Estimating a Dynamic Factor Model in EViews Using the Kalman Filter and Smoother

Martin Solberger[1,2] · Erik Spånberg[2,3]

Abstract

Dynamic factor models have become very popular for analyzing high-dimensional time series, and are now standard tools in, for instance, business cycle analysis and forecasting. Despite their popularity, most statistical software do not provide these models within standard packages. We briefly review the literature and show how to estimate a dynamic factor model in EViews. A subroutine that estimates the model is provided. In a simulation study, the precision of the estimated factors are evaluated, and in an empirical example, the usefulness of the model is illustrated.

Keywords Dynamic factor model · State space · Kalman filter · EViews

1 Introduction

Dynamic factor models are used in data-rich environments. The basic idea is to separate a possibly large number of observable time series into two independent and unobservable, yet estimable, components: a 'common component' that captures the main bulk of co-movement between the observable series, and an 'idiosyncratic component' that captures any remaining individual movement. The common component

✉ Martin Solberger
martin.solberger@statistics.uu.se

Erik Spånberg
erik.spanberg@stat.su.se

1 Department of Statistics, Uppsala University, Kyrkogårdsgatan 10, PO Box 513, 75120 Uppsala, Sweden

2 Ministry of Finance, Stockholm, Sweden

3 Department of Statistics, Stockholm University, Universitetsvägen 10b, 10691 Stockholm, Sweden

is assumed to be driven by a few common factors, thereby reducing the dimension of the system.

In Economics, dynamic factor models are motivated by theory, which predicts that macroeconomic shocks should be pervasive and affect most variables within an economic system. They have therefore become popular among macroeconometricians; see, e.g., Breitung and Eickmeier (2006), for an overview. Areas of economic analysis using dynamic factor models include, for example, yield curve modeling (e.g., Diebold and Li 2006; Diebold et al. 2006), financial risk-return analysis (Ludvigson and Ng 2007), monetary policy analysis (e.g., Bernanke et al. 2005; Boivin et al. 2009), business cycle analysis (e.g., Forni and Reichlin 1998; Eickmeier 2007; Ritschl et al. 2016), forecasting (e.g., Stock and Watson 2002a, b) and nowcasting the state of the economy, that is, forecasting of the very recent past, the present, or the very near future of indicators for economic activity, such as the gross domestic product (GDP) (see, e.g., Banbura et al. 2012, and references therein). Information of the economic activity is of great importance for decision makers in, for instance, governments, central banks and financial markets. However, first official estimates of GDP are published with a significant delay, usually about 6–8 weeks after the reference quarter, which makes nowcasting very useful. Despite the attractiveness of dynamic factor models for macroeconomists, statistical or econometric software do not in general provide these models within standard packages. In this paper, we illustrate how to, by means of programming, set up the popular two-step estimator of Doz et al. (2011) in EViews (IHS Global Inc. 2015a, b, c, d), a software specialized in time series analysis that is broadly used by economists, econometricians, and statisticians.

The parameters of dynamic factor models can be estimated by the method of principal components. This method is easy to compute, and is consistent under quite general assumptions as long as both the cross-section and time dimension grow large. It suffers, however, from a large drawback: the data set must be balanced, where the start and end points of the sample are the same across all observable time series. In practice, data are often released at different dates. A popular approach is therefore to cast the dynamic factor model in a state space representation and then estimate it using the Kalman filter, which allows unbalanced data sets and offers the possibility to smooth missing values. The state space representation contains a signal equation, which links observed series to latent *states*, and a state equation, which describes how the states evolve over time. Under the assumption of Gaussian noise, the Kalman filter and smoother provide mean-square optimal projections for both the signal and state equations. The method we set up in this paper is a two-step procedure, in which parameters are first estimated by principal components, and then, given these estimates, the factors are re-estimated as latent states by the Kalman smoother.

The rest of the paper is organized as follows. Section 2 outlines the notion and conventional estimation of dynamic factor models. Section 3 derives a state space solution. Section 4 describes the estimator considered in this paper. The estimator is evaluated in a simulation study in Sect. 5, and applied in an empirical example in Sect. 6. Section 7 concludes. A subroutine containing the estimator and programs that replicate our results are provided as supplementary material.

2 A Dynamic Factor Model

Let $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \ldots, x_{N,t})'$ be a vector of N time series, each of which is a real-valued stochastic process $\{x_{i,t}, t \in \mathbb{Z}\}$. Suppose we observe a finite realization of $\mathbf{x}_t$ over some time points $t = 1, 2, \ldots, T$, and let the empirical information available at time t be condensed into the information set $\mathcal{F}_t = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t\}$. A dynamic factor model is usually specified such that each observable $x_{i,t}$ ($i = 1, 2, \ldots, N$) is the sum of two independent and unobservable components: a common component $\chi_{i,t}$, which is driven by a small number of factors that are common to all individuals, and a remaining idiosyncratic (individual-specific) component $\epsilon_{i,t}$ (see, e.g., Bai and Ng [2008]). In panel notation, the model is

$$
\begin{aligned}
x_{i,t} &= \chi_{i,t} + \epsilon_{i,t}, \quad (i = 1, 2, \ldots, N; t = 1, 2, \ldots, T), \\
\chi_{i,t} &= \boldsymbol{v}_i(L)' \mathbf{z}_t,
\end{aligned}
\tag{1}
$$

where $\boldsymbol{v}_i(L) = \boldsymbol{v}_{i,0} + \boldsymbol{v}_{i,1}L + \cdots + \boldsymbol{v}_{i,\ell}L^\ell$ ($\ell < \infty$) is a vector lag-polynomial of constants loading onto a vector of $\mathcal{K}$ unobservable common factors, $\mathbf{z}_t = (z_{1,t}, z_{2,t}, \ldots, z_{\mathcal{K},t})'$.[1] Thus, only the left-hand side of (1) is observed; the right-hand side is unobserved. If the dimension of $\mathbf{z}_t$ is finite ($\mathcal{K} < \infty$), then there exists for every i an $\mathcal{R} \times 1$ vector ($\mathcal{R} \geq \mathcal{K}$) of constants $\boldsymbol{\lambda}_i = (\lambda_{i,1}, \lambda_{i,2}, \ldots, \lambda_{i,\mathcal{R}})'$, such that $\boldsymbol{v}_i(L)' = \boldsymbol{\lambda}_i' \mathbf{C}(L)$, where $\mathbf{C}(L)$ is an $\mathcal{R} \times \mathcal{K}$ matrix lag-polynomial, $\mathbf{C}(L) = \sum_{m=0}^{\infty} \mathbf{C}_m L^m$, that is absolutely summable, $\sum_{m=0}^{\infty} ||\mathbf{C}_m|| < \infty$ (see Forni et al. [2009]).[2] Thus, letting $\mathbf{f}_t = (f_{1,t}, f_{2,t}, \ldots, f_{\mathcal{R},t})' = \mathbf{C}(L)\mathbf{z}_t$, the dynamic factor model can be cast in the *static representation*

$$
\begin{aligned}
x_{i,t} &= c_{i,t} + \epsilon_{i,t}, \\
c_{i,t} &= \boldsymbol{\lambda}_i' \mathbf{f}_t,
\end{aligned}
\tag{2}
$$

which, equivalently, can be written in vector notation as

$$
\begin{aligned}
\mathbf{x}_t &= \mathbf{c}_t + \boldsymbol{\epsilon}_t, \\
\mathbf{c}_t &= \boldsymbol{\Lambda} \mathbf{f}_t,
\end{aligned}
\tag{3}
$$

where $\mathbf{c}_t = (c_{1,t}, c_{2,t}, \ldots, c_{N,t})'$, $\boldsymbol{\epsilon}_t = (\epsilon_{1,t}, \epsilon_{2,t}, \ldots, \epsilon_{N,t})'$ and $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2', \ldots, \boldsymbol{\lambda}_N')'$. The common factors in $\mathbf{z}_t$ are often referred to as *dynamic factors*, while the common factors in $\mathbf{f}_t$ are referred to as *static factors*. The number of static factors, $\mathcal{R}$, cannot be smaller than the number of dynamic factors, and is typically much smaller than the number of cross-sectional individuals, $\mathcal{K} \leq \mathcal{R} \ll N$. As with $\chi_{i,t}$ in the dynamic representation (1), we refer to the scalar process c_i in (2), or the multivariate process $\mathbf{c}_t$ in (3), as the *common component*.

---

[1] The transpose of $\boldsymbol{v}_i(L)$ is defined as $\boldsymbol{v}_i(L)' = \boldsymbol{v}_{i,0}' + \boldsymbol{v}_{i,1}'L + \cdots + \boldsymbol{v}_{i,\ell}'L^\ell$, where $\boldsymbol{v}_{i,q}$ ($q = 1, 2, \ldots, \ell$) are $\mathcal{K} \times 1$ vectors, so that $\chi_{i,t} = \boldsymbol{v}_{i,0}'\mathbf{z}_t + \boldsymbol{v}_{i,1}'\mathbf{z}_{t-1} + \cdots + \boldsymbol{v}_{i,\ell}'\mathbf{z}_{t-\ell}$.

[2] Here, $|| \cdot ||$ denotes the Frobenius norm, $||\mathbf{C}_m|| = \sqrt{\text{tr}(\mathbf{C}_m \mathbf{C}_m')}$.

In general, we suppose that $x_{i,t}$ is a weakly stationary process with mean zero that has finite second-order moments, $E(x_{i,t}) = 0$, $E(x_{i,t}x_{i,t-s}) < \infty$ $(s \in \mathbb{Z})$. To uphold this, something close to the following is usually assumed:

A1 (common factors) The q-variate process $\mathbf{z}_t$ is independent and identically distributed (i.i.d.) over both the cross-section and time dimension with zero mean: $E(\mathbf{z}_t) = \mathbf{0}$, $E(\mathbf{z}_t\mathbf{z}_s') = \mathbf{0}$ for all $t \neq s$ and $E(\mathbf{z}_t\mathbf{z}_t') = \text{diag}(\omega_1^2, \omega_2^2, \ldots, \omega_{\mathcal{R}}^2)$ is a diagonal matrix with main diagonal entries $\omega_1^2, \omega_2^2, \ldots, \omega_{\mathcal{R}}^2$, where $\omega_j^2 < \infty$ for all j.

A2 (idiosyncratic components) The process $\epsilon_{i,t}$ admits a Wold representation $\epsilon_{i,t} = \theta_i(L)u_{i,t} = \sum_{m=0}^{\infty} \theta_{i,m}u_{i,t-m}$, where $\sum_{m=0}^{\infty} |\theta_{i,m}| < \infty$ and $u_{i,t}$ is i.i.d. white noise with limited cross-sectional dependence: $E(u_{i,t}) = 0$, $E(u_{i,t}u_{j,s}) = 0$ for all $t \neq s$, and $E(u_{i,t}u_{j,t}) = \tau_{i,j}$, with $\sum_{i=0}^{N} |\tau_{i,j}| < \mathcal{J}$, where $\mathcal{J}$ is some positive number that does not depend on N or T.

A3 (independence) The idiosyncratic errors $\mathbf{u}_t = (u_{i,t}, u_{2,t}, \ldots, u_{N,t})'$ and common shocks $\mathbf{z}_t$ are mutually independent groups: $E(\mathbf{u}_t\mathbf{z}_s') = \mathbf{0}$ for all t, s.

