenaar, Gooijer, & Schmitz, 1992) and unequal measurement intervals (Voelkle & Oud, 2013), we limit our attention to the standard case. In this context modeling parameters are constrained to equality over time and this invariance is taken as a necessary condition for the stationarity assumption to hold. As ML and pseudo-ML estimation are the dominant approaches previously considered we highlight some notable aspects of each for single-subject dynamic factor models before introducing the MIIV-2SLS estimator.

**Pseudo-ML.**—The application of SEM software to block-Toeplitz autocovariance matrices is one of the most common methods for fitting dynamic factor models. A number of simulation studies have examined the nature of parameter estimates obtained from this approach (Chow et al., 2010; Hamaker, Dolan, & Molenaar, 2002; Hamaker et al., 2003; van Buuren, 1997; Voelkle et al., 2012; Z. Zhang et al., 2008). Results from these simulations have been mixed. For example, parameter estimates obtained from the pseudo-ML approach by Z. Zhang et al. (2008) showed mean absolute biases larger than the KF and least squares estimators, but smaller than those obtain from a Bayesian approach. A simulation conducted by Chow et al. (2010) showed larger biases for the structural weights and random shock variance and covariance parameters for the pseudo-ML approach compared to the KF. On the other hand, Hamaker et al. (2002) showed that for observed univariate time series measured without error pseudo-ML estimates were equivalent to ML estimates asymptotically for pure autoregressive processes, while decrements in efficiency and increases in bias were observed for models containing moving average parameters. Furthermore, the pseudo-ML approach is less prone to estimation problems such as nonconvergence relative to other full-information SEM-based approaches.

**Full Information Maximum Likelihood.**—Although not as widely recognized, it is also possible to obtain true ML parameter estimates for single-subject dynamic factor models using SEM software (Hamaker et al., 2003; Singer, 2010; Voelkle et al., 2012). The only requirement for obtaining ML estimates is a fitting function which does not contain the log determinant of the sample covariance matrix, which is singular in the case of  $N=1$ . The  $\log|S|$  does not occur in the exact likelihood function (Bollen, 1989) or the case-wise full information maximum likelihood (FIML) fitting function commonly employed for handling missing data in many SEM programs. In addition to obtaining true ML estimates, this approach also brings a number of additional benefits which we mention but do not address further. There are, however, also limitations associated with the ML approach. One potential difficulty results from the computational demands associated with the storage and inversion of a model-implied covariance matrix (du Toit & Browne, 2007) which becomes prohibitively large as the number of time points increases. Singer (2010) traced many of the

convergence problems associated with ML to very large, and negative, complex eigenvalues in the theoretically positive semi-definite manifest variable covariance matrix. Preliminary results from the present authors suggested FIML was not feasible for investigating the estimation of structurally misspecified models due to exceedingly high rates of nonconvergence and improper solutions.

**State Space and Kalman Filter.**—Methods for estimating dynamic factor models within the state-space modeling framework are increasingly popular and have fostered a number of promising extensions (Chow & Zhang, 2013; Chow et al., 2011). The State Space framework can also be used to obtain true maximum likelihood estimates. The mechanics of the KF and other associated filters and smoothers employed for estimating DFMs are beyond the scope of this paper. An introductory treatment of the relationship between the Kalman filter and smoother for those familiar with SEM-based modeling is given by Chow et al. (2010). Following Voelkle, Brose, Schmiedek, and Lindenberger (2014) we use the KF to obtain ML estimates since it circumvents many of the issues that arise with the row-wise likelihood expression discussed previously.

## The MIIV-2SLS Approach

Bollen's (1996; 2001) MIIV-2SLS estimator brings a number of advantages to the estimation of dynamic factor models. We will briefly discuss the properties of MIIV-2SLS most relevant to latent time series analysis before demonstrating the application of MIIV-2SLS to the specification and estimation of dynamic factor models.

MIIV-2SLS is a limited information estimator in that it estimates a subset of model parameters at a time, equation-by-equation. This stands in contrast to full-information estimators such as ML which simultaneously evaluate the entire system of equations, estimating all model parameters concurrently. A consequence of this difference is that limited information estimators tend to better isolate structural misspecifications, which are ubiquitous in practice. Misspecifications can result from the omission of paths or variables from a model, specification of the wrong functional form, or any situation where there is a failure to correctly map the true relationships between variables in a model. That full information estimators are less robust to structural misspecification follows logically. These estimators draw information from the entire system to estimate individual parameters, including any portion of the system which is incorrect. As a result errors from one part of the model can affect other areas, even those which are specified correctly. That the MIIV-2SLS estimator is more robust to structural misspecification than the maximum likelihood estimator is supported by recent simulation results (Bollen, Kirby, Curran, Paxton, & Chen, 2007; Nestler, 2014a, 2014b).

Overidentification tests are another valuable tool associated with the MIIV-2SLS approach. The structure of a given model implies that certain observed variables are uncorrelated with the composite error of an equation and some are not. The former are the Model Implied Instrumental Variables (MIIVs). In the context of MIIV-2SLS, overidentification tests evaluate the assumption that instruments for a given equation are uncorrelated with the composite disturbance for that equation. Overidentification tests are available for each

overidentified equation in the model. In terms of dynamic factor models more often than not all model equations will be overidentified. In the context of MIIV-2SLS, overidentification tests are tests of the current model specification as the model structure itself selects the MIIVs (Kirby & Bollen, 2009). For this reason overidentification tests can be used to identify problematic equations.

The MIIV-2SLS estimator is also noniterative and greatly increases computational efficiency. This quality is particularly relevant to DFA where problems with convergence are commonly encountered. A number of recent extensions have further increased the utility of MIIV-2SLS. These developments include the estimation of categorical endogenous variables (Bollen & Maydeu-Olivares, 2007; Jin, Luo, & Yang-Wallentin, 2016; Nestler, 2013); latent variable interactions (Bollen, 1995; Bollen & Paxton, 1998); second order growth curve models (Nestler, 2014b); higher order factor analysis (Bollen & Biesanz, 2002); specification error tests for nonlinearity and interactions (Nestler, 2015); testing the dimensionality of measures (Bollen, 2011) and generalized method of moments estimation (Bollen, Kolenikov, & Bauldry, 2014).

### Transforming the DFM

To estimate the parameters of the DFM using MIIV-2SLS we must transform the latent variable model into an observed variable form. Similar transformations have been examined by Bollen (1996) and Molenaar (2003). For the purpose of introducing the MIIV-2SLS estimator to dynamic factor analysis we demonstrate this transformation using a latent autoregressive process with multiple indicators, arguably the most common dynamic factor model implemented in psychological applications. We note this model can be equivalently written as DFM( $p, 0, 0$ ) or PFA( $p$ ), where  $p$  is the autoregressive order of the latent time series. To estimate this model using the MIIV-2SLS estimator we must first scale the latent time series by setting the intercept to zero and factor loading to one in (1) for a single indicator per factor. We can then partition the indicators such that  $\mathbf{y} = [\mathbf{y}^{(s)} \mathbf{y}^{(n)}]'$  where  $\mathbf{y}^{(s)}$  contains the indicators used to scale  $\boldsymbol{\eta}$  and  $\mathbf{y}^{(n)}$  contains the remaining nonscaling indicators. A consequence of this scaling choice is that we can now redefine the measurement equations for the scaling indicators such that each common factor can be expressed as the difference between its scaling indicator and unique factor.

Noting the mathematical relationship that follows from our scaling choice

$$\mathbf{y}_t^{(s)} = \boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t^{(s)} \Rightarrow \boldsymbol{\eta}_t = \mathbf{y}_t^{(s)} - \boldsymbol{\varepsilon}_t^{(s)} \quad (3)$$

we can rewrite the measurement model

$$\mathbf{y}_t^{(n)} = \boldsymbol{\alpha}_y^{(n)} + \boldsymbol{\Lambda}^{(n)} \mathbf{y}_t^{(s)} + \left( \boldsymbol{\varepsilon}_t^{(n)} - \boldsymbol{\Lambda}^{(n)} \boldsymbol{\varepsilon}_t^{(s)} \right), \quad (4)$$

and latent variable time series

$$\mathbf{y}_t^{(s)} = \boldsymbol{\alpha}_\eta + \sum_{u=1}^p \boldsymbol{\Phi}_u \mathbf{y}_{t-u}^{(s)} + \left( \boldsymbol{\varepsilon}_t^{(s)} - \sum_{u=1}^p \boldsymbol{\Phi}_u \boldsymbol{\varepsilon}_{t-u}^{(s)} + \boldsymbol{\zeta}_t \right) \quad (5)$$



with the composite disturbance term  $\tilde{\mathbf{u}} = (\epsilon_t^{(s)} - \sum_{u=1}^p \Phi_u \epsilon_{t-u}^{(s)} + \zeta_t)$  inside the parentheses. In (5) we have recast the system of latent variable AR( $p$ ) time series (2) into a set of manifest ARMA( $p,p$ ) processes, where the moving average order is equal to the latent variable autoregressive order (Box, Jenkins, & Reinsel, 2008; Granger & Morris, 1976). Here the moving average weights are functions of the autoregressive weights, composite disturbance variances and lagged covariances, while the autoregressive weights remain the same between the latent VAR( $p$ ) model and the manifest VARMA( $p,p$ ) series. The transformation of structural models with white-noise error terms into models where the disturbance is no longer white-noise is common in the time series literature (Bowden & Turkington, 1990). In addition to making the model estimable this transformation explicitly allows for the possibility of correlation between the regressor and the error, a realistic concern in the presence of lagged endogenous variables. Here, we are not concerned with estimating the MA parameters and treat them as nuisance parameters. Our main objective is to instead obtain consistent estimates of  $\Phi$  from the latent variable model in light of these composite disturbances. Although not common in the psychometrics literature, the estimation of AR parameters in the presence of moving-average measurement error is common in signal processing and system identification (Stoica, Friedlander, & Soderstrom, 1987; Stoica, Soderstrom, & Friedlander, 1985).

