Time-Varying Vector Autoregressive Models with Structural Dynamic Factors[1]

Paolo Gorgi[(a)] *Siem Jan Koopman*[(a,b)] *Julia Schaumburg*[(a)]

[(a)] Vrije Universiteit Amsterdam and Tinbergen Institute, The Netherlands

[(b)] CREATES, Aarhus University, Denmark

September 27, 2017

Abstract

We develop a transparent methodology for the estimation of time-varying parameters in vector autoregressive models. In contrast to the widely used Bayesian approaches, we base our analysis on a combination of time-varying autoregressive coefficient matrices depending on a flexible set of stochastic dynamic factors, and of time-varying variance matrices depending on score-driven factors. The resulting method for estimating static parameters and extracting the different factors is insightful, robust and computationally fast, while being easy to implement. In a simulation study, we demonstrate the good performance of the method. We further show that our approach is promising in the empirical modelling of time-varying macro-financial linkages using a data set of U.S. macroeconomic and financial variables.

Keywords: Dynamic factor model, Kalman filter, Score-driven models, Spill-over effects.

[1]Email addresses: p.gorgi@vu.nl (Paolo Gorgi), s.j.koopman@vu.nl (Siem Jan Koopman), j.schaumburg@vu.nl (Julia Schaumburg). Koopman acknowledges support from CREATES, the Center for Research in Econometric Analysis of Time Series (DNRF78) at Aarhus University, Denmark, funded by the Danish National Research Foundation. Schaumburg thanks the Dutch Science Foundation (NWO, grant VENI451-15-022) for financial support.

1 Introduction

Multivariate macroeconomic time series data are often observed over long time horizons, but at low frequencies. Methods based on the vector autoregression (VAR) model are probably the most widely used tools for the modelling and forecasting of aggregate macroeconomic variables. In most practical applications, the VAR model does not include more than five or six variables because otherwise the number of parameters becomes too large. The endowment of coefficient matrices in a macro VAR model with time-varying processes is even more challenging. It typically requires methods to reduce the dimensionality. Nevertheless, these time-varying extensions often deliver important insights. For example, a topical example is the study towards the dynamic, and potentially diverse, spillover channels from financial markets to the real economy.

This paper suggests a simple, transparent methodology to estimate time-varying parameter VAR models. We use an intuitive dynamic factor structure for the time-varying coefficient matrices. The covariance matrix of the error terms follows flexible score-driven dynamics. Our approach is in the spirit of the adaptive state space models of Delle Monache et al. (2016)

In the literature on time-varying coefficient VARs, often Bayesian approaches are employed, see for instance, Primiceri (2005), Canova and Ciccarelli (2009) and Canova and Ciccarelli (2004). In a recent article, Prieto et al. (2016) use the Bayesian time-varying parameter approach of Primiceri (2005) to model the dynamic interactions of two macro and four financial time series. They find that shocks to the financial variables and their transmission to the real economy, measured by GDP growth, are amplified during crisis periods. We address a similar research question as they do, but with a completely different, frequentist methodology.

[[[An extended literature review will be inserted here.]]]

The paper is structured as follows. The time-varying parameter model and our estimation approach are decribed in section 2. Section 3 presents a small Monte Carlo study, in which we investigate whether our method can filter out different patterns in the dynamic coefficient matrices. In section 4, the method is applied to the macro-financial data set of Prieto et al. (2016). Section 5 concludes.

2 The Time-Varying VAR model with dynamic factors

We denote the N-dimensional vector of time series variables by y_t for time index $t = 1, \ldots, T$ where T is the time series length. We define the vector autoregressive (VAR) model of order p, referred to as the VAR(p) model, by $y_t = \Phi_{[1]} y_{t-1} + \ldots + \Phi_{[p]} y_{t-p} + \varepsilon_t$, for $t = p+1, \ldots, T$, where $\Phi_{[i]}$ is the $N \times N$ autoregressive coefficient matrix, for $i = 1, \ldots, p$, and ε_t is the N-dimensional vector of disturbances that is identically and independently distributed as a multivariate normal with mean zero and variance matrix H, we can denote this by $\varepsilon_t \sim NID(0, H)$. Parameter estimation and impulse response analysis are widely explored in textbooks such as Hamilton (1994) and Lutkepohl (2005). The VAR(p) model can be efficiently formulated as $y_t = \Phi Y_{t-1:p} + \varepsilon_t$ where $\Phi = \left[\Phi_{[1]}, \ldots, \Phi_{[p]}\right]$ is the $N \times Np$ coefficient matrix and $Y_{t-1:p} = \left(y'_{t-1}, \ldots, y'_{t-p}\right)'$ is the $Np \times 1$ vector of lagged dependent variables, for $t = p+1, \ldots, T$. Throughout the discussions and developments below, we assume that the initial observation set $\{y_1, \ldots, y_p\}$ is fixed and given.