From Assumption A1, the static factors are stationary and variance-ergodic processes admitting a Wold representation $\mathbf{f}_t = \mathbf{C}(L)\mathbf{z}_t = \sum_{m=0}^{\infty} \mathbf{C}_m\mathbf{z}_{t-m}$. Assuming invertibility, the static factors may therefore follow some stationary VAR(p) process,

$$\mathbf{B}(L)\mathbf{f}_t = \mathbf{z}_t, \tag{4}$$

where $\mathbf{B}(L) = \mathbf{I}_{\mathcal{R}} - \mathbf{B}_1L - \mathbf{B}_2L^2 - \cdots - \mathbf{B}_pL^p = \mathbf{C}(L)^{-1}$, and $\mathbf{I}_{\mathcal{R}}$ denotes the $\mathcal{R} \times \mathcal{R}$ identity matrix.

We have assumed here, for the sake of argument, that the dynamic factors $\mathbf{z}_t$ (sometimes referred to as the *primitive shocks*) enters as errors in the static factor VAR process (4). This is unnecessarily strict. To be more precise, by suitably defining $\mathbf{f}_t$, the dynamic factor model (1) can in general be cast in the static representation (2), where $\mathbf{f}_t$ follows a VAR process which exact order depends on the specific dynamics of $\mathbf{z}_t$. The number of static factors is always $\mathcal{R} = \mathcal{K}(\ell + 1)$, where $\mathcal{K}$ is the number of dynamic factors and ℓ is the order of the vector lag-polynomial $\boldsymbol{\nu}_i(L)$ in the common component in (1); see Bai and Ng (2007). In practice, the static representation is typically stated without reference to a more general dynamic factor model. Additionally, it is often assumed that the static common factors follow a stationary VAR process. Assumption A1 is innocuous, as we may assume that $\mathbf{z}_t$ is an orthonormal error of the static factor VAR process, and that this, in general, relates to some dynamic factor model (1).

Now, from Assumptions A1–A3, the autocovariance of the panel data is

$$\boldsymbol{\Gamma}_x(h) = E(\mathbf{x}_t\mathbf{x}_{t-h}') = \boldsymbol{\Lambda}\boldsymbol{\Gamma}_f(h)\boldsymbol{\Lambda}' + \boldsymbol{\Gamma}_\epsilon(h),$$

where $\boldsymbol{\Gamma}_f(h) = E(\mathbf{f}_t\mathbf{f}_{t-h}')$ and $\boldsymbol{\Gamma}_\epsilon(h) = E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_{t-h}')$. By denoting $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}_x(0)$, $\boldsymbol{\Upsilon} = \boldsymbol{\Gamma}_f(0)$ and $\boldsymbol{\Psi} = \boldsymbol{\Gamma}_\epsilon(0)$, we can write the contemporary covariance matrix of $\mathbf{x}_t$ as $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Upsilon}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$. If the largest eigenvalue of $\boldsymbol{\Psi}$ is bounded as $N \to \infty$, then we have an *approximate* factor model as defined by Chamberlain and Rothschild (1983). Approximate factor models have become very popular within, for instance, panel

data econometrics, because they allow for a cross-sectional dependence among $\boldsymbol{\epsilon}_t$, whilst letting the factor structure be identified. If $\boldsymbol{\Psi}$ is diagonal, then we have the *exact* (or *strict*) factor model. Since the diagonal elements of $\boldsymbol{\Psi}$ are real and finite, the exact factor model is nested within the approximate factor model as a special case. The largest eigenvalue of $\boldsymbol{\Psi}$ is, in fact, smaller than $\max_j \sum_{i=1}^{N} |E(\epsilon_{i,t}\epsilon_{j,t})|$ (see Bai 2003), which is bounded with respect to both N and T by Assumption A2.[3] Hence, by assumption, we have an approximate factor model.

On top of this, some technical requirements are usually placed on the factor loadings, such that the common component is pervasive as the number of cross-sectional individuals increase. In essence, as $N \rightarrow \infty$, all eigenvalues of $N^{-1}\boldsymbol{\Lambda}'\boldsymbol{\Lambda}$ should be positive and finite, implying that the $\mathcal{R}$ largest eigenvalues of $\boldsymbol{\Sigma}$ are unbounded asymptotically.[4] Because the largest eigenvalue of $\boldsymbol{\Psi}$ is asymptotically bounded, the covariance decomposition in $\boldsymbol{\Sigma}$ is asymptotically identifiable.

Note finally that, without imposing restrictions, the factors and factor loadings are only identified up to pre-multiplication with an arbitrary $\mathcal{R} \times \mathcal{R}$ full rank matrix $\mathbf{M}$. That is, (3) is observationally equivalent to $\mathbf{x}_t = \check{\boldsymbol{\Lambda}}\check{\mathbf{f}}_t + \boldsymbol{\epsilon}_t$ where $\check{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}\mathbf{M}$ and $\check{\mathbf{f}}_t = \mathbf{M}^{-1}\mathbf{f}_t$, since, in terms of variances, $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Upsilon}\boldsymbol{\Lambda}' + \boldsymbol{\Psi} = \check{\boldsymbol{\Lambda}}Var(\check{\mathbf{f}}_t)\check{\boldsymbol{\Lambda}}' + \boldsymbol{\Psi}$. Due to this rotational indeterminacy in the factor space, we may without loss of generality impose the normalization $\boldsymbol{\Upsilon} = \mathbf{I}_{\mathcal{R}}$, restricting $\mathbf{M}$ to be an orthogonal matrix. This implies that only the spaces spanned by, respectively, the columns of $\boldsymbol{\Lambda}$ and those of $\mathbf{f}_t$, are identifiable from the contemporary covariance $\boldsymbol{\Sigma}$. In general, this identification requires a large N. In many cases, estimating the space spanned by the factors is as good as estimating the factors themselves. For instance, in forecasting under squared error loss, the object of interest is the conditional mean, which is unaffected by rotational multiplication. If, however, the actual factors or the coefficients associated with the factors are the parameters of interest, then we need to impose some identifying, yet arbitrary, restrictions such that the factors and factor loadings are exactly identified; see Bai and Ng (2013) and Bai and Wang (2014).

2.1 Estimation

Estimation of dynamic factor models concern foremost the common component; the idiosyncratic component is generally considered residual. The common component of the dynamic factor model (1) may be consistently estimated in the frequency domain by spectral analysis (see, e.g., Forni et al. 2004). The main benefit of the static represen-

---

[3] From Assumption A2, $E(\epsilon_{i,t}\epsilon_{j,t}) = \sum_{m=0}^{\infty}\theta_{i,m}E(u_{j,t-m}u_{j,t-m}) = \sum_{m=0}^{\infty}\theta_{i,m}\tau_{i,j} = \tau_{i,j}\sum_{m=0}^{\infty}\theta_{i,m}$. Thus, $|E(\epsilon_{i,t}\epsilon_{j,t})| = |\tau_{i,j}||\sum_{m=0}^{\infty}\theta_{i,m}|$, and so $\max_j \sum_{i=1}^{N}|E(\epsilon_{i,t}\epsilon_{j,t})| = |\sum_{m=0}^{\infty}\theta_{i,m}|\max_j \sum_{i=1}^{N}|\tau_{i,j}| < \infty$.

[4] To be more precise, Doz et al. (2011) assume that, as $N \rightarrow \infty$, the infimum of $\varphi_{min}(\boldsymbol{\Lambda}'\boldsymbol{\Lambda})/N$ exists and the supremum of $\varphi_{max}(\boldsymbol{\Lambda}'\boldsymbol{\Lambda})/N$ is finite, where $\varphi_{min}(\cdot)$ and $\varphi_{max}(\cdot)$ denote the minimum and maximum eigenvalue, respectively. By construction, the $\mathcal{R} \times \mathcal{R}$ matrix $\boldsymbol{\Lambda}'\boldsymbol{\Lambda}$ is positive definite, whereas the $N \times N$ matrix $\boldsymbol{\Lambda}\boldsymbol{\Lambda}'$ is positive semi-definite. It is easily established (see, e.g., Zhou and Solberger 2017, Lemma A1) that the (positive) ordered eigenvalues of $\boldsymbol{\Lambda}'\boldsymbol{\Lambda}$, $\varphi_1(\boldsymbol{\Lambda}'\boldsymbol{\Lambda}) \geq \varphi_2(\boldsymbol{\Lambda}'\boldsymbol{\Lambda}) \geq \cdots \geq \varphi_{\mathcal{R}}(\boldsymbol{\Lambda}'\boldsymbol{\Lambda})$, correspond exactly to the non-zero ordered eigenvalues of $\boldsymbol{\Lambda}\boldsymbol{\Lambda}'$. That is, for $i = 1, 2, \ldots, \mathcal{R}$, $\varphi_i(\boldsymbol{\Lambda}\boldsymbol{\Lambda}') = \varphi_i(\boldsymbol{\Lambda}'\boldsymbol{\Lambda})$, while for $i = \mathcal{R}+1, \mathcal{R}+2, \ldots, N$, $\varphi_i(\boldsymbol{\Lambda}\boldsymbol{\Lambda}') = 0$. Because the eigenvalues of $N^{-1}\boldsymbol{\Lambda}'\boldsymbol{\Lambda}$ are non-zero, the eigenvalues of $\boldsymbol{\Lambda}\boldsymbol{\Lambda}'$ are unbounded as $N \rightarrow \infty$.

tation (3), however, is that for an approximate factor model, the common component may be consistently estimated in the time domain, which computational methods are generally much easier to accomplish. Because the factor model is a panel (i.e., $x_{i,t}$ is doubly indexed over dimensions N and T), asymptotic theory has been developed as both dimensions N and T tend to infinity, which requires some special attention. Conceptually, estimation theory for panels can be derived in three ways: sequentially, diagonally, and jointly (see, e.g., Phillips and Moon 1999; Bai 2003). The methods presented in this paper all concern the latter limit. This limit is the most general, and its existence implies the existence of the other two limits.

From Assumption A2, each idiosyncratic component is a stationary and variance-ergodic process that, assuming invertibility, may be stated as a finite-order autoregressive process, $\phi_i(L)\epsilon_{i,t} = u_{i,t}$, where $\phi_i(L) = \theta_i(L)^{-1}$, so that the idiosyncratic process in (3) may be written as $\boldsymbol{\Phi}(L)\boldsymbol{\epsilon}_t = \mathbf{u}_t$, where $\mathbf{u}_t = (u_{1,t}, u_{2,t}, \ldots, u_{N,t})'$ and $\boldsymbol{\Phi}(L) = \mathrm{diag}(\phi_1(L), \phi_2(L), \ldots, \phi_N(L))$. In terms of parameters, the static factor model (3) may then be characterized by

$$\{\boldsymbol{\Lambda}, \mathbf{B}(L), \boldsymbol{\Phi}(L), \boldsymbol{\Gamma}_f(h), \boldsymbol{\Gamma}_\epsilon(h); h \in (-\infty, \infty)\}.$$

The autocovariances $\boldsymbol{\Gamma}_f(h)$ and $\boldsymbol{\Gamma}_\epsilon(h)$ for $h \neq 0$ are rarely of direct interest, and are not necessary for consistent estimation of the common component $\mathbf{c}_t$. In most cases, we are merely interested in $\{\boldsymbol{\Lambda}, \boldsymbol{\Upsilon}, \boldsymbol{\Psi}\}$, since $\boldsymbol{\Lambda}$ is, up to a rotation, asymptotically identifiable from $\boldsymbol{\Sigma}$ (that depends on $\boldsymbol{\Gamma}_f(0) = \boldsymbol{\Upsilon}$ and $\boldsymbol{\Gamma}_\epsilon(0) = \boldsymbol{\Psi}$). Given these parameters, the minimum mean square error (minimum MSE) predictor of the static factors is the projection (see, e.g., Anderson 2003, Sect. 14.7)

$$\hat{\mathbf{f}}_t = \left(\boldsymbol{\Upsilon}^{-1} + \boldsymbol{\Lambda}'\boldsymbol{\Psi}^{-1}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}'\boldsymbol{\Psi}^{-1}\mathbf{x}_t, \tag{5}$$

where $\boldsymbol{\Upsilon} = \mathbf{I}_\mathcal{R}$, under conventional normalization (see aforementioned).