To further illustrate this transformation of (4) and (5) we can write the observed form of a DFM(1, 0, 0) as

$$\begin{bmatrix} \mathbf{y}_t^{(s)} \\ \mathbf{y}_t^{(n)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_\eta \\ \boldsymbol{\alpha}_y^{(n)} \end{bmatrix} + \begin{bmatrix} \Phi \\ \Lambda^{(n)} \end{bmatrix} \mathbf{y}_{t-1}^{(s)} + \begin{bmatrix} \mathbf{I} & (\mathbf{I} - \Phi) & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -\Lambda^{(n)} & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \epsilon_t^{(s)} \\ \epsilon_{t-1}^{(s)} \\ \epsilon_t^{(n)} \\ \zeta_t \end{bmatrix}. \tag{6}$$

The transformations above have translated the structural relations from the original dynamic factor model into a system of estimating equations. Consolidating the composite disturbance we can further simplify our notation to express this system as  $\mathbf{y} = \mathbf{Z}\boldsymbol{\theta} + \tilde{\mathbf{u}}$  where  $\mathbf{y}$  is a stacked vector containing  $\mathbf{y}^{(s)}$  and  $\mathbf{y}^{(n)}$ ,  $\mathbf{Z}$  is block-diagonal and contains all relevant regressors from  $\mathbf{y}_{t-u}^{(s)}$ ,  $\boldsymbol{\theta}$  contains the free parameters in  $\Phi$  and  $\Lambda$  and  $\tilde{\mathbf{u}}$  contains the composite error terms for each equation. It is worth noting that for any equation  $i$ , the  $i$ th block of  $\mathbf{Z}$  contains the scaling indicators for any lagged or contemporaneous latent variable appearing on the right hand side of equation  $i$  and the corresponding  $\boldsymbol{\theta}_i$  will contain any estimated factor loadings or autoregression coefficients belonging to equation  $i$ . Furthermore we can define  $\tilde{\mathbf{u}}$  in terms of the correlated-shock  $(\xi_t + \epsilon_t^{(s)} - \sum_{u=1}^p \Phi_u \epsilon_{t-u}^{(s)})$  or independent-shock  $(\xi_t + \sum_{u=1}^p \Phi_u \epsilon_{t-u}^{(s)})$  ARMA( $p,p$ ) representation (Browne & Nesselrode, 2005). The difficulty in estimating  $\boldsymbol{\theta}$  from  $\mathbf{y} = \mathbf{Z}\boldsymbol{\theta} + \tilde{\mathbf{u}}$  results from the composite disturbance term  $\tilde{\mathbf{u}}$  which by construction will have a nonzero correlation with variables in  $\mathbf{Z}$ . To overcome this difficulty we can use the limited information MIV-2SLS estimator described below.

## Model-Implied Instrumental Variables

The 2SLS estimator was developed specifically for cases such as the one described above, where the equation error correlates with one or more explanatory variables in the model. 2SLS requires instrumental variables (IVs) for consistent estimation of  $\theta$ . For the matrix of IVs for equation  $i$ ,  $\mathbf{V}_i$ , to be valid the following conditions must be satisfied: (a)  $\text{Cov}(\mathbf{Z}_i, \mathbf{V}_i) \neq 0$ , (b) rank of  $\text{Cov}(\mathbf{V}_i, \mathbf{Z}_i)$  must be greater than or equal to the number of columns in  $\mathbf{Z}_i$ , (c)  $\text{Cov}(\mathbf{V}_i)$  is nonsingular, and (d)  $\text{Cov}(\tilde{\mathbf{u}}_i, \mathbf{V}_i) = 0$ .

When Condition (a) is only marginally satisfied this is referred to as the “weak instruments” problem. In this situation the actual sampling distributions of conventional statistics such as hypothesis tests, confidence intervals, and standard errors will generally be nonnormal, leading to unreliable inferences (Stock, Wright, & Yogo, 2002). Weak instrument diagnostics are described elsewhere (e.g., Stock, et al., 2002; Bollen, 2012) and these diagnostics apply here as well. Condition (c) can be assessed by checking to see if the sample IV covariance matrix has an inverse. Satisfaction of (d) requires all IVs to be uncorrelated with the composite disturbance for a given equation and this can be assessed empirically using overidentification tests such as Sargan’s  $\chi^2$ , which we describe in detail below. Unlike the majority of applications where IVs are drawn from a set of unmodeled external variables (or auxiliary instruments, see Bollen (2012)) the current procedures draw instruments from the pool of observed variables satisfying condition (d) based entirely on the model specification. Thus the MIIV-2SLS estimator is a model-based method for selecting instrumental variables and violation of (d) is a rejection of the model structure itself. In this case, the researcher will need to respecify the model. Although locating the source of a structural misspecification is always a challenge, research on the robustness to structural misspecification might help to eliminate those parts of the model that could not influence the test statistic. See Bollen (2001) and Bollen, Gates, and Fisher (2018). In addition, it is possible to calculate the Sargan test statistic for subsets of MIIVs for an equation so as to better isolate the source of the problem.

**MIIV Selection for the DFM.**—To illustrate the selection of MIIVs we first look at the measurement model for the DFM. Let the nonscaling indicator  $y_{j,t}^{(n)}$  be the  $j^{\text{th}}$  element of  $\mathbf{y}^{(n)}$ . The equation for  $y_{j,t}^{(n)}$  will include the composite disturbance term  $\tilde{\mathbf{u}} = \varepsilon_{j,t}^{(n)} - \lambda_{jk}^{(n)} \varepsilon_{k,t}^{(s)}$ . Any valid instrumental variables  $v_t$  will satisfy the following moment conditions:  $\text{Cov}(v_t, \varepsilon_{j,t}^{(n)}) = 0$ , and  $\text{Cov}(v_t, \varepsilon_{k,t}^{(s)}) = 0$ . If the factor complexity of  $y_{j,t}^{(n)}$  is greater than one, as can occur when specifying a dynamic factor model with lagged factor loadings, uniquenesses associated with the additional factors must also have a covariance of zero with the instrument. For some arbitrary lag  $\ell$ , if  $y_{j,t}^{(n)}$  also loads on  $\eta_{k,t-\ell}$ , the scaling indicator  $y_{k,t-\ell}^{(s)}$  will no longer be an eligible MIIV due to  $\varepsilon_{k,t-\ell}^{(s)}$ . From these moment conditions it is clear MIIVs for the measurement equation  $y_{j,t}^{(n)}$  can be obtained from any observed variable whose error does not correlate with  $\varepsilon_{j,t}^{(n)}$  or  $\varepsilon_{k,t}^{(s)}$  and any scaling indicator  $y_k^{(s)}$ , when  $\lambda_{jk}^{(n)} = 0$ .



For the latent variable model, let  $y_{j,t}^{(s)}$  be the  $j^{\text{th}}$  structural equation in  $\mathbf{y}^{(s)}$  with composite disturbance term  $\tilde{\mathbf{u}} = \varepsilon_{j,t}^{(s)} - \sum_{u=1}^p \phi_{jk,t-u} \varepsilon_{k,t-u}^{(s)} + \zeta_{j,t}$ . Any valid instrumental variables  $v_l$  for the  $j^{\text{th}}$  latent variable equation will satisfy the following moment conditions:  $\text{Cov}(v_l, \zeta_{j,t}^{(s)}) = 0$ ,  $\text{Cov}(v_l, \varepsilon_{j,t}^{(s)}) = 0$ , and  $\phi_{jk,t-u} \text{Cov}(v_l, \varepsilon_{k,t-u}^{(s)}) = 0$ . For single-indicator time series models a parsimonious solution for identifying IVs is to simply use lagged values of the endogenous variables. More specifically, the moving average structure of the disturbance in  $\tilde{\mathbf{u}}$  implies that  $\mathbb{E}(\tilde{\mathbf{u}} y_{k,t-u}^{(s)}) = 0$  only if  $u \geq p$ . In fact, for single indicator ARMA ( $p, q$ ) models it has been shown that estimates obtain from lagged endogenous variables corresponding to  $\mathbf{y}_{t-u+1}$  are the most asymptotically efficient of all possible lagged instrumental variables (Dolado, 1990). This is intuitive as we would expect the correlation between  $\mathbf{y}_t$  and  $\mathbf{y}_{t-u+1}$  to decay with increasing lags.

A drawback of this approach, however, is that although  $\mathbf{y}_{t-u+1}^{(s)}$  is certainly a MIIV, as its adequacy is implied at least partly by the MA( $p$ ) disturbance structure, it is not actually in the estimating model and therefore its use requires pruning the number of timepoints included in the analysis. For multiple indicator time series models another source of MIIVs for  $y_{j,t}^{(s)}$  comes from the set of nonscaling endogenous indicators. Although any lagged scaling indicator  $y_{k,t-u}^{(s)}$  with a nonzero auto- or cross-regression weight  $\phi_{jk,t-u}$  will not be an eligible MIIV, the nonscaling indicators of  $\phi_{jk,t-u} \neq 0$ , nonscaling indicators of  $\eta_{k,t-u}$  may be especially appropriate as there is some evidence to suggest estimates are superior when IVs span the factor space for a given equation (Cudeck, 1991). It is also likely these variables are highly correlated with  $y_{k,t-u}^{(s)}$ , and their implementation does not require any additional pruning of the sample size. An additional benefit of using non-scaling indicators is they will result in estimates which are more robust to omitted cross-regressive relations.

**Degree of Overidentification.**—In addition to determining the instrument validity implied by the DFM specification above we must also consider the number of MIIVs used in estimation. Condition (b) is often referred to as the rank condition and implies a necessary condition worth introducing, the order condition. The order condition states the number of instruments in  $\mathbf{V}_j$  must be greater than or equal to the number of endogenous regressors in  $\mathbf{Z}_j$ . When the rank condition is satisfied and the number of MIIVs for a given equation is equal to the number of endogenous regressors to be estimated that equation is said to be exactly identified. The degree of overidentification for each equation is equal to the number of MIIVs for that equation minus the number of endogenous regressors. As will often be the case with DFMs, the model specification will lead to more MIIVs than endogenous regressors, resulting in many overidentified equations. In the case of the DFM considered in our simulation study, for example, equation-specific degrees of overidentification implied by the model specification range from 2 to 15.