The VAR(p) model with time-varying parameters is then given by

$$y_t = \Phi_t Y_{t-1:p} + \varepsilon_t, \qquad \varepsilon_t \sim NID(0, H_t), \qquad t = p+1, \ldots, T, \tag{1}$$

where Φ_t is the $N \times Np$ time-varying matrix of autoregressive coefficients and H_t is the $N \times N$ time-varying variance matrix. The two time-varying matrices are modelled separately and very differently in nature. We let Φ_t be dependent on a set of dynamic factors which are specified as stochastic processes. We show that the resulting model can be formulated as a linear Gaussian state space model. Hence this part of the analysis can be based on the Kalman filter and related methods, see Harvey (1989) and Durbin and Koopman (2012). The specification of the time-varying variance matrix H_t is based on the score-driven approach as introduced in Creal et al. (2013) and Harvey (2013). In this approach, the variance matrix H_t is implicitly defined as a nonlinear function of past observations $\{y_{p+1}, \ldots, y_{t-1}\}$ through the score function of the predictive loglikelihood function of y_t, with respect to H_t. We have opted for these different specifications of time-varying matrices out of convenience; it facilitates the relatively easy and straightforward implementation of methods for estimation, analysis and forecasting. This combination is explored in more generality by Delle Monache et al. (2016) who refer to this approach as *adaptive state space models*. We follow their approach, but consider direct updating of the variance matrix, show the specific details for our model specification and show that the required methods are basic and easy to implement.

2.1 Time-varying autoregressive coefficient matrix

We specify Φ_t in (1) as the $N \times Np$ matrix function $\Phi(\cdot)$ with its argument being the $r \times 1$ vector of stochastic dynamic factors f_t. We adopt a general linear specification for Φ_t and is given by

$$\Phi_t = \Phi(f_t) = \Phi^c + \Phi_1^f f_{t,1} + \cdots + \Phi_r^f f_{t,r} \tag{2}$$

where Φ^c and Φ_i^f are $n \times Np$ coefficient matrices, for $i = 1, \ldots, r$, and $f_{t,i}$ is the ith element of vector f_t, for $i = 1, \ldots, r$, that is $f_t = (f_{t,1} \ldots, f_{t,r})'$. The coefficient matrices are subject to restrictions such that possible unknown elements in these matrices can be treated as parameters that are identified. The restrictions can be formulated on case-by-case basis. The stochastic dynamic factors can be specified as stationary autoregressive processes. For example, we can have

$$f_{t+1} = \varphi f_t + \eta_t, \qquad \eta_t \sim NID(0, \Sigma_\eta), \qquad t = p, \ldots, T, \tag{3}$$

with $r \times r$ autoregressive coefficient matrix φ and $r \times 1$ disturbance vector η_t that is assumed normally distributed with mean zero and $r \times r$ variance matrix Σ_η. To keep the model parsimonious but also for identification purposes, we assume that the dynamic factors are independent and are standardised (unconditional mean is zero vector and variance is identity matrix) by imposing $\Sigma_\eta = \mathrm{I} - \varphi\varphi'$. All elements of the disturbance vectors ε_t in (1) and η_t in (3) are serially and mutually uncorrelated, at all leads and lags, that is $\mathbb{E}(\varepsilon_t \, \eta_s') = 0$ for all $t, s = p+1, \ldots, T$. It is implied that the initial condition for f_p is given by its unconditional properties, that is $f_p \sim N(0, \mathrm{I})$. The generality of this specification and additional generalisations for the dynamic specifications of f_t are discussed in Appendix A.