Naturally, the parameters $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ are unknown, and need to be estimated. On theoretical grounds, maximum likelihood (ML) estimation is attractive. It is generally efficient and provides means for incorporating restrictions based on theory. However, ML estimators for dynamic factor models tend to be very complicated to derive, and full ML estimation is only available for special cases (see, e.g., Stoica and Jansson 2009; Doz et al. 2012). When the idiosyncratic component exhibits either time series dynamics, cross-sectional heteroscedasticity, or cross-sectional correlations, then full ML estimation is not attainable. However, by imposing misspecifying restrictions, subsets of the parameters may be consistently estimated by quasi-ML in the sense of White (1982). For example, by falsely assuming an exact factor model when the true model is an approximate factor model, the diagonal elements of $\boldsymbol{\Psi}$ (i.e., the contemporary idiosyncratic variances) and the space spanned by the columns of $\boldsymbol{\Lambda}$ may be consistently estimated by numerical quasi-ML estimation based on the iterative Expectation-Maximization (EM) algorithm; see Bai and Li (2012, 2016). In a similar fashion, Doz et al. (2012) have shown that the space spanned by the factors may be directly and consistently estimated by quasi-ML using the Kalman filter. If the procedure is iterated, then it is equivalent to the EM algorithm.

The workhorse for the static factor model is the method of principal components (PC). Consider the covariance matrix of $\mathbf{x}_t$, $\boldsymbol{\Sigma}$. Because every covariance matrix is positive semi-definite, it may be decomposed as $\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Pi}\mathbf{V}'$, where $\boldsymbol{\Pi} = \text{diag}(\varphi_1(\boldsymbol{\Sigma}), \varphi_2(\boldsymbol{\Sigma}), \ldots, \varphi_N(\boldsymbol{\Sigma}))$ is a diagonal matrix with the ordered positive eigenvalues of $\boldsymbol{\Sigma}$ (the principle components) on its main diagonal, and $\mathbf{V}$ is a matrix with the associated eigenvectors as columns, such that $\mathbf{V}'\mathbf{V} = \mathbf{I}_N$. Under the normalization $\boldsymbol{\Upsilon} = \mathbf{I}_{\mathcal{R}}$, the linear transformation $\mathbf{m}_t = \mathbf{V}'\mathbf{x}_t$ is the population PC estimator of the factors $\mathbf{f}_t$. It has contemporary covariance $Var(\mathbf{m}_t) = \mathbf{V}'\boldsymbol{\Sigma}\mathbf{V} = \boldsymbol{\Pi}$. Because $\boldsymbol{\Pi}$ is a diagonal matrix, the population factors are uncorrelated. Now, let $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)$. The first PC factor, $\hat{f}_{1,t} = \mathbf{v}_1'\mathbf{x}_t$, is the projection which maximizes the variance among all linear projections from unit vectors. Its variance is the first principal component $\varphi_1(\boldsymbol{\Sigma})$. The second PC factor, $\hat{f}_{2,t} = \mathbf{v}_2'\mathbf{x}_t$, maximizes the variance under the restriction of being orthogonal to the first PC factor. Its variance is the second principle component $\varphi_2(\boldsymbol{\Sigma})$, and so on. The PC estimator of the factor loadings is found be setting $\hat{\boldsymbol{\Lambda}}$ equal to the eigenvectors of $\boldsymbol{\Sigma}$ associated with its $\mathcal{R}$ largest eigenvalues. Replacing $\boldsymbol{\Sigma}$ with its sample counterpart $\mathbf{S} = T^{-1}\sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t'$, gives the sample PC estimators. Under an approximate factor model, they are consistent as $N, T \to \infty$ (i.e., as N and T tend to infinity simultaneously) for the spaces spanned by the factors and factor loadings, respectively; see, foremost, Stock and Watson (2002a) and Bai (2003).

The method of PC is a dimension reducing technique, and does, as opposed to ML, not require the existence of the static factor model (3). Yet, the PC and ML estimators are closely related. Under a Gaussian static factor model with spherical noise, the ML estimator of the factors is proportional to the PC estimator. The PC estimators of the factors and factor loadings are therefore often used as initial estimators for ML algorithms. For approximate factor models, the largest drawbacks of the PC estimators are that (i) they are inconsistent for fixed N, and (ii) they require a balanced panel. Meanwhile, ML estimation can be consistent for fixed N, and numerical procedures, such as the Kalman filter and the EM algorithm can smooth over missing values, allowing an unbalanced panel with missing values at the end or start of the panel; the so called "ragged edge" or "jagged edge" problem. This feature is very valuable in, for instance, economic forecasting, because key economic indicators tend to be released at different dates. In particular, and of special interest in this paper, Doz et al. (2011) show that, by consistently estimating $\mathbf{B}(L)$, the precision in estimating the factor space may be improved by setting up a state space solution and perform one run with the Kalman smoother to re-estimate the factors $\mathbf{f}_t$ for $t = 1, 2, \ldots, T$. This method is presented in detail in Sect. 4, and can be implemented in EViews by using our code in the supplementary material.

For any estimation approach, the number of factors $\mathcal{R}$ is generally unknown, and needs to be either estimated or assumed. Popular estimators for the number of factors in approximate factor models can be found in Bai and Ng (2002), Onatski (2010) and Ahn and Horenstein (2013). Throughout the paper, we will treat the number of factors as known.

3 A State Space Representation of the Static Factor Model

The linear time series process $\mathbf{x}_t$ can be cast in the state space form

$$\mathbf{x}_t = \mathbf{H}_t\boldsymbol{\alpha}_t + \boldsymbol{\xi}_t, \tag{6}$$

$$\boldsymbol{\alpha}_{t+1} = \mathbf{T}_t\boldsymbol{\alpha}_t + \mathbf{R}_t\boldsymbol{\eta}_t, \tag{7}$$

where $\boldsymbol{\alpha}_t$ ($k \times 1$) is a latent state vector, $\mathbf{H}_t$ ($N \times k$) and $\mathbf{T}_t$ ($k \times k$) are possibly time-varying parameter matrices and $\mathbf{R}_t$ ($k \times q; q \leq k$) is, in general, either the identity matrix or a selection matrix consisting of a subset of the columns of the identity matrix (see, e.g., Durbin and Koopman 2012). The system is stochastic through the $N \times 1$ vector $\boldsymbol{\xi}_t$ and the $k \times 1$ vector $\boldsymbol{\eta}_t$, which are mutually and serially uncorrelated with zero mean and contemporary covariance matrices $\boldsymbol{\Sigma}_\xi$ and $\boldsymbol{\Sigma}_\eta$, respectively. In the Gaussian state space form, which is what EViews handles, the errors are normally distributed: $\boldsymbol{\xi}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\xi)$, $\boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$. We refer to (6) as the *signal equation*, and to (7) as the *state equation*. Note that the state equation is a VAR(1) process.

The static factor model (3) can be written as a state space solution defined by (6) and (7), where the number of states relates to the latent components of the model, that is, the common factors and the idiosyncratic components. Moreover, as shown by Doz et al. (2011), neglecting the idiosyncratic time series dynamics, and thereby possibly misspecifying the underlying model, can still lead to consistent estimation of the central parameters of the factor model, given by the common component. Specifically, imposing the misspecification that $\boldsymbol{\epsilon}_t$ in (3) is white noise, the static factor model can be written in state space form where the number of states k is equal to the number of factors $\mathcal{R}$ times the number of VAR lags p: $k = \mathcal{R}p$. To see this, note that the factor VAR(p) process (4) can be written in stacked form as the VAR(1) process (see, e.g., Lütkepohl 2007, p. 15)

$$\tilde{\mathbf{f}}_t = \tilde{\mathbf{B}}\tilde{\mathbf{f}}_{t-1} + \tilde{\mathbf{z}}_t, \tag{8}$$

where

$$\tilde{\mathbf{f}}_t = \underbrace{\begin{pmatrix} \mathbf{f}_t \\ \mathbf{f}_{t-1} \\ \vdots \\ \mathbf{f}_{t-p+1} \end{pmatrix}}_{\mathcal{R}p \times 1}, \quad \tilde{\mathbf{B}} = \underbrace{\begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_{p-1} & \mathbf{B}_p \\ \mathbf{I}_\mathcal{R} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_\mathcal{R} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_\mathcal{R} & \mathbf{0} \end{pmatrix}}_{\mathcal{R}p \times \mathcal{R}p}, \quad \tilde{\mathbf{z}}_t = \underbrace{\begin{pmatrix} \mathbf{z}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}}_{\mathcal{R}p \times 1}. \tag{9}$$

Thus, if the static factors follow the VAR process (4) and the idiosyncratic components are serially uncorrelated, then the static factor model (3) has a state space representation defined by (6) and (7) with

$$\mathbf{H}_t = \underbrace{\begin{pmatrix} \boldsymbol{\Lambda} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}}_{N \times \mathcal{R}p}, \ \boldsymbol{\alpha}_t = \tilde{\mathbf{f}}_t, \ \boldsymbol{\xi}_t = \boldsymbol{\epsilon}_t, \ \mathbf{R}'_t = \underbrace{\begin{pmatrix} \mathbf{I}_\mathcal{R} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}}_{\mathcal{R} \times \mathcal{R}p}, \ \mathbf{T}_t = \tilde{\mathbf{B}}, \ \boldsymbol{\eta}_t = \mathbf{z}_{t+1}, \tag{10}$$

where $\boldsymbol{\Lambda}$ and $\boldsymbol{\epsilon}_t$ are the factor loadings and idiosyncratic error in (3), $\tilde{\mathbf{f}}_t$ and $\tilde{\mathbf{B}}$ are the stacked factors and parameters in (9) and $\mathbf{z}_t$ is the error in (4). Here, the subscripts may be dropped from $\mathbf{H}_t$, $\mathbf{T}_t$ and $\mathbf{R}_t$, since in this case, the parameters are not time-varying.

Specification and estimation of state space models in EViews is outlined in IHS Global Inc. (2015d, Chapter 39). A recent demonstration is found in van den Bossche (2011). The estimation concerns two aspects: (i) measuring the unknown states $\boldsymbol{\alpha}_t$ for $t = 1, 2, \ldots, T$, involving prediction, filtering and smoothing (see Sect. 3.1), and (ii) estimation of the unknown parameters $\mathbf{H}_t$, $\mathbf{T}_t$, $\boldsymbol{\Sigma}_\xi$ and $\boldsymbol{\Sigma}_\eta$. Doz et al. (2011) propose to estimate the parameters by PC (see Sect. 2.1), and leave only the estimation of the factors (i.e., the states) to the state space form (see Sect. 4).

Remark 3.1 The Eq. (7) is specified for the states in period $t + 1$, given the errors in period t, which requires some consideration when modeling correlations between the signal and state errors in EViews. However, by Assumption A3, the state and signal errors are mutually uncorrelated. Therefore, the construction of the temporal indices of the state and signal equations do not affect the methods we use in this paper. For a state space specification with contemporary error indices, see, e.g., Hamilton (1994, p. 372).