Evidence from the econometric literature has shown the bias of the 2SLS estimator is minimal in large samples but tends to increase with the degree of overidentification when the sample is small (Hall, 2005; Mikhail, 1972). This can be especially true when the

contribution of additional MIIVs to the first stage regression  $R^2$  is minimal (Bollen, 2012, p. 59). A similar pattern has been observed for the MIIV-2SLS estimator (Bollen et al., 2007). It is important to remember that even though the model specification may imply a large number of valid instruments (Condition (d)), those instruments will differ in how strongly they relate to the endogenous variables in that equation. In some cases it is possible to have a number of MIIVs which are only weakly correlated with the endogenous regressors. For this reason it is important to mention strategies for pruning excess MIIVs when necessary.

In the present context two viable options for pruning instruments were available, however, due to space considerations only the latter is considered here. The first method involves using variable selection methods in the first stage regression to identify the smallest set of instruments that jointly explain the most variation in the endogenous regressors while satisfying (d) (Serena & Jushan, 2009). A simpler option, and the one used in this paper, is to select a degree of overidentification (e.g. one) and then select MIIVs. For example, the nonscaling indicators of the same lagged latent variable are good potential MIIVs in that they should be moderately to highly correlated with the scaling indicator of the same latent variable. It is worth mentioning the pruning of MIIVs discussed here applies to the estimation of DFM model parameters only, and not to the model specification testing which will be discussed later. In the case of evaluating model specification the full set of MIIVs is recommended for detecting misspecifications.

### The MIIV-2SLS Estimator

With instruments in hand we now define the limited information MIIV-2SLS estimator in the context of generalized method of moments (GMM) estimation. GMM estimation is based on the idea that parameters  $\theta$  from an overidentified model are related to a set of data  $\mathbf{X}$  through a series of orthogonality conditions on the population moments,  $\mathbb{E}[\mathbf{g}(\mathbf{x}_i; \theta)] = 0$  (Hansen, 1982) for equation  $i$ . Here we can define the vector  $\mathbf{g}_i()$  as

$$\mathbf{g}_i = \mathbf{g}(\mathbf{x}_i; \theta_i) = \mathbf{V}_i(\mathbf{y}_i - \mathbf{Z}_i\theta_i) \quad (7)$$

where  $\mathbf{V}_i$  ( $T \times \#$  of MIIVs),  $\mathbf{y}_i$  ( $T \times 1$ ), and  $\mathbf{Z}_i$  ( $T \times \#$  of endogenous covariates) are the instrumental, outcome, and explanatory variable quantities associated with equation  $i$ . To further simplify our notation we can define  $\mathbf{S}_{V_i y_i} = T^{-1} \sum_{t=1}^T V_{it}' y_{it}$  and  $\mathbf{S}_{V_i Z_i} = T^{-1} \sum_{t=1}^T V_{it}' Z_{it}$  as the sample moments corresponding to  $\sigma_{V_i y_i}$  and  $\Sigma_{V_i Z_i}$ , respectively. The goal in estimation is to choose  $\hat{\theta}_i$  such that  $\mathbf{g}(\mathbf{x}_i; \theta_i)$  or more specifically the distance between  $\mathbf{S}_{V_i y_i}$  and  $\mathbf{S}_{V_i Z_i} \hat{\theta}_i$  is as close to zero as possible. Note that in an overidentified equation it is not possible to identify  $\hat{\theta}_i$  such that  $\mathbf{g}(\hat{\theta}_i) = 0$ . In this case we instead minimize the joint distance between the sample orthogonality conditions using a symmetric, positive-definite weight matrix  $\mathbf{W}$ ,  $T(\mathbf{S}_{V_i y_i} - \mathbf{S}_{V_i Z_i} \hat{\theta}_i)' \widehat{\mathbf{W}} (\mathbf{S}_{V_i y_i} - \mathbf{S}_{V_i Z_i} \hat{\theta}_i)$ . The choice of  $\mathbf{W}$  leads to a number of different GMM estimators, however, for the remainder of this paper we will focus on the weight matrix ( $\mathbf{W} = T(\mathbf{V}'\mathbf{V})^{-1}$ ) which gives the 2SLS estimator for equation  $i$ . Interested readers should see Hayashi (2011) for a general treatment of the various single and multiple equation GMM

estimators associated with  $\mathbf{W}$  and Bollen et al. (2014) for a MIIV-GMM estimator for latent variable models. After identifying the MIIVs for a given equation, a unique solution to the objective function given above can be obtained by solving for  $\hat{\theta}_i$

$$\hat{\theta}_i = (\mathbf{S}'_{X_i V_i} \mathbf{S}^{-1}_{V_i V_i} \mathbf{S}_{X_i V_i})^{-1} \mathbf{S}'_{X_i V_i} \mathbf{S}_{X_i y_i}. \quad (8)$$

As mentioned previously, in standard dynamic factor analysis it is a necessary, but not sufficient, condition for stationarity that modeling parameters be time-invariant. Expanding our notation to the full system of equations we can write the MIIV-2SLS estimator capable of handling cross-equation equality restrictions as

$$\begin{bmatrix} \hat{\theta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{S}'_{XV} \mathbf{S}^{-1}_{VV} \mathbf{S}_{XV} & \mathbf{R}' \\ \mathbf{R} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{S}'_{XV} \mathbf{S}_{Xy} \\ \mathbf{q} \end{bmatrix} \quad (9)$$

where  $\hat{\theta}$  again contains all the estimated structural parameters,  $\mathbf{S}_{XV}$  is a block diagonal matrix containing the equation-specific covariances of each RHS (Right Hand Side) endogenous variables with each MIIV,  $\mathbf{S}_{VV}$  is a block diagonal MIIV covariance matrix, and  $\mathbf{S}_{Xy}$  is a block diagonal matrix containing the equation-specific covariances of each RHS endogenous variables with the dependent variable. The  $(\# \text{ of restrictions} \times 1)$  vector  $\mathbf{q}$  and the  $(\# \text{ of restrictions} \times \# \text{ of estimated coefficients})$  restriction matrix  $\mathbf{R}$  are used to impose the cross-equation equality constraints implied by the dynamic factor model. Each row of  $\mathbf{R}$  and element of  $\mathbf{q}$  can represent a linear equality restriction made on the coefficient vector. If the  $i^{\text{th}}$  row of  $\mathbf{R}$  does not contain a restriction or represents an equality restriction only (i.e.,  $\lambda_{21,t-1} = \lambda_{21,t}$ ) then  $\mathbf{q}_i = 0$ . If a subset of coefficients are instead fixed to a constant (i.e.,  $\lambda_{21,t-1} = \lambda_{21,t} = 2$ ) then the corresponding element of  $\mathbf{q}$  would be set equal to 2. A vector of estimated Lagrangean multipliers,  $\hat{\gamma}$ , is also produced as a result of these restrictions but will not be discussed further here.

### Overidentification Tests for MIIV-2SLS

An advantage of the MIIV-2SLS approach for estimating DFMs is the availability of overidentification tests when the model specification leads to an excess of instruments. For dynamic factor models any non-scaling lagged indicators will be valid MIIVs, making the availability of excess instruments likely for identified models. These tests can be used to assess the assumption of orthogonality between residuals and instruments (i.e.  $\mathbb{E}(\mathbf{V}'\tilde{\mathbf{u}}) = 0$ ). Rejection of the null implies a problem with the logic underlying the choice of instruments (Woolridge, 2010), and importantly in the context of MIIV-2SLS this logic is the model specification itself. Therefore, overidentification tests can serve dual functions in SEM, identifying misspecified models and also diagnosing specific structural misspecifications through the examination of restrictions leading to the problematic instruments.

Kirby and Bollen (2009) examined the performance of overidentification tests commonly employed in simultaneous equation models with observed variables in the context of latent

variable structural equation models. In their simulations Sargan's  $\chi^2$  (Sargan, 1958) test performed well compared to the other tests considered:

$$S_{\chi^2} = \frac{T(\mathbf{S}_{V_i y_i} - \mathbf{S}_{V_i Z_i} \hat{\boldsymbol{\theta}}_i)' \mathbf{S}_{V_i V_i}^{-1} (\mathbf{S}_{V_i y_i} - \mathbf{S}_{V_i Z_i} \hat{\boldsymbol{\theta}}_i)}{\sigma_{ii}} \quad (10)$$

where all quantities are as previously defined. The null hypothesis of Sargan's test is that all MIIVs are uncorrelated with the equation's composite disturbance. The alternative hypothesis is that at least one MIIV correlates with the disturbance. If we reject the null, then this is evidence that there is a specification error in the structure that led us to select the MIIVs. For this reason it is advisable to include all valid instruments when using Sargan's test. Of course, just like the  $\chi^2$  test used in ML estimators in SEM, we need to keep in mind that the statistical power of the test increases with additional timepoints, in addition to other factors, so that high statistical power might lead to the rejection of MIIVs that are only weakly correlated with the error.

## Monte Carlo Simulations

Monte Carlo simulations were conducted to examine the finite-sample properties of the MIIV-2SLS estimator for dynamic factor models. Comparisons of MIIV-2SLS were made to the pseudo-ML and ML (via KF) estimates under varying time series lengths ( $T = 50, 100, 200, 500$ ) with and without structural misspecifications. Within each simulation condition the finite-sample properties of Sargan's  $\chi^2$  overidentification test were also examined to determine its utility in diagnosing the structural misspecifications. In addition to the simulations presented below additional simulations were conducted to assess the generality of our results under alternative model specifications. As the pattern of results proved similar across the two simulations we have included results from the latter in Appendix A.

## Data Generation

Continuous multivariate latent time series were generated in accordance with equation (2). Random shock vectors  $\boldsymbol{\xi}_t$  were generated from  $\mathcal{N}(\mathbf{0}, \boldsymbol{\Psi})$ . Subsequently, the observed multivariate time series,  $\mathbf{y}_t = 1, \mathbf{y}_t = 2, \dots, \mathbf{y}_t = T$ , were generated according to (1). The vector of measurement errors at each time point were generated from  $\mathcal{N}(\mathbf{0}, \boldsymbol{\Xi})$ . As in Z. Zhang et al. (2008) and Voelkle et al. (2012) a burn-in of 1,000 time points was used to attenuate any persisting effects of the initial parameters. 500 datasets were generated for each block of the simulation design. The individual model specifications used to generate and fit individual datasets will be described in detail below.

As previously mentioned we are concerned with the application of dynamic factor analysis to covariance-stationary processes such that  $\mathbb{E}(\mathbf{y}_t) = \mathbb{E}(\mathbf{y}_{t'})$  and  $\text{Cov}(\mathbf{y}_t, \mathbf{y}_{t'+u}) = \text{Cov}(\mathbf{y}_{t'}, \mathbf{y}_{t'+u})$  for  $u = \pm 0, \pm 1, \pm 2, \dots$ . Therefore, for each condition elements of the autoregressive weight matrix,  $\boldsymbol{\Phi}$ , were chosen both to span a range of plausible coefficient values, but also to ensure the generated time series are stationary. The stability conditions for a VARMA( $p, q$ ) process require that all eigenvalues of  $\boldsymbol{\Phi}$  have modulus less than 1 (Lutkepohl, 2007). If this requirement holds, the process can be considered covariance-stationary.

## Model Specification

The following data generating parameters were employed across all simulation conditions

$$\Phi = \begin{bmatrix} 0.7 & -0.2 \\ -0.2 & 0.5 \end{bmatrix}, \quad \Lambda = [\Lambda_0 \Lambda_1] = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 1.0 & 0 & 0 & 0 \\ 1.0 & 0 & 0.5 & 0 \\ 0 & 1.0 & 0 & 0 \\ 0 & 1.0 & 0 & 0 \\ 0 & 1.0 & 0 & 0 \end{bmatrix}, \quad \Psi = \begin{bmatrix} 0.36 & 0.18 \\ 0.18 & 0.36 \end{bmatrix},$$

with diagonal uniqueness covariance matrix  $\Xi = \text{diag}[0.3, 0.3, 0.3, 0.3, 0.3, 0, 3]$ . Here we have included both cross-regressive relationships among the latent variables, as well as a lagged factor loading, to demonstrate estimation of a DFM model with components from common to the PFA and SFA model classes. In addition, this model allows us to examine the consequences of model misspecifications which are likely to occur in practice: omission of latent variable cross-regressions and lagged factor loadings. To assess the performance of each estimator we considered three different use cases: (1) the correct model in terms of free and fixed parameters was fit to the data, (2) a model where a single lagged factor loading,  $\lambda_{33}$ , is incorrectly omitted, and (3) the case where  $\Phi$  is diagonal, the small lagged cross-regressive paths,  $\phi_{21}$  and  $\phi_{12}$ , are incorrectly omitted from the model specification. An important consequence of these choices is that we can assess the effect of each type of misspecification on the different estimators.

## Model Estimation

Pseudo-ML estimates were obtained using the structural equation modeling software lavaan (Rosseel, 2012). State space estimation using the Kalman filter and smoother in combination with an Expectation-Maximization algorithm was implemented using MARSS (Holmes, Ward, & Wills, 2012). To estimate the model described above the standard state space model specification was modified to allow for the observed indicators to depend on lagged states (Nimark, 2015, p.10). MIIV-2SLS search and estimation procedures were conducted using MIIVsem (Fisher, Bollen, Gates, & Rönkkö, 2017).

## MIIVs

To maintain consistency across all simulation conditions, including each model specification, the identical set of MIIVs was used to estimate the parameters of each DFM. Based on previous simulation results we chose to use a single degree of overidentification for all equations in the correctly specified model. Although the equation-specific degree of overidentification implied by the model exceeds one across all equations it is only required we have as many instruments as endogenous covariates. In the case of the correctly specified model this means we employed two MIIVs for the  $y$  equations with a complexity of one, and three MIIVs for the  $y_3$ ,  $\eta_1$  and  $\eta_2$  equations. In the model where  $\lambda_{33}$  is incorrectly omitted from the model specification there is one less RHS variable, leading to an additional degree of overidentification for that equation. In addition, for the case where the cross-regressive paths  $\phi_{21}$  and  $\phi_{12}$  are mistakenly omitted, both latent variable equations no longer include a cross-lagged path and thus each of those equations will also have one additional degree of

overidentification. In all cases the MIIVs themselves are identical across these parameterizations.

For the measurement model equations we chose MIIVs from the lagged versions of non-scaling indicators originating from the same factor. For example, for the  $y_{2,t}$  equation we used  $y_{3,t-1}$  and  $y_{3,t-2}$  as MIIVs. This choice makes sense for the DFM model because it is robust to both the possibility of autoregressions among the measurement errors and lagged factor loadings. For the latent variable model we chose MIIVs from the set of lagged non-scaling indicators. For example, for the  $\eta_{1,t}$  equation, which included the predictors  $\eta_{1,t-1}$  and  $\eta_{2,t-1}$ , we chose two non-scaling indicators of the lagged dependent variable,  $y_{2,t-1}$  and  $y_{3,t-1}$  and one indicator from  $\eta_{2,t-1}$ ,  $y_{6,t-1}$ . This choice also makes sense in the context of DFMs because it means the set of MIIVs will not change based on any omitted latent variable relations. To evaluate model specification using Sargan's test all valid MIIVs were employed for each equation.

## Measures

To compare the performance of each estimator we examined the relative bias and efficiency within each simulation condition. Mean relative bias was calculated as  $[\sum_{k=1}^N (\hat{\theta}_{ak} - \theta_a) / \theta_a] / N$  where  $\theta_a$  is the data generating parameter in a given simulation condition,  $\hat{\theta}_{ak}$  is the estimate for parameter  $a$  in the  $k^{\text{th}}$  Monte Carlo replication, and  $N$  is the total number of replications. Second, we examined the variability associated with each model parameter using the standard deviation of  $\hat{\theta}_a$  within a single block of the simulation,  $SD(\hat{\theta}_a)$ . To evaluate the finite sample properties of Sargan's  $\chi^2$  overidentification test in diagnosing structural misspecification two situations were considered: (1) how often does Sargan's  $\chi^2$  test incorrectly identify a correctly specified equation as misspecified, and (2) how often does Sargan's  $\chi^2$  test correctly identify an equation when it is misspecified. In both circumstances we used an  $\alpha$  of 0.05.

## Results

**Improper Solutions.**—Improper solutions were defined as correlations whose absolute magnitude was greater than 1, negative variance parameters, or autoregressive coefficients implying a non-stationary solution. Any dataset on which an estimator did not converge or produced an improper solution was dropped from the analysis. The MIIV-2SLS estimator produced 9 improper solutions, while the pseudo-ML estimator produced 12, both at the smallest sample size,  $T = 50$ . The pattern of improper solutions for the KF estimator was more complicated. For the correctly specified model, the KF did not converge for 21 datasets at  $T = 50$ , 7 datasets at  $T = 100$ , and 2 datasets at  $T = 500$ . For the model where the lagged factor loading  $\lambda_{33}$  was incorrectly omitted the KF estimator did not converge for 22 datasets at  $T = 50$ , 19 at  $T = 100$ , 10 at  $T = 250$ , and 3 at  $T = 500$ . For the model where  $\phi_{21}$  and  $\phi_{12}$  were incorrectly omitted from the model specification, the KF did not converge in 8 datasets at  $T = 50$ . To make comparisons across the estimators and model specifications only those datasets on which all estimators converged with no improper solutions were used. This left 440 datasets at  $T = 50$ , 476 datasets at  $T = 100$ , 489 datasets at  $T = 250$ , and 497 datasets at  $T = 500$  to be used for comparing the point estimates and efficiency. One side effect of this



choice is that performance for the KF estimator may appear better than otherwise as we have trimmed the most problematic cases for that estimator specifically. For the evaluation of Sargan's test all 500 datasets were used for each simulation block.

**Percentage of Relative Bias.**—The percentage of relative bias across all estimators and simulation conditions is presented in Table 1. For the correctly specified model, the absolute percentage of relative bias for the factor loadings did not exceed 2% for the KF and MIIV-2SLS estimators across all time series length. Likewise, the relative bias for the factor loadings did not exceed 5% for the pseudo-ML estimator, suggesting trivial levels of bias for the measurement model parameters across estimators. In general, the pattern of results for the time series model parameters was also similar across estimators. For example, all estimators exhibited a decreasing negative bias for the autoregressive coefficients as time series lengths increased. The pattern for the cross-regressive coefficients was not consistent in regards to sign, but compared to the autoregressive coefficients showed a smaller magnitude of relative bias, which also decreased with longer time series lengths.

For the first misspecified model, the MIIV-2SLS coefficient estimates are largely immune to the omitted lagged factor loading,  $\lambda_{33}$ . In fact, the  $\lambda_{31}$  parameter was the only DFM parameter whose estimates were affected by the misspecification, resulting in a larger positive bias. Here the effect of a misspecification in the measurement model is isolated to the measurement model, and more specifically to the equation in which the misspecification occurs. The system-wide estimators (pseudo-ML and KF) show an even larger average positive bias in the estimates of  $\lambda_{31}$  as a result of this misspecification. For all estimators this increase is relatively stable across time series lengths. In the system-wide estimators, however, the effect of the misspecified factor loading is not isolated to the misspecified equation. As a result of omitting  $\lambda_{33}$  both regression coefficients for the  $\eta_1$  equation showed a substantial increase in bias. Estimates of the autoregressive relationship  $\phi_{11}$  became positively inflated, while estimates of the negative cross-regressive coefficient  $\phi_{12}$  became smaller in magnitude, adjusting to the omitted positive loading.