3.1 The Kalman Filter and Smoother

The latent state vector $\boldsymbol{\alpha}_t$ can be estimated numerically by the Kalman filter and smoother (see, e.g., Harvey 1989; Durbin and Koopman 2012, for thorough treatments). Consider the conditional distribution of $\boldsymbol{\alpha}_t$, based on the available information at time $t - 1$, $\mathcal{F}_{t-1}$. Under the Gaussian state space model, the distribution is normal $\boldsymbol{\alpha}_t | \mathcal{F}_{t-1} \sim \mathcal{N}(\mathbf{a}_{t|t-1}, \mathbf{P}_{t|t-1})$, where

$$\mathbf{a}_{t|t-1} = E(\boldsymbol{\alpha}_t | \mathcal{F}_{t-1}), \quad \mathbf{P}_{t|t-1} = Var(\boldsymbol{\alpha}_t | \mathcal{F}_{t-1}).$$

By construction, $\mathbf{a}_{t|t-1}$ is the minimum MSE estimator of (the Gaussian) $\boldsymbol{\alpha}_t$, with MSE matrix $\mathbf{P}_{t|t-1}$. Given the conditional mean, we can find the minimum MSE estimator of $\mathbf{x}_t$ from (6),

$$\hat{\mathbf{x}}_t = E(\mathbf{x}_t | \mathbf{a}_{t|t-1}) = \mathbf{H}_t \mathbf{a}_{t|t-1},$$

with error $\mathbf{v}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$ and associated $N \times N$ error covariance matrix

$$\mathbf{F}_t = Var(\mathbf{v}_t) = \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + \boldsymbol{\Sigma}_\xi,$$

where $\boldsymbol{\Sigma}_\xi$ was defined in relation to (6).

The Kalman filter is a recursion over $t = 1, 2, \ldots, T$ that, based on the error $\mathbf{v}_t$ and the dispersion matrix $\mathbf{F}_t$, sequentially updates the means and variances in the conditional distributions $\boldsymbol{\alpha}_t | \mathcal{F}_t \sim \mathcal{N}(\mathbf{a}_{t|t}, \mathbf{P}_{t|t})$ and $\boldsymbol{\alpha}_{t+1} | \mathcal{F}_t \sim \mathcal{N}(\mathbf{a}_{t+1|t}, \mathbf{P}_{t+1|t})$ by

$$\mathbf{a}_{t|t} = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{H}_t'\mathbf{F}_t^{-1}\mathbf{v}_t,$$
$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{H}_t'\mathbf{F}_t^{-1}\mathbf{H}_t\mathbf{P}_t,$$
$$\mathbf{a}_{t+1|t} = \mathbf{T}_t\mathbf{a}_{t|t},$$
$$\mathbf{P}_{t+1|t} = \mathbf{T}_t\mathbf{P}_{t|t}\mathbf{T}_t' + \mathbf{R}_t\mathbf{\Sigma}_\eta\mathbf{R}_t', \tag{11}$$

where $\mathbf{T}_t$ and $\mathbf{\Sigma}_\eta$ were defined in relation to (7). The computational complexity of the recursion depends largely on the inversion of $\mathbf{F}_t$. Note that the second term in (11) may be viewed as a correction term, based on the last observed error $\mathbf{v}_t$. EViews refers to $\mathbf{a}_{t|t}$ as the *filtered estimate* of $\boldsymbol{\alpha}_t$, and to $\mathbf{a}_{t+1|t}$ as the *one-step ahead prediction* of $\boldsymbol{\alpha}_t$. The recursion requires the initial one-step ahead predicted values $\mathbf{a}_{1|0}$ and its associated covariance matrix $\mathbf{P}_{1|0}$. If $\mathbf{x}_t$ is a stationary process, then they can be found analytically. Otherwise, EViews uses *diffuse priors*, following Koopman et al. (1999). The user may also provide the initial conditions; see van den Bossche (2011) and IHS Global Inc. (2015d, p. 683). Additionally, we need estimates of the parameters $\mathbf{H}_t$, $\mathbf{T}_t$, $\mathbf{\Sigma}_\xi$ and $\mathbf{\Sigma}_\eta$. Conveniently, the Kalman filter provides the likelihood function as a by-product from the one-step ahead prediction errors (see Harvey 1989, Sect. 3.4.), and the recursion can therefore be based on maximum likelihood estimators of the parameters; EViews offers various numerical optimization routines to find the associated estimates. In this paper, however, we follow Doz et al. (2011) and estimate the main parameters of the dynamic factor model (3) by PC; only the factors are estimated using the Kalman filter. This method is consistent and generally much faster than a recursion that includes parameter estimation.

The Kalman filter is a forward recursion. By applying a backward (smoothing) recursion using the output from the Kalman filter, we can find $\mathbf{a}_{t|T}$, the $\mathcal{F}_T$-conditional (i.e., conditional on *all* available observations) minimum MSE estimator of $\boldsymbol{\alpha}_t$, and its associated MSE matrix $\mathbf{P}_{t|T}$. There are different kinds of smoothing algorithms (see, e.g., Durbin and Koopman 2012, Sect. 4.4). EViews uses a fixed-interval smoothing, which in its classical form returns the estimated means and variances of the conditional distributions $\boldsymbol{\alpha}_t|\mathcal{F}_T \sim \mathcal{N}(\mathbf{a}_{t|T}, \mathbf{P}_{t|T})$ by

$$\mathbf{a}_{t|T} = \mathbf{a}_{t|t} + \mathbf{P}_{t|t}\mathbf{T}_t'\mathbf{P}_{t+1|t}^{-1}\left(\mathbf{a}_{t+1|T} - \mathbf{a}_{t+1|t}\right),$$
$$\mathbf{P}_{t|T} = \mathbf{P}_{t|t} - \mathbf{P}_{t|t}\mathbf{T}_t'\mathbf{P}_{t+1|t}^{-1}\left(\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t}\right)\mathbf{P}_{t+1|t}^{-1}\mathbf{T}_t\mathbf{P}_{t|t},$$

for $t = T, T-1, \ldots, 1$.

Because the smoothed estimator $\mathbf{a}_{t|T}$ is based on all observations, its MSE cannot be larger than the MSE from the filtered estimator $\mathbf{a}_{t|t}$, in the sense that the MSE matrix of the latter is the MSE matrix of the smoothed estimator plus some positive definite matrix.[5]

The Kalman filter and smoother offer an exceptionally easy way of handling missing values, whereby the matrix $\mathbf{H}_t$ is simply set to zero (see, e.g., Durbin and Koopman

---

[5] If the state space model is not Gaussian, then the Kalman filter and smoother do not in general provide the conditional means, and the associated estimators are no longer the minimum MSE estimators. They are, however, the *linear* minimum MSE estimators.

2012, Sect. 4.1). This treatment preserves optimality. Similarly, MSE-optimal forecasts are conducted by treating future values of $\mathbf{x}_t$ as missing observations.

4 The Two-Step Estimator of the Common Factors

Doz et al. (2011) show that misspecifying the static factor model (3) with respect to some of the dynamics and cross-sectional properties may still lead to consistent estimation of the space spanned by the common factors. They propose to estimate the common factors in two steps. In the first step, preliminary parameter estimates are computed by PC. In the second step, the factors are re-estimated by MSE-optimal linear projection from one run of the Kalman smoother, allowing for idiosyncratic cross-sectional heteroscedasticity, common factor dynamics, as well as an unbalanced panel.

Because the factors and factor loadings are not uniquely identified, Doz et al. (2011) consider a specific rotation, outlined in the following way. Under the normalization $\mathbf{\Upsilon} = \mathbf{I}_\mathcal{R}$, we have that $\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Psi}$, so that $\mathbf{\Lambda}$ is identified up to an orthogonal multiplication (see Sect. 2). Let $\mathbf{\Lambda}'\mathbf{\Lambda}$ have spectral decomposition $\mathbf{\Lambda}'\mathbf{\Lambda} = \mathbf{Q}\mathbf{D}\mathbf{Q}'$, where $\mathbf{D} = \operatorname{diag}(\varphi_1(\mathbf{\Lambda}'\mathbf{\Lambda}), \ldots, \varphi_\mathcal{R}(\mathbf{\Lambda}'\mathbf{\Lambda}))'$ and $\mathbf{Q}\mathbf{Q}' = \mathbf{I}_N$, and consider the following representation of the factor model (3):

$$\mathbf{x}_t = \mathbf{\Lambda}_+ \mathbf{g}_t + \boldsymbol{\epsilon}_t,$$

where $\mathbf{\Lambda}_+ = \mathbf{\Lambda}\mathbf{Q}$ and $\mathbf{g}_t = \mathbf{Q}'\mathbf{f}_t$. By construction, $\mathbf{\Lambda}'_+\mathbf{\Lambda}_+ = \mathbf{D}$ and $\mathbf{\Lambda}_+\mathbf{\Lambda}'_+ = \mathbf{\Lambda}\mathbf{\Lambda}'$, such that there exists a matrix $\mathbf{P} = \mathbf{\Lambda}_+\mathbf{D}^{-1/2}$ with the property $\mathbf{P}'\mathbf{P} = \mathbf{I}_\mathcal{R}$. Because $\mathbf{f}_t$ has a VAR representation (4), $\mathbf{g}_t$ also has a VAR representation,

$$\mathbf{B}^+(L)\mathbf{g}_t = \mathbf{w}_t, \tag{12}$$

where $\mathbf{B}^+(L) = \mathbf{Q}'\mathbf{B}(L)\mathbf{Q}$ and $\mathbf{w}_t = \mathbf{Q}'\mathbf{z}_t$. In many cases, estimating $\mathbf{g}_t$, or any other rotation of $\mathbf{f}_t$, is as good as estimating $\mathbf{f}_t$ itself (see Sect. 2). In particular, the conditional mean, which is used when forecasting under squared error loss, is unaffected by the rotation.

Suppose $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T)$ is the $N \times T$ matrix of standardized and balanced panel data with sample covariance matrix $\mathbf{S} = T^{-1}\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}'_t = T^{-1}\mathbf{X}\mathbf{X}'$. Let $\hat{\mathbf{D}} = \operatorname{diag}(d_1, d_2, \ldots, d_\mathcal{R})$ be a diagonal matrix with the $\mathcal{R}$ largest eigenvalues of $\mathbf{S}$ on its main diagonal, and let $\hat{\mathbf{P}}$ be the $\mathcal{R} \times \mathcal{R}$ matrix with the associated eigenvectors as columns. Under the specific rotation $\mathbf{Q}'\mathbf{f}_t$, the PC estimators of the factors and factor loadings (see Sect. 2.1) are

$$\hat{\mathbf{g}}_t = \hat{\mathbf{D}}^{-1/2}\hat{\mathbf{P}}'\mathbf{x}_t, \tag{13}$$

$$\hat{\mathbf{\Lambda}}_+ = \hat{\mathbf{P}}\hat{\mathbf{D}}^{1/2}. \tag{14}$$

If the covariance decomposition in $\mathbf{\Sigma}$ is identifiable (see Sect. 2), then under Assumptions A1–A3 in Sect. 2 it holds that $\hat{\mathbf{g}}_t \overset{p}{\to} \mathbf{g}_t$ and $\hat{\mathbf{\Lambda}}_+ \overset{p}{\to} \mathbf{\Lambda}_+$, as $N, T \to \infty$, where

$\overset{p}{\to}$ denotes convergence in probability. The PC estimators are thus consistent, and in that sense they suffice. They do not, however, exploit the factor time series dynamics imposed by $\mathbf{B}(L)$, or the fact that the idiosyncratic components are potentially cross-sectionally heteroscedastic (i.e., that the diagonal elements of $\boldsymbol{\Psi}$ are possibly different). Also, they require a balanced panel. Conveniently, the Kalman filter offers a solution to these issues.