For the second misspecified model, where  $\phi_{12}$  and  $\phi_{21}$  are incorrectly omitted, the MIIV-2SLS measurement model estimates are unaffected by the misspecification. For the system-wide estimators, coefficient estimates for the lagged factor loading  $\lambda_{33}$  show a small decrease in magnitude. Effects of this misspecification on the latent variable model are more pronounced. For the MIIV-2SLS estimator the autoregressive coefficients become increasingly negatively biased. Here the positive autoregressive relationship is becoming smaller to compensate for the omitted negative cross-lagged relationship. On the other hand, for the system-wide estimators, estimates of  $\phi_{11}$  and  $\phi_{22}$  become larger in response to the omitted negative relationships. The increase in bias associated with the system-wide estimators is roughly twice as large as that of the MIIV-2SLS.

**Standard Deviation of Coefficient Estimates.**—Efficiency of the estimators across all model specifications was assessed using the standard deviation of parameter estimates within each simulation block (see Table 2). No meaningful changes in efficiency were observed across the different model specifications so only the results for the correctly

specified model are reported. The pattern of variability within each estimator was similar, with increasing time series lengths resulting in less variable point estimates. Across estimators, variability was smallest for the KF, followed by the pseudo-ML and MIIV-2SLS estimators, although these differences are more pronounced at shorter time series lengths. This finding is consistent with previous results for instrumental variable estimators which have shown that small degrees of overidentification combined with small sample sizes can result in more variable estimators.

**Sargan's  $\chi^2$  Test.**—Results for Sargan's test are given in Table 3. For the correctly specified model the proportion of overidentification tests rejecting the null hypothesis ranged from 0.03 to 0.08 suggesting the test performs well in terms of Type I error across the sample sizes considered here. For the model with the omitted lagged factor loading  $\lambda_{33}$  the overidentification test should indicate a problem with the  $y_3$  equation, as the lagged  $y_1$  variable is incorrectly included in the MIIV set. At the smaller sample sizes,  $T = 50, 100$ , the test rejects approximately 20% and 32% of datasets at an  $\alpha = 0.05$ . This suggests at smaller time series lengths the test will not always detect a misspecified equation if one is present. At the larger time series lengths the test performs considerably better, detecting the misspecified equation in 74% of the datasets at  $T = 250$  and 98% at  $T = 500$ . For the model with omitted cross-regressive coefficients, the  $\eta_1$  equation would incorrectly include  $y_4$ , and the  $\eta_2$  equation would incorrectly include  $y_1$ , as MIIVs. At time series lengths of  $T = 50$ ,  $T = 100$ , and  $T = 250$  the test rejects far below the nominal rates, suggesting the test may have low power for detecting the omission of small cross-lagged relations. At the larger sample size of  $T = 500$ , the test accurately detects a problem with the  $\eta_1$  equation in 57% of datasets, and the  $\eta_2$  equation in 88%.

## An Empirical Example

To demonstrate estimation of a DFM using the MIIV-2SLS estimator we used data from Subject 1 described by Lebo and Nesselrode (1978). The data include 112 daily ratings on six items labelled *cheerful*, *happy*, *content*, *tired*, *sluggish*, and *weary*. The first three items are hypothesized to represent *well-being* and the second three items *fatigue*. The scaling indicator for each latent variable should be chosen based on theory. If no theory is available to guide this choice scaling indicators can also be chosen based on empirical information. Here, we choose the scaling indicator for each latent series as the indicator with the largest  $R^2$  (*cheerful*,  $R^2 = 0.86$  and *tired*,  $R^2 = 0.93$ ). Another consideration would be the correlation of the scaling indicator with the model-implied instruments; larger correlations would lead to stronger instruments and better performance.

In this example MIIV-2SLS point estimates were obtained using all model-implied instruments. This specification allows us to simultaneously compare the estimated coefficients with results from local overidentification tests. To begin our analysis we adopt a model specification similar to that used by Ram, Brose, and Molenaar (2013) for these data, a bivariate DFM with both lagged and cross-regressive effects among the latent variable time series (Model 1). Point estimates and associated test statistics for all models presented in Table 4. For this model we observe one significant Sargan's  $\chi^2$  test statistic,  $\chi^2(9) = 24.4$ ,  $p < 0.01$ , for the *sluggish* equation. This result suggests that at least one of the model-implied

instruments is correlated with the equation error and therefore there are problems with the larger model specification. A number of specification errors might be responsible for the observed result and we explore two of these possibilities: (1) there is an omitted lagged factor loading from *fatigue* at  $t - 1$  to *sluggish* at  $t$  (Model 2), and (2) there is an omitted residual covariance between the errors of *sluggish* and *weary* (Model 3).

Based on the specification of Model 1, lagged and contemporaneous variables *cheerful*, *happy*, *content*, and *weary*, along with lagged variables *tired* and *weary* are all valid instruments for the *sluggish* equation. In Model 2 *sluggish* also loads directly on *fatigue* at  $t - 1$  which means *tired* at  $t - 1$  would be ruled out as a valid instrument for *sluggish*. Based on 95% confidence interval constructed using a moving block bootstrap procedure the lagged loading for *sluggish* ( $\lambda_{52}^{(1)}$ ) was not significant. Furthermore, Sargan's  $\chi^2$  test statistic is still significant for the *sluggish* equation, with *tired* at  $t - 1$  removed from the instrument set,  $\chi^2(7) = 21.7, p < 0.01$ , suggesting there are still problems with this specification.

In Model 3 we hypothesize the existence of an omitted residual covariance between *sluggish* and *weary*. This might occur due to some ambiguity regarding perceived distinctions between the two adjectives. As a result of this additional covariance, *weary* is no longer considered as a valid MIIV for the *sluggish* equation. As a result of this new specification, Sargan's  $\chi^2$  test statistic is no longer significant,  $\chi^2(8) = 12.9, p = 0.12$ . We can follow-up this test with a more formal comparison of the two specifications using the C statistic described by Hayashi (2011, p. 218). This test involves taking the difference of the restricted and unrestricted Sargan's  $\chi^2$  statistic, in this case the model where *weary* is a valid instrument for the *sluggish* equation and one where it is not, which itself is  $\chi^2$  distributed with degrees of freedom equal to the number of suspect instruments. We also find the C statistic comparing the orthogonality conditions implied by Model 1 and 3 to be significant,  $\chi^2(1) = 11.5, p < 0.001$ , providing additional evidence of some omitted covariation between the two items.

The steps detailed here are not intended to provide an exhaustive or rigorous model respecification procedure, but to demonstrate at a broad level the relation between overidentification tests and the larger model specification, and how these tests might be used to localize misfit. It is also worth noting that across the three model specifications the time series parameter estimates were robust to the various measurement model specifications. Code for fitting the empirical example using the MIIV-2SLS estimator can be obtained from the first author.

## Discussion

In this paper we introduced a limited-information estimator for dynamic factor analysis based on Bollen's (1996; 2001) MIIV-2SLS estimator. We demonstrated how a DFM can be transformed from a model containing latent variables into one composed of observed time series only. We showed how instrumental variables implied by the model specification itself can be used to obtain consistent estimates of the original DFM model parameters. As latent variable time series models imply certain equality constraints we established how these constraints can be imposed on the contemporaneous and lagged sample covariance matrices

during estimation. Furthermore, we examined the performance of local misspecification tests as a means for detecting structural misspecifications in both the time series and factor analysis models.

Although there is a long tradition of employing SEM to analyze multivariate time series data from a single individual very little attention has been paid to the effects of structural misspecification in these models. In the case of dynamic factor models this is unfortunate as misspecification is possible in both the factor analysis and time series components of the model. When the DFM is correctly specified we expect all estimators to produce consistent estimates, and this result was shown empirically. Unfortunately, structural misspecification is inevitable and therefore it is essential to understand which DFM model parameters remain consistent in the face of various misspecifications. The method introduced here differs from the system-wide estimators previously considered for DFMs in that it does not require all model parameters to be estimated simultaneously and may better isolate structural errors.

To better understand the robustness properties of the MIIV-2SLS and system-wide estimators we examined two plausible structural misspecifications of a bivariate DFM. The first misspecification involved the omission of a lagged factor loading. As models without lagged factor loadings tacitly imply all indicators return to baseline levels at equivalent rates we viewed this as an important misspecification to consider. As a result of this misspecification all estimators produced biased estimates of the contemporaneous loading associated with the omitted indicator. However, for the system-wide estimators, the effects were not isolated to the measurement model parameters. Incorrect omission of a single lagged factor loading resulted in some auto and cross-regressive parameters that were no longer consistent. This was not the case for the MIIV-2SLS estimator whose time series model parameters were unaffected. The second misspecification we considered was the omission of negative lagged cross-regressive weights amongst the latent variable time series. We view this misspecification as important to consider as the degrees of freedom associated with the auto and cross-regressive paths are considerably fewer than those associated with the factor analysis model. In practice this could lead to adequate overall model fit statistics despite having incorrect constraints on the time series model parameters. In our simulations, this misspecification resulted in autoregressive weights which were no longer consistent across all estimators. This misspecification had little to no affect on the measurement model parameters.

Through our simulations we also considered the efficiency of each estimator and found the pattern of standard deviation to be similar across all model specifications within each estimator. Furthermore, the variability of the Kalman filter and smoother based estimates were the least variable, followed by pseudo-ML and MIIV-2SLS. For the case of MIIV-2SLS this result is not surprising as we chose to use a small degree of overidentification in our simulations, which is known to decrease bias but increase variability in small samples. To obtain a constant degree of overidentification across equations we eliminated MIIVs based on robustness concerns. Theoretically, MIIVs could be selected to reduce bias and variability, and future work should consider both the equation-specific degree of overidentification and MIIV selection criteria more systematically.

In addition to introducing the MIIV-2SLS estimator we also demonstrated how overidentification tests can be used to detect local misspecification in latent time series models. For the longer time series lengths considered here ( $T=250$  and  $T=500$ ) Sargan's  $\chi^2$  test performed well in detecting the misspecified factor loading. Interestingly, the test did not perform as well at detecting the misspecified time series components, only approaching the nominal rejection rate at the largest time series length,  $T=500$ . Given the effects of structural misspecification on DFM model parameters observed here we view these results as a promising first step in providing new methods for diagnosing model misspecification across all time series estimators.