Suppose that the true model (3) fulfills Assumptions A1–A3. Let this model be characterized by $\Omega = \{\boldsymbol{\Lambda}, \mathbf{B}(L), \boldsymbol{\Phi}(L), \boldsymbol{\Psi}\}$. Doz et al. (2011) consider two misspecifications of Ω where the Kalman smoother can be used to exploit the dynamics of the common factors composed in $\mathbf{B}(L)$: one model that is characterized by $\Omega^{\mathcal{A}3} = \{\boldsymbol{\Lambda}, \mathbf{B}(L), \mathbf{I}_N, \psi \mathbf{I}_N\}$, where ψ is a constant, and a second model that is characterized by $\Omega^{\mathcal{A}4} = \{\boldsymbol{\Lambda}, \mathbf{B}(L), \mathbf{I}_N, \boldsymbol{\Psi}_d\}$, where $\boldsymbol{\Psi}_d = \mathrm{diag}(\psi_{1,1}, \ldots, \psi_{N,N})$ is a diagonal matrix with the diagonal elements of $\boldsymbol{\Psi}$ on its main diagonal. The parameters of the the approximating models $\Omega^{\mathcal{A}3}$ and $\Omega^{\mathcal{A}4}$ can be consistently estimated even though the true model is characterized by Ω. Moreover, given the parameters, the precision in the factors can be improved by linear projection using the Kalman filter based on the state space solution in Sect. 3.

Here, imposing the rotation $\mathbf{g}_t = \mathbf{Q}' \mathbf{f}_t$, implies that $\mathbf{g}_t$ by assumption follows the VAR representation (12). Hence, for the state space solution in Sect. 3, $\{\mathbf{g}_t, \mathbf{B}_1^+, \ldots, \mathbf{B}_p^+, \mathbf{w}_t, \boldsymbol{\Lambda}^+\}$ replace $\{\mathbf{f}_t, \mathbf{B}_1, \ldots, \mathbf{B}_p, \mathbf{z}_t, \boldsymbol{\Lambda}\}$ in the representation (10). The signal and state equations (6) and (7) are then, respectively,

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{N,t} \end{pmatrix} = (\boldsymbol{\Lambda}_+ \; \mathbf{0} \; \cdots \; \mathbf{0}) \begin{pmatrix} \mathbf{g}_t \\ \mathbf{g}_{t-1} \\ \vdots \\ \mathbf{g}_{t-p+1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{N,t} \end{pmatrix}, \tag{15}$$

$$\begin{pmatrix} \mathbf{g}_t \\ \mathbf{g}_{t-1} \\ \vdots \\ \mathbf{g}_{t-p+1} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1^+ & \mathbf{B}_2^+ & \cdots & \mathbf{B}_{p-1}^+ & \mathbf{B}_p^+ \\ \mathbf{I}_{\mathcal{R}} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\mathcal{R}} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{\mathcal{R}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{g}_{t-1} \\ \mathbf{g}_{t-2} \\ \vdots \\ \mathbf{g}_{t-p} \end{pmatrix} + \begin{pmatrix} \mathbf{w}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}. \tag{16}$$

As shown by Doz et al. (2011), the following steps provide a consistent method for estimating the factors under $\Omega^{\mathcal{A}3}$ and $\Omega^{\mathcal{A}4}$:

1. Estimate $\boldsymbol{\Lambda}_+$ and $\mathbf{g}_t$ for $t = 1, 2, \ldots, T$ by the rotated PC estimators (13) and (14).
2. Estimate the VAR polynomial $\mathbf{B}^+(L)$ by the ordinary least squares (OLS) regression

$$\hat{\mathbf{g}}_t = \mathbf{B}_1^+ \hat{\mathbf{g}}_{t-1} + \mathbf{B}_2^+ \hat{\mathbf{g}}_{t-2} + \cdots + \mathbf{B}_p^+ \hat{\mathbf{g}}_{t-p} + \hat{\mathbf{w}}_t. \tag{17}$$

Under standard aforementioned assumptions it holds as $N, T \to \infty$ that $\hat{\mathbf{B}}_m^+ \overset{p}{\to} \mathbf{B}_m^+$, for $m = 1, 2, \ldots, p$.

3. Run the Kalman smoother over the state space model defined by (15) and (16) to re-estimate the factors $\mathbf{g}_t$ for $t = 1, 2, \ldots, T$, conditional on the estimates $\hat{\mathbf{\Lambda}}_+$ and $\hat{\mathbf{B}}^+(L)$.

To apply the Kalman filter we need to know the covariance matrices $\mathbf{\Sigma}_\xi$ and $\mathbf{\Sigma}_\eta$ (see Sect. 3), which here correspond to, respectively, the idiosyncratic covariance matrix $\mathbf{\Psi}$ and the covariance matrix for $\mathbf{w}_t$ in (12), $Var(\mathbf{w}_t)$. Under Assumptions A1–A3, $\mathbf{S} \overset{p}{\to} \mathbf{\Sigma}$, as $N, T \to \infty$. Thus, we may consistently estimate $\mathbf{\Psi}$ by $\hat{\mathbf{\Psi}} = \mathbf{S} - \hat{\mathbf{\Lambda}}_+ \hat{\mathbf{\Lambda}}'_+ \overset{p}{\to} \mathbf{\Sigma} - \mathbf{\Lambda}_+ \mathbf{\Lambda}'_+ = \mathbf{\Sigma} - \mathbf{\Lambda} \mathbf{\Lambda}' = \mathbf{\Psi}$. Similarly, a consistent estimator of $Var(\mathbf{w}_t)$ may be based on the residuals in (17).

The model defined by $\Omega^{\cdot A4}$ is the most general of the two approximating models, and is therefore expected to have the highest precision in estimating the factors, unless the model defined by $\Omega^{\cdot A3}$ is in fact the true (or very close to the true) model.

4.1 Implementing the Two-Step Estimator in EViews

In the supplementary material, we provide the file `subroutine_dfm` containing a subroutine named `DFM` that estimates the approximating models characterized by $\Omega^{\cdot A3}$ and $\Omega^{\cdot A4}$. To call the subroutine, either the subroutine code should be placed directly into a user's main script, or the file `subroutine_dfm` should be included using the command `Include`; see Chapter 6 in IHS Global Inc. (2015a).

EViews works around *objects*, consisting of information related to a specific choice of analysis. Two types of objects are frequently used when programming in EViews: string objects (i.e., sequences of characters) and scalar objects (i.e., numbers). However, in the code we make more extensive use of the related concepts *string variables* and *control variables*, which are temporary variables whose values are strings and scalars, respectively, and that only exist while the EViews program executes. String variables have names that begin with a "%" (see IHS Global Inc. 2015a, p. 92). Control variables have names that begin with a "!" (see IHS Global Inc. 2015a, p. 126). For example, the commands `%ser = "gdp"` and `!n = 5` create a string variable `%ser` containing the characters 'gdp' and a control variable `!n` containing the number 5. By enclosing these variables in curly brackets, "{" and "}", EViews will replace the expressions with the underlying variable value (see section on *replacement variables* in IHS Global Inc. 2015a, p. 130). For example, the commands `series {%ser}` and `Group G{!n}` create a series object named `gdp` and a group object named `G5`.

The subroutine `DFM` is defined by

```
DFM(Group XGrp, Scalar FNum, Scalar VLag, Sample S,
 String Model)
```

where each argument is specified by an EViews object: a group object `XGrp` containing the observable time series in $\mathbf{x}_t$, a scalar object `FNum` containing the number of factors, a scalar object `VLag` containing the number of lags in the factor VAR representation (12), a sample object `S` over which the Kalman smoother should estimate the states, and a string object `Model` that should be set to "A3" for the model characterized by

Table 1 Summary of objects that are stored by the subroutine `DFM`

Object name	Description	Object type
`LambdaHat`	PC-estimated factor loadings: $\hat{\mathbf{\Lambda}}_{+}$	Matrix ($N \times \mathcal{R}$)
`CovEpsHat`	PC-estimated idiosyncratic covariance matrix: $\hat{\mathbf{\Psi}}$	Matrix ($N \times N$)
`BHat`	Estimated factor VAR coefficients: $(\hat{\mathbf{B}}_{(1)}, \hat{\mathbf{B}}_{(2)}, \ldots, \hat{\mathbf{B}}_{(\mathcal{R})})'$	Matrix ($\mathcal{R}p \times \mathcal{R}$)
`CovWHat`	Estimated factor VAR error covariance matrix: $\hat{Var}(\mathbf{w}_t)$	Matrix ($\mathcal{R} \times \mathcal{R}$)
`DFMSS`	State space specification (state space object)	Sspace
`SVk_m`	State relating to factor k, lag m	Series

$\Omega^{\mathcal{A}3}$, or to "A4" for the model characterized by $\Omega^{\mathcal{A}4}$. The subroutine is called by the keyword `Call`. Suppose, for example, that we want to estimate a dynamic factor model under $\Omega^{\mathcal{A}4}$ using two factors that are VAR(1). If there is a time series group object `G` and a sample object `J`, then the subroutine may be called by

```
Call DFM(G, 2, 1, J, "A4")
```

Within the subroutine, `XGrp` is assigned the series in the group `G`, `FNum` is assigned the number 2, `VLag` is assigned the number 1, `S` is assigned the sample `J`, and `Model` is assigned the characters 'A4'. The objects within the subroutine are *global*, meaning that any changes to the objects inside the subroutine will change the very objects or variables that are passed into the subroutine. For instance, in the subroutine, the series in `XGrp` are standardized prior to the PC estimation. Hence, this will standardize the series in the group that is used to call the subroutine (`G`, in the example). EViews offers also the possibility to use *local* subroutines; see IHS Global Inc. (2015a, p. 156).

Table 1 displays a summary of the objects that are stored in the Eviews workfile. These objects belong to the state space specification of the dynamic factor model, implying that they are required to produce certain output, such as state series graphs, and state or signal forecasts (see IHS Global Inc. 2015d, Chapter 39). Note that once the subroutine is called, the objects will be written over. Thus, if the subroutine is called several times in an active workfile, say in a loop, then only the objects from the last run will be available, unless they are consecutively stored by the user.

The subroutine executes the steps 1–3 as outlined above.[6] First, the data are standardized over the balanced panel, and the factors and factor loadings are estimated by (13) and (14).[7] Next, we estimate the factor VAR (17), producing the matrix `BHat` (see Table 1). By default, `BHat` is the $\mathcal{R}p \times \mathcal{R}$ matrix $(\hat{\mathbf{B}}_{(1)}, \hat{\mathbf{B}}_{(2)}, \ldots, \hat{\mathbf{B}}_{(\mathcal{R})})'$, where

---

[6] The data set could contain missing values within the balanced panel. In that case, EViews will disregard entire rows (each relating to an observation number) for matrix operations, whereby the one-to-one relation between time and the data positions in matrices will be lost. To avoid such cases, we check each series for missing values within the balanced panel. If, for a series, missing values are found, then the corresponding series is removed from the group.

[7] The sample covariance matrix $\mathbf{S}$ is stored in lower triangular form using a `Sym` object, which is required by EViews for collecting its eigenvalues and eigenvectors.

$\hat{\mathbf{B}}_{(j)}$ is a matrix containing the jth columns of each of the estimated autoregressive coefficient matrices $\hat{\mathbf{B}}_k^+ (k = 1, 2, \ldots, p)$ from the regression (17). That is, the first row of BHat is the first column of $\hat{\mathbf{B}}_1^+$, the second row is the first column of $\hat{\mathbf{B}}_2^+$, and so on. For more details, see IHS Global Inc. (2015b, p. 815).[8]

Lastly we set up a state space object and run the Kalman smoother. For the state space object we need to declare signal and state properties (see van den Bossche 2011), where every line in (15) and (16) has to be specified. Moreover, if an error should exist for a specific equation line, then it must be specified. In the signal equation (15), all lines should have errors, whereas in the state equation (16), only the first $\mathcal{R}$ lines should have errors (relating to the elements in $\mathbf{w}_t$). These errors are named by the keyword @ename. The signal errors are named e1, e2, and so on, and the state errors are named w1, w2, and so on.