## Appendix

### Appendix A.: Additional Monte Carlo Simulations

The data generating procedure, time series lengths, estimators, equation-specific MIIVs, and summary measures were identical across the two simulations.

#### A.1 Model Specification

In Simulation 2 we varied the data generating parameters contained in  $\Lambda$  and  $\Phi$ ,

$$\Phi = \begin{bmatrix} 0.7 & -0.2 \\ 0.0 & 0.5 \end{bmatrix}, \quad \Lambda = [\Lambda_0 \Lambda_1] = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0 \\ 0.6 & 0 & 0.5 & 0 \\ 0 & 1.0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0.6 & 0 & 0 \end{bmatrix}.$$

while keeping  $\Psi$  and  $\Xi$  consistent with Simulation 1. This data generating model allows us to examine consequences of model misspecification not permitted by Simulation 1, such as the estimation of a cross-regressive path,  $\phi_{21}$ , whose data generating value was zero. To assess the performance of each estimator we again considered three different use cases: (1) the correct measurement model in terms of free and fixed parameters and a latent variable model with all autoregressive and cross-regressive effects present, (2) a misspecified measurement model where a single lagged factor loading,  $\lambda_{33}$ , is incorrectly omitted, and the latent variable model is identical to (1), and (3) the case where the measurement model is correctly specified, but the cross-regressive path,  $\phi_{12}$  is incorrectly omitted from the model specification.

#### A.2 Results

As in Simulation 1 to make comparisons across the estimators and model specifications only those datasets on which all estimators converged with no improper solutions were used. Results were consistent across the two simulations with the majority of improper solutions occurring at the smallest time series length. This left 437 datasets at  $T=50$ , and 500 datasets at  $T=100, 250, 500$  for the following comparisons. For Simulation 2 the percentage of mean relative bias is presented in Table A1, the standard deviation of coefficient estimates in

Table A2, and the proportion of equations failing Sargan’s  $\chi^2$  test in Table A3. These results are generally consistent with those observed in Simulation 1. For this reason we do not discuss the individual measures exhaustively but highlight areas where the results differed.

First, Simulation 2 provides a test of the MIV-2SLS estimator when a cross-regressive path is erroneously included in the model specification. Specifically, in Models 1 and 2,  $\phi_{21}$  is estimated despite the population value being zero. Since it is not possible to calculate the percentage of relative bias in this case we give the range of mean coefficient estimates across all time series lengths for Models 1 and 2. Here the coefficient estimates were similar across estimators, ranging from 0.004 to 0.021 for MIV-2SLS, 0.002 to 0.026 for the pseudo-ML, and 0.003 to 0.029 for the Kalman Filter. No differences in the coefficient estimates of  $\phi_{21}$  were observed as a result of omitting  $\lambda_{33}$  from the model specification. For the pseudo-ML and KF estimators, the omission of  $\lambda_{33}$  (Model 2) had a less dramatic impact on coefficient estimates of  $\phi_{21}$ , while estimates of  $\lambda_{31}$  were generally more biased, when compared to the results from Simulation 1. The pattern of standard deviations (Table A2) observed in Simulation 2 was also similar to the pattern seen in Simulation 1. Results for Sargan’s test showed a slight increase in power to detect the misspecified measurement model equation and a decrease to detect the omitted cross-regressive path. The consistency of our findings across these two simulations designs provide additional confidence in the generalizability of our results.

**Table A1**

Mean Percentage of Relative Bias

Parameter	Estimator											
	MIV-2SLS				pseudo-ML				Kalman Filter			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
Correct Measurement Model, All Elements of $\Phi$ Estimated												
$\phi_{11}$	-15	-7	-2	-1	-11	-6	-2	-1	-13	-7	-2	-1
$\phi_{12}$	-17	-7	-0	-2	12	1	1	-1	-15	-11	-4	-2
$\phi_{22}$	-17	-9	-4	-2	-21	-10	-4	-2	-17	-8	-3	-2
$\lambda_{21}$	1	1	0	0	3	2	0	1	1	1	0	1
$\lambda_{31}$	1	-1	1	-0	3	2	1	0	0	0	0	0
$\lambda_{55}$	-6	-2	-2	-0	0	-1	-0	-0	1	0	0	0
$\lambda_{52}$	-2	1	-1	-0	5	2	0	0	3	L	0	0
$\lambda_{62}$	1	1	-0	-0	5	3	0	0	2	2	0	0
Percentage of Relative Bias from Model 1 Due to Omission of $\lambda_{33}$												
$\phi_{11}$	0	0	0	0	13	12	12	12	11	10	10	10
$\phi_{12}$	0	0	0	0	-31	-30	-31	-31	-23	-22	-22	-23



Parameter	Estimator											
	MIIV-2SLS				pseudo-ML				Kalman Filter			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
$\phi_{22}$	0	0	0	0	-1	-0	0	0	1	1	1	L
$\lambda_{21}$	0	0	0	0	L	L	1	1	-1	0	0	0
$\lambda_{31}$	53	59	61	63	71	72	74	75	65	67	69	69
$\lambda_{52}$	0	0	0	0	0	0	0	-0	1	1	1	0
$\lambda_{62}$	0	0	0	0	1	1	0	-0	1	1	1	0
Percentage of Bias from Model 1 Due to Omission of $\phi_{21}$ and $\phi_{12}$												
$\phi_{11}$	-6	-9	-10	-10	-10	-10	-10	-10	-6	-7	-9	-9
$\phi_{22}$	6	3	2	1	17	13	11	11	15	13	13	12
$\lambda_{21}$	0	0	0	0	-0	-0	-0	-0	0	0	0	-0
$\lambda_{31}$	0	0	0	0	2	1	L	1	2	2	2	2
$\lambda_{55}$	0	0	0	0	-2	-2	-2	-2	-1	-2	-2	-2
$\lambda_{52}$	0	0	0	0	-1	0	-0	-0	1	1	0	0
$\lambda_{62}$	0	0	0	0	0	0	-0	-0	1	1	0	0

**Table A2**

Standard Deviation of Coefficient Estimates

Parameter	Estimator											
	MIIV-2SLS				pseudo-ML				Kalman Filter			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
Correct Measurement Model, All Elements of $\Phi$ Estimated												
$\phi_{11}$	0.23	0.15	0.09	0.07	0.19	0.11	0.07	0.05	0.16	0.10	0.06	0.05
$\phi_{12}$	0.38	0.27	0.15	0.11	0.26	0.14	0.08	0.06	0.19	0.12	0.08	0.05
$\phi_{21}$	0.27	0.17	0.10	0.07	0.19	0.12	0.07	0.05	0.16	0.11	0.06	0.04
$\phi_{22}$	0.28	0.18	0.11	0.07	0.22	0.14	0.08	0.05	0.18	0.12	0.07	0.05
$\lambda_{21}$	0.17	0.12	0.07	0.05	0.17	0.12	0.07	0.05	0.16	0.11	0.07	0.05
$\lambda_{31}$	0.24	0.17	0.11	0.08	0.20	0.15	0.09	0.06	0.19	0.14	0.09	0.06
$\lambda_{33}$	0.25	0.18	0.12	0.08	0.19	0.13	0.08	0.06	0.19	0.13	0.08	0.06
$\lambda_{52}$	0.21	0.17	0.10	0.08	0.24	0.15	0.09	0.07	0.21	0.14	0.09	0.06
$\lambda_{62}$	0.19	0.13	0.08	0.06	0.21	0.13	0.08	0.05	0.17	0.12	0.07	0.05

**Table A3**  
Proportion of Equations Failing Sargan's  $\chi^2$  Test

Parameter	Model Specification											
	Model 1				Model 2				Model 3			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
$y_2$	0.03	0.06	0.06	0.06	0.03	0.06	0.06	0.06	0.03	0.06	0.06	0.06
$y_3^a$	0.06	0.06	0.07	0.06	0.17	0.43	0.86	1.00	0.06	0.06	0.07	0.06
$y_5$	0.05	0.05	0.04	0.06	0.05	0.05	0.04	0.06	0.05	0.05	0.04	0.06
$y_6$	0.04	0.06	0.05	0.05	0.04	0.06	0.05	0.05	0.04	0.06	0.05	0.05
$\eta_1^b$	0.05	0.03	0.04	0.06	0.05	0.03	0.04	0.06	0.05	0.06	0.15	0.37
$\eta_2^b$	0.04	0.05	0.03	0.05	0.04	0.05	0.03	0.05	0.05	0.04	0.04	0.05

<sup>a</sup>Equation is misspecified in Model 2.

<sup>b</sup>Equation is misspecified in Model 3.