The error variances and covariances are specified using the keyword @evar. Variances and covariances that are not specified are by default zero. Following the estimation procedure for the models defined by $\Omega^{\mathcal{A}4}$ and $\Omega^{\mathcal{A}3}$, the signal errors are, by misspecification, uncorrelated over both time and the cross-section. As such, we are only concerned with the contemporary signal error variances (the diagonal elements of $\mathbf{\Psi}$, that also correspond to the diagonal elements of $\mathbf{\Sigma}_\xi$ in Sect. 3). For the model $\Omega^{\mathcal{A}4}$, these parameters are assigned the PC-estimated variances $\hat{\psi}_{1,1}, \hat{\psi}_{2,2}, \ldots, \hat{\psi}_{N,N}$, collected in the diagonal of CovEpsHat (see Table 1). For the model $\Omega^{\mathcal{A}3}$, we impose the restriction $\mathbf{\Psi} = \psi \mathbf{I}_N$, where, following Doz et al. (2011), ψ is estimated by the mean of the estimated idiosyncratic variances, $\hat{\psi} = N^{-1} \sum_{i=1}^N \hat{\psi}_{i,i} = N^{-1} \text{trace}(\hat{\mathbf{\Psi}})$. The state errors relate to the factor VAR residuals, and are, under correct specification, also white noise. The variances and covariances of the $\mathcal{R}$ first state errors are assigned the elements in the residual covariance matrix from the estimated regression (17), CovWHat (see Table 1).

After declaring the signal and state error properties, the subroutine DFM defines and appends the signal and state equations to the state space object DFMSS using the keywords @signal and @state. The output of main interest are the estimated factors. The subroutine lets SVk_m refer to factor k, lag m. Thus, we are primarily interested in SV1_0, SV2_0, …, SVR_0, i.e., the states referring the contemporary factors. The remaining states, SV1_1, SV2_1, …, SVR_p, refer to lags of the factors, and are simply created for the sole reason to complete the markovian state space solution for the Kalman filter and smoother (see Sect. 4). Accordingly, the state space object allows only one-period lags of the states.

As an example, say that we wish to model 50 time series in the vector $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \ldots, x_{50,t})'$ by the static factor model (3) with two factors that follow a VAR(2). There are then $\mathcal{R}p = 2 \times 2 = 4$ states in the state vector $\boldsymbol{\alpha}_t = (g_{1,t}, g_{2,t}, g_{1,t-1}, g_{2,t-1})'$, which by the subroutine are named SV1_0, SV2_0, SV1_1 and SV2_1, respectively. The signal equation (15) is now

---

[8] We do not check that the estimated VAR is stationary, which for instance could be done by checking that the eigenvalues of the companion matrix $\tilde{\mathbf{B}}$ in (8) all lie within the complex unit circle. This procedure is easily accomplished in EViews version 9.5 and later, since the companion matrix is then available as a function return.

$$
\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{50,t} \end{pmatrix} = \begin{pmatrix} \lambda_{1,1}^+ & \lambda_{1,2}^+ & 0 & 0 \\ \lambda_{2,1}^+ & \lambda_{2,2}^+ & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{50,1}^+ & \lambda_{50,2}^+ & 0 & 0 \end{pmatrix} \begin{pmatrix} g_{1,t} \\ g_{2,t} \\ g_{1,t-1} \\ g_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{50,t} \end{pmatrix}, \tag{18}
$$

where $[\lambda_{i,j}^+] = \mathbf{\Lambda}_+$, a matrix with elements $\lambda_{i,j}^+$ corresponding to the ith row and jth column, and the state equation (16) is

$$
\begin{pmatrix} g_{1,t} \\ g_{2,t} \\ g_{1,t-1} \\ g_{2,t-1} \end{pmatrix} = \begin{pmatrix} b_{(1)1,1}^+ & b_{(1)1,2}^+ & b_{(2)1,1}^+ & b_{(2)1,2}^+ \\ b_{(1)2,1}^+ & b_{(1)2,2}^+ & b_{(2)2,1}^+ & b_{(2)2,2}^+ \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} g_{1,t-1} \\ g_{2,t-1} \\ g_{1,t-2} \\ g_{2,t-2} \end{pmatrix} + \begin{pmatrix} w_{1,t} \\ w_{2,t} \\ 0 \\ 0 \end{pmatrix}, \tag{19}
$$

where $[b_{(1)j,k}^+] = \mathbf{B}_1^+$ and $[b_{(2)j,k}^+] = \mathbf{B}_2^+$.

Suppose, for simplicity, that the time series in $\mathbf{x}_t$ are named x1, x2, …, x50 in the EViews workfile. The subroutine declares (18) and (19) line by line. First, the $N = 50$ equations are declared as

```
x1 = LambdaHat(1, 1) * SV1_0 + LambdaHat(1, 2) * SV2_0 + e1
x2 = LambdaHat(2, 1) * SV1_0 + LambdaHat(2, 2) * SV2_0 + e2
.
x50 = LambdaHat(50, 1) * SV1_0 + LambdaHat(50, 2) * SV2_0 + e50
```

Then, the $\mathcal{R} = 2$ first state equations are declared as

```
SV1_0 = BHat(1, 1) * SV1_0(-1) + BHat(2, 1) * SV1_1(-1) + _
         BHat(3, 1) * SV2_0(-1) + BHat(4, 1) * SV2_1(-1) + w1
SV2_0 = BHat(1, 2) * SV1_0(-1) + BHat(2, 2) * SV1_1(-1) + _
         BHat(3, 2) * SV2_0(-1) + BHat(4, 2) * SV2_1(-1) + w2
```

(where the underscore "_" joins the lines) and, lastly, the remaining $\mathcal{R}(p - 1) = 2$ state equations are declared as

```
SV1_1 = SV1_0(-1)
SV2_1 = SV2_0(-1)
```

To run the Kalman smoother in EViews, we need to set up the ML estimation procedure provided by the Kalman filter (see Sect. 3), even though there are no state space parameters to be estimated, as we only seek the smoothed states, given the PC-estimated parameters. The smoothing algorithm is provided by the command `Makestates(t = smooth) *`, where the * implies that output will have the same name as the input, meaning that the smoothed states will have names `SV1_0`, `SV2_0`, `SV3_0`, …, `SV1_1`, `SV2_1`, `SV3_1`, and so on, each relating to a factor and its lags. If we want filtered states instead of smoothed states (see Sect. 3.1), then we can change the option to `t = filt`. The subroutine code has comments that explain individual steps. A few final remarks are presented below.

Remark 4.1 We have assumed that the number of factors $\mathcal{R}$ is known. In practice, this is rarely the case. We could of course estimate the number of factors (see references in

Sect. 2.1); though, for forecasting or nowcasting (see Sect. 6), the appropriate number of factors is rather a practical concern, and can be found from forecasting evaluations. The same reasoning applies to the number of lags in the factor VAR regression (17). In case we wish to estimate the number of lags, however, we may use the criteria available in EViews; see IHS Global Inc. (2015b, p. 838) and IHS Global Inc. (2015d, Sect. 38). These criteria are also discussed in Lütkepohl (2007, Sect. 4.3).

Remark 4.2 Forecasts can be made for the signal and state equations using the EViews built-in procedures; see IHS Global Inc. (2015d), p. 676. Because future values are treated as missing data by the Kalman smoother, smoothed state forecasts can be carried out by simply extending the sample to include the forecast period.

5 Monte Carlo Simulation

For consistency, we replicate the parts of the simulation study by Doz et al. (2011) that compares the precision of the models defined by $\Omega^{\cdot A4}$ and $\Omega^{\cdot A3}$. Code that reproduces the simulation study is available in the supplementary material.

Doz et al. (2011) consider a one-factor ($\mathcal{R} = 1$) setup with the following data generating process (DGP):

$$x_{i,t} = \lambda_i f_t + \epsilon_t \ (i = 1, 2, \ldots, N; \ t = 1, 2, \ldots, T),$$
$$\lambda_i \sim \ \text{i.i.d.} \ \mathcal{N}(0, 1),$$
$$(1 - bL)f_t = z_t; \ z_t \ \text{i.i.d.} \ \mathcal{N}(0, 1 - b^2),$$
$$\boldsymbol{\Phi}(L)\epsilon_t = \mathbf{u}_t; \ \mathbf{u}_t = (u_{1,t}, u_{2,t}, \ldots, u_{N,t})' \ \text{i.i.d.} \ \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_u), \ \boldsymbol{\Phi}(L) = (1 - \phi L)\mathbf{I}_N,$$
$$\boldsymbol{\Sigma}_u = [\tau_{i,j}]; \ \tau_{i,j} = \sqrt{\kappa_i \kappa_j}\delta^{|i-j|}(1 - \phi^2), \kappa_i = \frac{\beta_i}{1 - \beta_i}\lambda_i^2,$$
$$\beta_i \ \text{i.i.d.} \ \mathcal{U}(m, 1 - m), (i, j = 1, 2, \ldots, N),$$

where $\mathcal{U}$ denotes the uniform distribution. That is, the factor and idiosyncratic components are AR(1) processes, $f_t = bf_{t-1} + z_t, \epsilon_{i,t} = \phi\epsilon_{i,t-1} + u_{i,t}$ (i.e., the idiosyncratic AR-coefficients are the same over the cross-section), where the idiosyncratic components are possibly cross-sectionally dependent through the covariance matrix $\boldsymbol{\Sigma}_u$. The constant δ controls for the amount of idiosyncratic cross-correlation, where $\delta = 0$ implies the exact factor model with no cross-sectional dependence. The constant β_i controls the signal-to-noise ratio, i.e., the ratio between the variance of $x_{i,t}$ and the variance of the idiosyncratic component $\epsilon_{i,t}$. For this constant we have used the well-known result that a uniformly distributed variable in the interval $[m_l, m_u]$ may be generated as $m_l + (m_u - m_l)\mu$, where $\mu \sim \mathcal{U}(0, 1)$. The correlated idiosyncratic errors are generated using a Cholesky decomposition, $\boldsymbol{\Sigma}_u = \mathbf{LL}'$, where $\mathbf{L}$ is a lower diagonal matrix. The multivariate normal error $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_u)$ is then generated as $\mathbf{L}\mathbf{e}_t$, where $\mathbf{e}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$, for $t = 1, 2, \ldots, T$.

The autoregressive coefficients are set to $b = 0.9$ and $\phi = 0.5$, and the constants δ and m are set to $\delta = 0.5$ and $m = 0.1$. Note that the variances of the errors z_t and

$u_{i,t}$ are scaled, so that the variances of f_t and $\epsilon_{i,t}$ are 1 and κ_i, respectively.[9] For a coherent DGP, the starting values for the factors and idiosyncratic components, f_0 and $\epsilon_{i,0}$, should come from the stationary distributions of f_t and $\epsilon_{i,t}$, respectively. By construction of the DGP, we have that $f_t \sim \mathcal{N}(0, 1)$ and $\epsilon_{i,t} \sim \mathcal{N}(0, \kappa_i)$, for all t. Hence, f_0 and $\epsilon_{i,0}$ should also be $\mathcal{N}(0, 1)$ and $\mathcal{N}(0, \kappa_i)$, respectively. For convenience, we generate $T + 1$ observations for periods $t = 0, 1, 2, \ldots, T$ and discard the starting values $x_{i,0}$.