## References

- Anderson TW (1963). The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika*, 28(1), 1–25. doi: 10.1007/bf02289543
- Bollen KA (1989). *Structural Equations with Latent Variables*. New York: Wiley-Interscience. doi: 10.1002/9781118619179
- Bollen KA (1995). Structural equation models that are nonlinear in latent variables: A least-squares estimator. *Sociological Methodology*, 25, 223–251. doi: 10.2307/271068
- Bollen KA (1996, 3). An alternative two stage least squares (2SLS) estimator for latent variable equations. *Psychometrika*, 61 (1), 109–121. doi: 10.1007/bf02296961
- Bollen KA (2001). Two-stage least squares and latent variable models: Simultaneous estimation and robustness to misspecifications In Cudeck R, Jöreskog KG, & Sörbom D (Eds.), *Structural equation modeling: Present and future : a festschrift in honor of karl jöreskog*. Scientific Software International.
- Bollen KA (2011). Evaluating effect, composite, and causal indicators in structural equation models. *MIS Quarterly*, 35(2), 359–372. doi: 10.2307/23044047
- Bollen KA (2012). Instrumental variables in sociology and the social sciences. *Annual Review of Sociology*, 38, 37–72. doi: 10.1146/annurev-soc-081309-150141
- Bollen KA, & Biesanz JC (2002). A note on a two-stage least squares estimator for higher-order factor analyses. *Sociological methods and research*, 30(4), 568–579. doi: 10.1177/0049124102030004004
- Bollen KA, Gates KM, & Fisher Z (2018, 5). Robustness Conditions for MIIV-2sls When the Latent Variable or Measurement Model is Structurally Misspecified. *Structural Equation Modeling: A Multidisciplinary Journal*, 0 (0), 1–12. Retrieved 2018–06-25, from 10.1080/10705511.2018.1456341 doi: 10.1080/10705511.2018.1456341
- Bollen KA, Kirby JB, Curran PJ, Paxton PM, & Chen F (2007, 8). Latent variable models under misspecification: Two-stage least squares (2sls) and maximum likelihood (ML) estimators. *Sociological Methods & Research*, 36(1), 48–86. doi: 10.1177/0049124107301947
- Bollen KA, Kolenikov S, & Bauldry S (2014, 1). Model-implied instrumental variable generalized method of moments (MIIV-GMM) estimators for latent variable models. *Psychometrika*, 79(1), 20–50. doi: 10.1007/s11336-013-9335-3 [PubMed: 24532165]

- Bollen KA, & Maydeu-Olivares A (2007, 6). A Polychoric Instrumental Variable (PIV) Estimator for Structural Equation Models with Categorical Variables. *Psychometrika*, 72(3), 309–326. doi: 10.1007/s11336-007-9006-3
- Bollen KA, & Paxton P (1998). Interactions of latent variables in structural equation models. *Structural Equation Modeling*, 5(3), 267–293. doi: 10.1080/10705519809540105
- Bowden RJ, & Turkington DA (1990). *Instrumental variables*. Cambridge University Press. doi: 10.1007/978-1-4899-4477-1\_5
- Box G, Jenkins G, & Reinsel G (2008). *Time series analysis: Forecasting and control* (4th Edition ed). Wiley. doi: 10.1111/j.1467-9892.2009.00643.x
- Browne MW (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37(1), 62–83. doi: 10.1111/j.2044-8317.1984.tb00789.x [PubMed: 6733054]
- Browne MW, & Nesselroade JR (2005, 4). Representing Psychological Processes with Dynamic Factor Models: Some Promising Uses and Extensions of Autoregressive Moving Average Time Series Models In Maydeu-Olivares A & McArdle JJ (Eds.), *Contemporary Psychometrics*. Mahwah, N.J: Psychology Press.
- Browne MW, & Zhang G (2005). Dyfa: Dynamic factor analysis of lagged correlation matrices [Computer software manual]. Retrieved from <http://quantrm2.psy.ohio-state.edu/browne/> (Version 2.03)
- Browne MW, & Zhang G (2007). Developments in the Factor Analysis of Individual Time Series In Cudeck R & MacCallum RC (Eds.), *Factor Analysis at 100: Historical Developments and Future Directions* (1edition ed.). Mahwah, N.J: Routledge.
- Cattell RB (1963). The structure of change by p-technique and incremental r-technique. In *Problems in measuring change*. The University of Wisconsin Press.
- Cattell RB, Cattell AKS, & Rhymer RM (1947, 12). P-technique demonstrated in determining psychophysiological source traits in a normal individual. *Psychometrika*, 12(4), 267–288. doi: 10.1007/bf02288941 [PubMed: 18921433]
- Chow S-M, Ho M.-h. R, Hamaker EL, & Dolan CV. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling*, 17(2), 303–332. doi: 10.1080/10705511003661553
- Chow S-M, & Zhang G (2013, 10). Nonlinear Regime-Switching State-Space (RSSS) models. *Psychometrika*, 78(4), 740–768. doi: 10.1007/s11336-013-9330-8 [PubMed: 24092487]
- Chow S-M, Zu J, Shifren K, & Zhang G (2011). Dynamic Factor Analysis Models With Time-Varying Parameters. *Multivariate Behavioral Research*, 46(2), 303–339. doi: 10.1080/00273171.2011.563697 [PubMed: 26741330]
- Cudeck R (1991). Noniterative factor analysis estimators, with algorithms for subset and instrumental variable selection. *Journal of Educational Statistics*, 16(1), 35–52. doi: 10.2307/1165098
- Dolado JJ (1990). Optimal instrumental variable estimator of the AR parameter of an ARMA(1,1). *Econometric Theory*, 6(1), 117–119. doi: 10.1017/s0266466600005016
- du Toit SHC, & Browne MW (2007, 6). Structural Equation Modeling of Multivariate Time Series. *Multivariate Behavioral Research*, 42(1), 67–101. doi: 10.1080/00273170701340953 [PubMed: 26821077]
- Fisher ZF, Bollen KA, Gates KM, & Rönkkö M (2017). *Miivsem: Model implied instrumental variable estimation of structural equation models* [Computer software manual]. (R package version 0.52)
- Granger CWJ, & Morris MJ (1976). Time series modelling and interpretation. *Journal of the Royal Statistical Society. Series A (General)*, 139(2), 246–257. doi: 10.2307/2345178
- Hall AR (2005). *Generalized Method of Moments*. Oxford University Press.
- Hamaker EL, Dolan CV, & Molenaar PCM (2002, 7). On the Nature of SEM Estimates of ARMA Parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(3), 347–368. doi: 10.1207/s15328007sem0903\_3
- Hamaker EL, Dolan CV, & Molenaar PCM (2003, 7). ARMA-Based SEM When the Number of Time Points T Exceeds the Number of Cases N: Raw Data Maximum Likelihood. *Structural Equation Modeling: A Multidisciplinary Journal*, 10(3), 352–379. doi: 10.1207/s15328007sem1003\_2

- Hansen LP (1982). Large sample properties of generalized method of moments estimators., 50(4), 1029–54. doi: 10.2307/1912775
- Harvey AC (1981). *The econometric analysis of time series*. Wiley. doi: 10.2307/2290620
- Hayashi F (2011). *Econometrics*. Princeton University Press.
- Holmes EE, Ward EJ, & Wills K (2012). Marss: Multivariate autoregressive state-space models for analyzing time-series data. *The R Journal*, 4 (1), 30.
- Holtzman W (1963). Statistical models for the study of change in the single case. In *Problems in measuring change*. The University of Wisconsin Press.
- Jin S, Luo H, & Yang-Wallentin F (2016). A simulation study of polychoric instrumental variable estimation in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 680–694. doi: 10.1080/10705511.2016.1189334
- Justiano A (2004). Estimation and model selection in dynamic factor analysis. (Unpublished doctoral dissertation). Princeton.
- Kalman R (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, 82, 35–45. doi: 10.1115/1.3662552
- Kirby JB, & Bollen KA (2009). Using instrumental variable (IV) tests to evaluate model specification in latent variable structural equation models. *Sociological Methodology*, 39 (1), 327–355. [PubMed: 20419054]
- Lebo MA, & Nesselroade JR (1978, 6). Intraindividual differences dimensions of mood change during pregnancy identified in five P-technique factor analyses. *Journal of Research in Personality*, 12(2), 205–224. doi: 10.1016/0092-6566(78)90098-3
- Lee T (2010). An em algorithm for maximum likelihood estimation of process factor analysis models (Unpublished doctoral dissertation). University of North Carolina at Chapel Hill.
- Lütkepohl H (2007). *New introduction to multiple time series analysis*. Springer.
- Mikhail WM (1972, 9). The Bias of the Two-Stage Least Squares Estimator. *Journal of the American Statistical Association*, 67(339), 625–627. doi: 10.1080/01621459.1972.10481263
- Molenaar PCM (1994, 8). Dynamic latent variable models in developmental psychology In Eye A. v.& Clogg CC (Eds.), *Latent Variables Analysis: Applications for Developmental Research* (pp. 155–180). Thousand Oaks, Calif: SAGE Publications, Inc.
- Molenaar PCM (2003). State space techniques in structural equation modeling. Molenaar, P. C. M. (2017). Equivalent dynamic models. *Multivariate Behavioral Research*, 52(2), 242–258. doi: 10.1080/00273171.2016.1277681
- Molenaar PCM, Gooijer JGD, & Schmitz B (1992, 9). Dynamic factor analysis of nonstationary multivariate time series. *Psychometrika*, 57(3), 333–349. doi: 10.1007/bf02295422
- Molenaar PCM, & Nesselroade JR (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(3), 313–342. doi: 10.1207/s15327906mbr3303\_1 [PubMed: 26782717]
- Nestler S (2013). A monte carlo study comparing PIV, ULS and DWLS in the estimation of dichotomous confirmatory factor analysis. *British Journal of Mathematical and Statistical Psychology*, 66(1), 127–143. doi: 10.1111/j.2044-8317.2012.02044.x [PubMed: 22524532]
- Nestler S (2014a, 5). How the 2sls/IV estimator can handle equality constraints in structural equation models: A system-of-equations approach. *British Journal of Mathematical and Statistical Psychology*, 67(2), 353–369. doi: 10.1111/bmsp.12023 [PubMed: 2403324]
- Nestler S (2014b). Using Instrumental Variables to Estimate the Parameters in Unconditional and Conditional Second-Order Latent Growth Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 0(0), 1–13. doi: 10.1080/10705511.2014.934948
- Nestler S (2015). A specification error test that uses instrumental variables to detect latent quadratic and latent interaction effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(4), 542–551. doi: 10.1080/10705511.2014.994744
- Nimark KP (2015, 2). A low dimensional Kalman filter for systems with lagged states in the measurement equation. *Economics Letters*, 127(Supplement C), 10–13. doi: 10.1016/j.econlet.2014.12.016

- Ram N, Brose A, & Molenaar PCM (2013, 3). Dynamic Factor Analysis: Modeling Person-Specific Process. *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*. doi: 10.1093/oxfordhb/9780199934898.013.0021
- Rosseel Y (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. Retrieved from <http://www.jstatsoft.org/v48/i02/> doi: 10.18637/jss.v048.i02
- Sargan JD (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, 26(3), 393–415. doi: 10.2307/1907619
- Serena N, & Jushan B (2009). Selecting Instrumental Variables in a Data Rich Environment. *Journal of Time Series Econometrics*, 1 (1), 1–34. doi: 10.2202/1941-1928.1014
- Singer H (2010, 9). SEM Modeling with Singular Moment Matrices Part I: ML-Estimation of Time Series. *The Journal of Mathematical Sociology*, 34 (4), 301–320. doi: 10.1080/0022250x.2010.509524
- Stock JH, Wright JH, & Yogo M (2002). A survey of weak instruments and weak identification in generalized method of moments., 20 (4), 518–529. doi: 10.1198/073500102288618658
- Stoica P, Friedlander B, & Soderstrom T (1987). Optimal instrumental variable multistep algorithms for estimation of the AR parameters of an ARMA process. *International Journal of Control*, 45(6), 2083–2107. doi: 10.1080/00207178708933869
- Stoica P, Soderstrom T, & Friedlander B (1985). Optimal instrumental variable estimates of the AR parameters of an ARMA process. *IEEE Transactions on Automatic Control*, 30(11), 1066–1074. doi: 10.1109/tac.1985.1103839
- van Buuren S (1997, 6). Fitting arma time series by structural equation models. *Psychometrika*, 62(2), 215–236. doi: 10.1007/bf02295276
- Voelkle MC, Brose A, Schmiedek F, & Lindenberger U (2014, 5). Toward a Unified Framework for the Study of Between-Person and Within-Person Structures: Building a Bridge Between Two Research Paradigms. *Multivariate Behavioral Research*, 49(3), 193–213. doi: 10.1080/00273171.2014.889593 [PubMed: 26735189]
- Voelkle MC, & Oud JHL (2013). Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes., 66 (1), 103–126. doi: 10.1111/j.2044-8317.2012.02043.x
- Voelkle MC, Oud JHL, von Oertzen T, & Lindenberger U (2012, 7). Maximum Likelihood Dynamic Factor Modeling for Arbitrary N and T Using SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(3), 329–350. doi: 10.1080/10705511.2012.687656
- Woolridge JM (2010). *Econometric analysis of cross section and panel data (Second Edition ed)*. The MIT Press.
- Zhang G, & Browne MW (2010). Bootstrap standard error estimates in dynamic factor analysis., 45(3). doi: 10.1080/00273171.2010.483375
- Zhang G, Browne MW, Ong AD, & Chow SM (2014). Analytic standard errors for exploratory process factor analysis., 79(3), 444–469. doi: 10.1007/s11336-013-9365-x
- Zhang Z, Hamaker EL, & Nesselroade JR (2008, 7). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling: A Multidisciplinary Journal*, 15(3), 377–402. doi: 10.1080/10705510802154281
- Zhang Z, & Nesselroade JR (2007, 12). Bayesian Estimation of Categorical Dynamic Factor Models. *Multivariate Behavioral Research*, 42 (4), 729–756. doi: 10.1080/00273170701715998

**Table 1**

Mean Percentage of Relative Bias

Parameter	Estimator											
	MIIV-2SLS				pseudo-ML				Kalman Filter			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
Correctly Specified Model												
$\phi_{11}$	-15	-6	-3	-2	-14	-6	-3	-1	-9	-6	-2	-1
$\phi_{12}$	-6	-4	-3	0	-10	-2	2	2	-5	-5	-3	0
$\phi_{21}$	3	-1	-2	1	5	2	-0	1	-3	-3	1	-0
$\phi_{22}$	-11	-7	-4	-1	-14	-7	-4	-1	-13	-5	-2	-2
$\lambda_{21}$	-0	-0	-0	0	1	0	0	0	-0	-0	0	0
$\lambda_{31}$	-1	0	0	0	2	1	-0	0	-1	1	0	1
$\lambda_{33}$	2	-1	-1	0	4	0	1	0	0	-1	1	1
$\lambda_{52}$	1	1	-0	-0	3	1	0	-0	-0	0	0	0
$\lambda_{62}$	0	1	-1	0	2	2	-0	0	-1	1	0	0
Percentage of Relative Bias from Correctly Specified Model Due to Omission of $\lambda_{33}$												
$\phi_{11}$	0	0	0	0	13	12	11	11	11	10	10	10
$\phi_{12}$	0	0	0	0	-42	-44	-45	-46	-33	-34	-35	-36
$\phi_{21}$	0	0	0	0	4	2	1	1	-0	-3	-2	-3
$\phi_{22}$	0	0	0	0	-3	-3	-3	-3	-1	-1	-1	-1
$\lambda_{21}$	0	0	0	0	0	-0	-0	0	-2	-1	-1	-1
$\lambda_{31}$	35	37	39	40	44	44	45	45	39	40	42	42
$\lambda_{52}$	0	0	0	0	0	0	0	0	0	0	0	0
$\lambda_{62}$	0	0	0	0	-0	0	-0	-0	0	0	0	0
Percentage of Relative Bias from Correctly Specified Model Due to Omission of $\phi_{12}$ and $\phi_{21}$												
$\phi_{11}$	-3	-2	-2	-2	9	9	9	9	10	10	10	10
$\phi_{22}$	-5	-6	-4	-4	9	8	8	9	13	10	10	10
$\lambda_{21}$	0	0	0	0	0	0	0	0	1	0	0	-0
$\lambda_{31}$	0	0	0	0	-0	0	0	0	1	1	1	1
$\lambda_{33}$	0	0	0	0	-1	-1	-1	-1	-2	-3	-3	-4
$\lambda_{52}$	0	0	0	0	0	0	0	0	1	1	0	0
$\lambda_{62}$	0	0	0	0	0	0	0	0	1	1	0	0



**Table 2**

Standard Deviation of Coefficient Estimates

Parameter	Estimator											
	MIIV-2SLS				pseudo-ML				Kalman Filter			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
Correctly Specified Model												
$\phi_{11}$	0.20	0.12	0.07	0.05	0.16	0.10	0.06	0.04	0.12	0.09	0.05	0.04
$\phi_{12}$	0.29	0.16	0.09	0.07	0.21	0.12	0.07	0.05	0.16	0.11	0.07	0.05
$\phi_{21}$	0.21	0.13	0.08	0.05	0.17	0.10	0.06	0.04	0.13	0.09	0.05	0.04
$\phi_{22}$	0.21	0.14	0.09	0.06	0.17	0.11	0.07	0.05	0.14	0.10	0.07	0.05
$\lambda_{21}$	0.18	0.11	0.07	0.05	0.17	0.11	0.07	0.04	0.15	0.10	0.06	0.04
$\lambda_{31}$	0.29	0.20	0.13	0.08	0.24	0.16	0.09	0.07	0.21	0.15	0.09	0.06
$\lambda_{33}$	0.28	0.20	0.12	0.08	0.21	0.14	0.08	0.06	0.19	0.13	0.08	0.06
$\lambda_{52}$	0.22	0.15	0.09	0.06	0.23	0.15	0.08	0.06	0.18	0.14	0.08	0.06
$\lambda_{62}$	0.21	0.14	0.08	0.06	0.22	0.14	0.08	0.06	0.18	0.13	0.08	0.06

**Table 3**

Proportion of Equations Failing Saragan's  $\chi^2$  Test

Dependent Variable	Model Specification											
	Model 1				Model 2				Model 3			
	Correctly Specified				$\lambda_{33}$ Incorrectly Omitted				$\phi_{12}$ and $\phi_{21}$ Incorrectly Omitted			
	Time Series Length				Time Series Length				Time Series Length			
	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500	T = 50	T = 100	T = 250	T = 500
$y_2$	0.03	0.05	0.06	0.06	0.03	0.05	0.06	0.06	0.03	0.05	0.06	0.06
$y_3^a$	0.08	0.06	0.06	0.05	0.20	0.32	0.74	0.98	0.08	0.06	0.06	0.05
$y_5$	0.05	0.07	0.06	0.05	0.05	0.07	0.06	0.05	0.05	0.07	0.06	0.05
$y_6$	0.05	0.05	0.04	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.04	0.05
$\eta_1^b$	0.05	0.05	0.05	0.08	0.05	0.05	0.05	0.08	0.05	0.11	0.24	0.57
$\eta_2^b$	0.05	0.04	0.06	0.05	0.05	0.04	0.06	0.05	0.06	0.12	0.39	0.88

<sup>a</sup>Equation is misspecified in Model 2.

<sup>b</sup>Equation is misspecified in Model 3.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4**Coefficient Estimates and Sargan's  $\chi^2$  Test P-Values for Subject 1

Equation	Parameter	Model					
		1		2		3	
		Coefficient	Sargan $\chi^2$ P-Value	Coefficient	Sargan $\chi^2$ P-Value	Coefficient	Sargan $\chi^2$ P-Value
wellbeing	$\phi_{11}$	0.36 *	0.98	0.36 *	0.98	0.36 *	0.98
	$\phi_{12}$	0.11		0.11		0.11	
fatigue	$\phi_{22}$	0.32	0.47	0.32	0.47	0.32	0.47
	$\phi_{21}$	0.07		0.07		0.07	
happy	$\lambda_{21}^{(0)}$	0.85 *	0.82	0.85 *	0.82	0.85 *	0.82
content	$\lambda_{31}^{(0)}$	0.65 *	0.34	0.65 *	0.34	0.65 *	0.34
sluggish	$\lambda_{52}^{(0)}$	0.78 *	0.00	0.76 *	0.00	1.13 *	0.12
	$\lambda_{52}^{(0)}$	—	—	0.14		—	—
weary	$\lambda_{62}^{(0)}$	0.87 *	0.27	0.87 *	0.27	0.99 *	0.41

Note: The Sargan  $\chi^2$  P-Values are equation specific. A — indicates a coefficient was not estimated in the corresponding model.

\* An indicates a 95% CI for the coefficient estimates constructed using a moving block bootstrap (1,000 replications) did not contain zero.