The dimensions are $N = 5, 10, 25, 50, 100$ and $T = 50, 100$. The panel is unbalanced between the time points $T - 3$ and T by letting $x_{i,t}$ be available for $t = 1, 2, \ldots, T - j$ if $i \le (j + 1) \frac{N}{5}$. That is, at time $T - 3$, 80% of the data are available; at time $T - 2$, 60% are available; at time $T - 1$, 40% are available; and at time T, 20% are available. For each set of dimensions $\{N, T\}$, the parameters β_i and λ_i are drawn 50 times. Then, for each such draw, the error processes z_t and $\mathbf{u}_t$ are generated 50 times. Hence, the total number of replications is 2500. At each replication, we execute steps 1–3 from Sect. 4 using our subroutine. To evaluate the precision of the smoothed factors $\hat{\mathbf{g}}_t$, the following measure is used:

$$\Delta_t = \text{Trace}\left[(\mathbf{f}_t - \hat{\mathbf{Q}}\hat{\mathbf{g}}_t)(\mathbf{f}_t - \hat{\mathbf{Q}}'\hat{\mathbf{g}}_t)' \right], \tag{20}$$

where $\hat{\mathbf{Q}}$ is the estimated coefficient from an OLS regression of $\mathbf{f}_t$ on $\hat{\mathbf{g}}_t$, using the sample between $t = 1$ and $t = T - 4$, $\hat{\mathbf{Q}} = \sum_{t=1}^{T-4} \mathbf{f}_t \hat{\mathbf{g}}_t' \left(\sum_{t=1}^{T-4} \hat{\mathbf{g}}_t \hat{\mathbf{g}}_t' \right)^{-1} = (\hat{\mathbf{G}}'\hat{\mathbf{G}})^{-1}\hat{\mathbf{G}}'\mathbf{F}$, where $\hat{\mathbf{G}} = (\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \ldots, \hat{\mathbf{g}}_{T-4})'$ and $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{T-4})'$. A value closer to zero indicates higher squared error precision (see Doz et al. 2011). Because $\mathbf{f}_t$ is in this particular study a univariate process, we have that $\Delta_t = (f_t - \hat{Q}\hat{g}_t)^2$.

In the simulation study we run a large number of replications. This increases the probability that the OLS estimator of the VAR coefficients in (17) by chance may provide estimates that are inconsistent with a stationary VAR. That is to say, whereas we use a single factor with true parameter $b = 0.9$, the estimate of b may occasionally be 1 or larger. To avoid these cases (they are very rare), we assign such estimates with a value that is near the boundary of the stationary parameter space (a value close to 1, from below).

Table 2 shows a replication of the parts of Table 1 in Doz et al. (2011) that are related to models $\Omega^{\mathcal{A}4}$ and $\Omega^{\mathcal{A}3}$. Here, $\Delta_{T-s}^{\mathcal{A}4}$ and $\Delta_{T-s}^{\mathcal{A}3}$ denote the evaluation by (20) of a smoothed factor in period $T - s$ ($s = 0, 1, \ldots, 4$) from models $\Omega^{\mathcal{A}4}$ and $\Omega^{\mathcal{A}3}$, respectively, and $\bar{\Delta}_{T-s}^{\mathcal{A}4}$ and $\bar{\Delta}_{T-s}^{\mathcal{A}3}$ denote their averages over the 2500 replications. The panel is unbalanced for $s = 0, 1, 2, 3$. The upper part of Table 2 displays the averages $\bar{\Delta}_{T-s}^{\mathcal{A}4}$, and the lower part displays the ratios $\bar{\Delta}_{T-s}^{\mathcal{A}4}/\bar{\Delta}_{T-s}^{\mathcal{A}3}$ (a value below 1 indicates that $\Omega^{\mathcal{A}4}$ is, on average, more accurate). Our simulation results are close to the results in Doz et al. (2011). However, we note that we have somewhat higher precision for small N, while slightly less precision for large N. As expected, the average precision of $\Omega^{\mathcal{A}4}$ is uniformly better than the average precision of $\Omega^{\mathcal{A}3}$.

---

[9] The variance of an AR(1) process $v_t = \varphi v_{t-1} + \varepsilon_t$ is $Var(v_t) = (1 - \varphi^2)^{-1} Var(\varepsilon_t)$. Hence, $Var(f_t) = (1 - b^2)^{-1} Var(z_t) = 1$, and $Var(\epsilon_{i,t}) = (1 - \phi^2)^{-1} Var(u_{i,t}) = (1 - \phi^2)^{-1} \tau_{i,i} = \kappa_i$.

Table 2 Monte Carlo evaluation

	$T = 50$					$T = 100$				
s	$N = 5$	$N = 10$	$N = 25$	$N = 50$	$N = 100$	$N = 5$	$N = 10$	$N = 25$	$N = 50$	$N = 100$
Kalman smoother with cross-sectional heteroscedasticity: $\bar{\Delta}_{T-s}^{\mathcal{A}4}$										
4	0.33	0.32	0.32	0.34	0.33	0.20	0.19	0.17	0.18	0.18
3	0.33	0.32	0.32	0.34	0.33	0.20	0.18	0.17	0.18	0.18
2	0.35	0.33	0.32	0.34	0.33	0.21	0.19	0.17	0.18	0.18
1	0.34	0.33	0.32	0.33	0.33	0.22	0.19	0.18	0.18	0.18
0	0.37	0.34	0.32	0.34	0.33	0.25	0.20	0.18	0.19	0.18
Relative performance to Kalman smoother with cross-sectional homoscedasticity: $\bar{\Delta}_{T-s}^{\mathcal{A}4}/\bar{\Delta}_{T-s}^{\mathcal{A}3}$										
4	0.99	0.99	0.99	1.00	1.00	0.99	0.98	0.99	1.00	0.99
3	0.99	0.98	0.98	0.99	0.99	0.98	0.98	0.99	0.99	0.99
2	0.98	0.98	0.98	0.99	0.99	0.96	0.97	0.99	0.99	0.99
1	0.98	0.98	0.98	0.99	0.99	0.97	0.97	0.98	0.99	0.99
0	0.97	0.97	0.97	0.98	0.99	0.97	0.94	0.96	0.97	0.98

6 An Empirical Example: Nowcasting GDP Growth

The usefulness of our subroutine is illustrated in an empirical example where we nowcast Swedish GDP growth using a large number of economic indicators. All data are freely available for download, and are provided as supplementary material, together with code that replicates the empirical example.

A nowcast is a forecast of the very recent past, the present, or the near future. Nowcasting the state of the economy is important for economic analysts, because key statistics on economic activity, such as GDP, are published with a significant delay. For example, in the euro area, first official estimates of GDP are published 6–8 weeks after the reference quarter. Additionally, GDP is subject to substantial revisions as more source data become available. Meanwhile, a large number of indicators related to economic activity tend to be released well before official estimates of GDP are available, and typically at higher frequencies. For instance, whereas GDP is typically measured at quarterly frequency, data on industrial production (relating to the production side of GDP), personal consumption (relating to the expenditure side of GDP), unemployment, business survey data, and various financial data are available on monthly, weekly, or daily frequencies (for overviews, see, e.g., Banbura et al. 2011, 2012). As an illustration, we nowcast the Swedish GDP growth for the second quarter of 2017 (2017Q2), following in spirit the popular procedure by Giannone et al. (2008). We use data that were available in late August 2017, about 3 weeks before the Swedish national accounts were released on September 13.

Let y_t denote the quarterly GDP growth at time t, measured as percentage change from period $t - 1$, and let $x_{i,t}$ $(i = 1, 2, \ldots, N)$ denote the monthly indicators of economic activity outlined in Table 3. The number of indicators is large, $N = 124$. Hence, constructing models based directly on the indicators would quickly suffer from the curse of dimensionality, with limited degrees of freedom and large uncertainty

Table 3 Monthly time series for nowcasting GDP growth

Description	No.	Transformation	Publication lag	Producer	Downloaded from
GDP, calendar adjusted, seasonally adjusted, constant prices	1	Percentage change	Max 60 days	Statistics Sweden	www.scb.se
Economic tendency survey: choices of indicators and subsectors, seasonally adjusted, net balances	48	–	No lag	National Institute of Economic Research	www.konj.se
Export and import of goods and associated price indices	4	Percentage change	3–4 weeks	Statistics Sweden	www.scb.se
Industrial new orders, calendar adjusted, seasonally adjusted, constant prices	7	First-difference	5–6 weeks	Statistics Sweden	www.scb.se
Industrial production index and choices of subcategories, calendar adjusted, seasonally adjusted, constant prices	22	Percentage change	5–6 weeks	Statistics Sweden	www.scb.se
Index of service production and choices of subcategories, calendar adjusted, seasonally adjusted, constant prices	11	Percentage change	5–6 weeks	Statistics Sweden	www.scb.se
Household consumption indicator, seasonally adjusted, constant prices	1	Percentage change	5–6 weeks	Statistics Sweden	www.scb.se

Table 3 continued

Description	No.	Transformation	Publication lag	Producer	Downloaded from
Retail sales index, and and choices of subcategories, seasonally adjusted	21	Percentage change	3–4 weeks	Statistics Sweden & HUI Research	www.scb.se
New registrations of cars	1	Percentage change	2 days	Statistics Sweden & BIL Sweden	www.bilsweden.se
Producer price index	1	Percentage change	2–3 weeks	Statistics Sweden	www.scb.se
Consumer price index	1	Percentage change	2–3 weeks	Statistics Sweden	www.scb.se
Labor statistics: employment and unemployment, seasonally adjusted	2	Percentage change	2–3 weeks	Statistics Sweden	www.scb.se
Housing index, HOX	1	Percentage change	2–3 weeks	Nasdaq OMX Valueguard-KTH	www.valueguard.se
Effective exchange rate, KIX	1	First-difference	No lag	The Riksbank	www.riksbank.se
Short-term interest rates: repo rate and treasury bills	3	–	No lag	The Riksbank	www.riksbank.se

in parameter estimation. Therefore, a dynamic factor model is a natural choice of model, as an efficient dimension reducing technique. Furthermore, while collinearity is generally bad for conventional estimation methods, such as OLS, collinearity is rather preferred when extracting factors, since the goal for the extracted factors is to cover the bulk of variation in the elements of $\mathbf{x}_t$. As can be seen in Table 3, all 124 indicators would be available within 5–6 weeks after the last day of the reference quarter. This is 2–3 weeks before the first outcome of GDP is released. We could also choose to nowcast the GDP growth well before this date; either by smoothing missing data or by simply removing all data with later publication dates. A nowcast for y_t may be found by first estimating the monthly factors $\hat{f}_{j,t}$ for $j = 1, 2, \ldots, \mathcal{R}$, using the two-step estimator outlined in Sect. 4, and then by regressing the GDP growth on the quarterly (collapsed) factors,

$$\hat{y}_t = \hat{\alpha} + \sum_{j=1}^{\mathcal{R}} \hat{\beta}_j \, \hat{f}_{j,t}^Q, \tag{21}$$

where $\hat{\alpha}$ and $\hat{\beta}_j$ are estimated regression parameters, and $\hat{f}_{j,t}^Q$ are quarterly averages of the monthly factors $\hat{f}_{j,t}$.[10] We could also consider lagging the quarterly factors. However, as advocated by Giannone et al. (2008), the contemporary factors should capture the bulk of dynamic interaction among the indicators, and hopefully also the bulk of dynamics in GDP growth.[11]

We deflate export and import of goods using their respective price indices. The deflated series are then seasonally adjusted using the U.S. Census Bureau's x12-ARIMA procedure, accessible from EViews (see IHS Global Inc. 2015c, p. 416). Likewise, we seasonally adjust and replace all price indices and new registrations of cars. Lastly, before estimating the dynamic factor model, we make the transformations outlined in Table 3 to achieve stationarity. The series that we use from the Economic Tendency Survey are measured in net balances; they are therefore stationary (or approximately stationary) by construction (see, e.g., Hansson et al. 2005; Österholm 2014, for descriptions of the survey). Based on the (now) stationary monthly indicators, we estimate two monthly factors ($\mathcal{R} = 2$) that follow a VAR(2) over the sample 2002M01–2017M06. For seven of the monthly indicators, the start date is later than 2002M01. For these series, we let the Kalman filter smooth the missing values.

There are numerous outputs and tools available for analyzing a state space object in EViews; see van den Bossche (2011) and IHS Global Inc. (2015d, Chapter 39). Here, our main interest is in the estimated factors and their associated error bounds. As discussed in Sect. 3.1, EViews allows the user to create filtered states, one-step ahead

---

[10] Another approach is to let (a monthly) y_t be a part of the observed vector so that $\mathbf{x}_t = (y_t, x_{1,t}, \ldots, x_{N,t})'$ (see, e.g., Schumacher and Breitung 2008; Banbura and Runstler 2011). At point t, y_t is missing. By projection from the Kalman filter or the EM algorithm, the minimum mean-square error estimator of y_t can be achieved.

[11] Naturally, we could use the same principle to make a forecast of y_{t+h}, for some $h > 0$ (see, e.g., Stock and Watson 2002a, b). The main difference between forecasting and nowcasting is that, for nowcasting we exploit contemporary (rather than leading) indicators. For forecasting, Stock and Watson (2002b) suggest to use lags of both the factors and y_t as predictors.
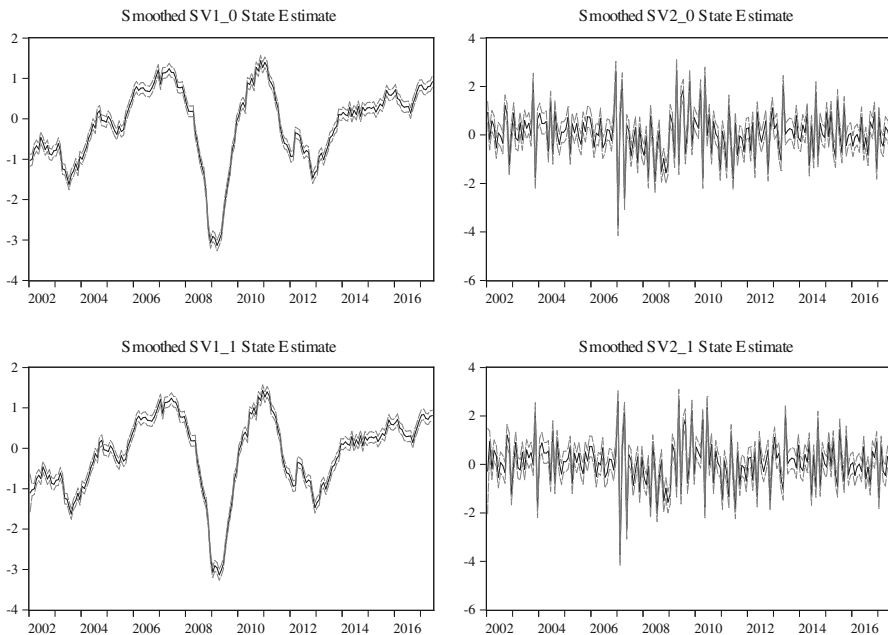
Fig. 1 Smoothed monthly state estimates (black solid lines) with error bounds (gray dashed lines)

predictions of states or smoothed states. In the subroutine, we use smoothed states (see Sect. 4.1). By default, EViews provides uncertainty bounds constructed from 2 standard errors (see Fig. 1). Recall from Sect. 4.1 that there are $\mathcal{R}p = 2 \times 2 = 4$ states, where the first two (SV1_0 and SV2_0) are our estimated factors, and the last two are their lags. In our case, the estimated uncertainty in terms of standard errors is somewhat larger for the second factor than for the first.

Given the estimated factors (that cover the nowcast period), the estimated regression (21) and its nowcast can easily be obtained in EViews using the interface menus (see IHS Global Inc. 2015c, Chapter 2), or by programming (see IHS Global Inc. 2015a, pp. 360–362). Carrying out the regression, our nowcast for the GDP growth 2017Q2 is 1.29 percent; the first published outcome by mid-September 2017 was 1.27 percent. Naturally, it would be desirable to have the associated error bounds for the nowcast. However, basing the error bounds on theory is a bit more cumbersome than usual, since the regressors in (21) (i.e., the factors) are estimated, and not observed series (see Bai and Ng 2006). EViews provides conventional error bounds based on a regression (see IHS Global Inc. 2015d, p. 144). In this particular case, they would be too narrow. Moreover, the error bounds are complicated by GDP revisions, where the most recent quarterly outcomes tend to also be revised, affecting the current growth rate. Figure 2 shows the quarterly collapsed factors and the GDP growth nowcast for 2017Q2 (including outcome up to 2017Q1). Judging from the figure, it is tempting to conclude that the second factor is partly lagging the first factor. However, its movement is rather a consequence of being (almost) orthogonal to the first factor.
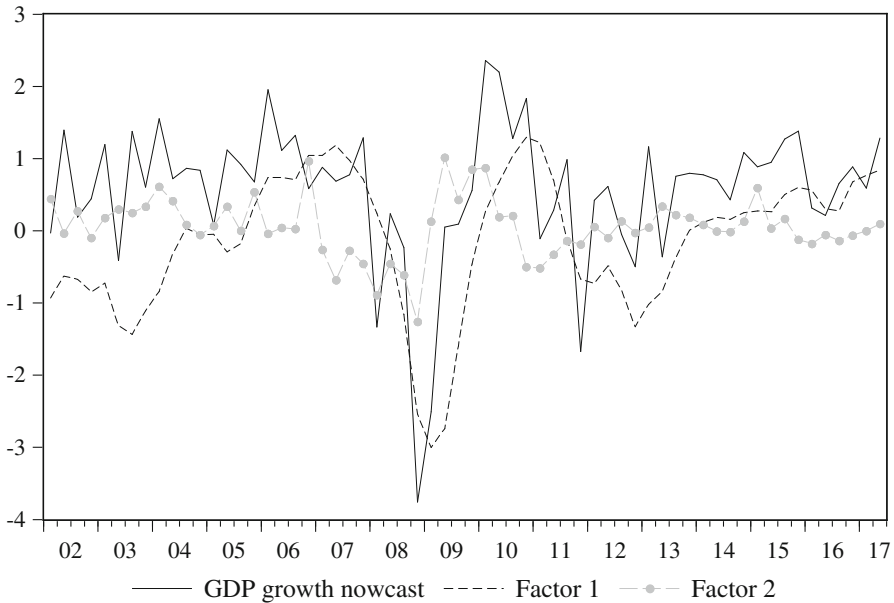
Fig. 2 GDP growth nowcast and quarterly factors

7 Concluding Remarks

We have demonstrated how to estimate a dynamic factor model by the Kalman filter and smoother in EViews and provided a global subroutine that can be useful to a broad range of economists or statisticians using large panel data to extract dynamic factors. Because the estimation requires only to compute principal components and then make one run with the Kalman smoother, the procedure is fast. Several extensions are possible. For instance, the code can be used to make an EViews Add-in. The code could also easily be altered to meet specific user needs. For example, a user may find it more convenient to work with a local subroutine. Or a user may wish to add estimation procedures for the number of factors and the number of lags in the factor autoregression. These modifications would require only basic knowledge in the EViews programming language.

References

Ahn, S. C., & Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, *81*(3), 1203–1227.

Anderson, T. W. (2003). *An introduction to multivariate statistical analysis*. New York: Wiley.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.

Bai, J., & Li, K. (2012). Statistical analysis of factor models of high dimension. *The Annals of Statistics*, *40*(1), 436–465.

Bai, J., & Li, K. (2016). Maximum likelihood estimation and inference for approximate factor models of high dimension. *The Review of Economics and Statistics*, *98*(2), 298–309.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221.

Bai, J., & Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, *74*(4), 1133–1150.

Bai, J., & Ng, S. (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics*, *25*(1), 52–60.

Bai, J., & Ng, S. (2008). Large dimensional factor analysis. *Foundations and Trends in Econometrics*, *3*(2), 89–163.

Bai, J., & Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, *176*(1), 18–29.

Bai, J., & Wang, P. (2014). Identification theory for high dimensional static and dynamic factor models. *Journal of Econometrics*, *178*(3), 794–804.

Banbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2012). Now-casting and the real-time data flow. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2A). Amsterdam: Elsevier-North Holland.

Banbura, M., Giannone, D., & Reichlin, L. (2011). Nowcasting. In M. Clements & D. Hendry (Eds.), *The Oxford handbook of economic forecasting*. Oxford: Oxford University Press.

Banbura, M., & Runstler, G. (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting*, *27*(2), 333–346.

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics*, *120*(1), 387–422.

Boivin, J., Giannoni, M. P., & Mihov, I. (2009). Sticky prices and monetary policy: Evidence from disaggregated us data. *The American Economic Review*, *99*(1), 350–384.

Breitung, J., & Eickmeier, S. (2006). Dynamic factor models. *Allgemeines Statistisches Archiv*, *90*(1), 27–42.

Chamberlain, G., & Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, *51*(5), 1281–1304.

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, *130*(2), 337–364.

Diebold, F. X., Rudebusch, G. D., & Aruoba, S. B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, *131*(1–2), 309–338.

Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, *164*(1), 180–205.

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics*, *94*(4), 1014–1024.

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (2nd ed.). Oxford: Oxford University Press.

Eickmeier, S. (2007). Business cycle transmission from the us to germany—A structural factor approach. *European Economic Review*, *51*(3), 521–551.

Forni, M., Giannone, D., Lippi, M., & Reichlin, L. (2009). Opening the black box: Structural factor models with large cross sections. *Econometric Theory*, *25*(5), 1319–1347.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2004). The generalized dynamic factor model consistency and rates. *Journal of Econometrics*, *119*(2), 231–255.

Forni, M., & Reichlin, L. (1998). Let's get real: A factor analytical approach to disaggregated business cycle dynamics. *The Review of Economic Studies*, *65*(3), 453–473.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, *55*(4), 665–676.

Hamilton, J. D. (1994). *Time series analysis*. New Jersey: Princeton University Press.

Hansson, J., Jansson, P., & Löf, M. (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting*, *21*(2), 377–389.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.

IHS Global Inc. (2015a). *EViews 9 command and programming reference*. Irvine: IHS Global Inc.

IHS Global Inc. (2015b). *EViews 9 object reference*. Irvine: IHS Global Inc.

IHS Global Inc. (2015c). *EViews 9 user's guide I*. Irvine: IHS Global Inc.

IHS Global Inc. (2015d). *EViews 9 user's guide II*. Irvine: IHS Global Inc.

Koopman, S. J., Shephard, N., & Doornik, J. A. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *The Econometrics Journal*, *2*(1), 107–160.

Ludvigson, S. C., & Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics*, *83*(1), 171–222.

Lütkepohl, H. (2007). *New introduction to multiple time series analysis*. Berlin: Springer.

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics*, *92*(4), 1004–1016.

Österholm, P. (2014). Survey data and short-term forecasts of swedish GDP growth. *Applied Economics Letters*, *21*(2), 135–139.

Phillips, P. C. B., & Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, *67*(5), 1057–1111.

Ritschl, A., Sarferaz, S., & Uebele, M. (2016). The U.S. business cycle, 1867–2006: A dynamic factor approach. *The Review of Economics and Statistics*, *98*(1), 159–172.

Schumacher, C., & Breitung, J. (2008). Real-time forecasting of german GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, *24*(3), 386–398.

Stock, J. H., & Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179.

Stock, J. H., & Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, *20*(2), 147–162.

Stoica, P., & Jansson, M. (2009). On maximum likelihood estimation in factor analysis—An algebraic derivation. *Signal Processing*, *89*(6), 1260–1262.

van den Bossche, F. A. M. (2011). Fitting state space models with EViews. *Journal of Statistical Software*, *41*(8), 1–16.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, *50*(1), 1–25.

Zhou, X., & Solberger, M. (2017). A lagrange multiplier-type test for idiosyncratic unit roots in the exact factor model. *Journal of Time Series Analysis*, *38*(1), 22–50.