When we assume that the sequence of variance matrices $H_{p+1}, \ldots, H_T$ is known and fixed, we can represent the model (1), (2) and (3) in space space form. In order to obtain the state space form of the model, we define $\tilde{y}_t = y_t - \Phi^c Y_{t-1:p}$ and consider the following equation equalities

$$\begin{aligned}
\tilde{y}_t &= \left[\Phi_1^f f_{t,1} + \ldots + \Phi_r^f f_{t,r}\right] Y_{t-1:p} + \varepsilon_t \\
&= \left[\Phi_1^f, \cdots, \Phi_r^f\right] (f_t \otimes \mathrm{I}_{Np}) \, Y_{t-1:p} + \varepsilon_t \\
&= \left(Y_{t-1:p}' \otimes \left[\Phi_1^f, \cdots, \Phi_r^f\right]\right) \mathrm{vec}\,(f_t \otimes \mathrm{I}_{Np}) + \varepsilon_t \\
&= \left(Y_{t-1:p}' \otimes \left[\Phi_1^f, \cdots, \Phi_r^f\right]\right) Q \, f_t + \varepsilon_t,
\end{aligned}$$

where I_q is the $q \times q$ unity matrix and where the $N^2p^2r \times r$ matrix Q consists of zero and unity values such that $Q\, f_t = \mathrm{vec}\,(f_t \otimes \mathrm{I}_{Np})$. We let

$$Z_t = \left(Y'_{t-1:p} \otimes \left[\Phi_1^f, \cdots, \Phi_r^f\right]\right) Q,$$

to obtain the linear Gaussian state space form

$$\tilde{y}_t = Z_t f_t + \varepsilon_t, \qquad f_{t+1} = \varphi f_t + \eta_t, \tag{4}$$

where the properties of the disturbances ε_t and η_t are discussed above. The Kalman filter and related methods can be applied towards the state space model (4); a textbook treatment on state space methods is provided by Durbin and Koopman (2012).

The vector of prediction errors is defined as $v_t = y_t - \mathbb{E}(y_t|\mathcal{F}_{t-1};\psi)$ where $\mathcal{F}_t$ is the set of all information that is contained from the past upto time t, including all observations, and where ψ is the parameter vector that collects all unknown coefficients in the autoregressive and variance matrices including Φ^c, Φ_1^f, ..., Φ_r^f, H_{p+1}, ..., H_T, φ and Σ_η. Under correct model specification, the prediction error series $v_{p+1}, \ldots, v_T$ is serially uncorrelated. The variance matrix of the prediction error is defined as $F_t = \mathbb{V}\mathrm{ar}(v_t|\mathcal{F}_{t-1};\psi) = \mathbb{V}\mathrm{ar}(v_t;\psi)$. For a given vector ψ, the Kalman filter can numerically be implemented and the prediction error and its variance are computed by

$$\begin{array}{rclrclrcl} v_t &=& \tilde{y}_t - Z_t a_t, & F_t &=& Z_t P_t Z_t' + H_t, & K_t &=& \varphi P_t Z_t' F_t^{-1}, \\ a_{t+1} &=& \varphi a_t + K_t v_t, & P_{t+1} &=& \varphi P_t(\varphi - K_t Z_t)' + \Sigma_\eta, & & & \end{array} \tag{5}$$

where the prediction of the state vector is given by $a_t = \mathbb{E}(f_t|\mathcal{F}_{t-1};\psi)$ with variance matrix $P_t = \mathbb{V}\mathrm{ar}(f_t - a_t|\mathcal{F}_{t-1};\psi)$, for $t = p+1, \ldots, T$. The weighting matrix K_t is referred to as the Kalman gain. The Kalman filter is instrumental for the computation of the loglikelihood function of the state space model as given by

$$\ell(\psi) = \sum_{t=p+1}^{T} \ell_t(\psi), \qquad \ell_t(\psi) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\log|F_t| - \frac{1}{2} v_t' F_t^{-1} v_t. \tag{6}$$

The maximum likelihood estimation of ψ reduces to the numerical maximisation of the likelihood $\ell(\psi)$ with respect to ψ.

The parameter vector ψ can easily have a large dimension, even for a moderate value of N. A practical and realistic model will treat matrix Φ^c as a full square matrix with unknown coefficients that all need to be estimated. The unconditional expectation of Φ_t equals Φ^c in this specification for f_t in (3), that is $\mathbb{E}(\Phi_t) = \Phi^c$ since $\mathbb{E}(f_t) = 0$, for all t. The matrix Φ^c is then clearly related, "in the long-run", to Φ in the static VAR model as introduced in the beginning of this section. The maximum likelihood estimation of Φ in a static VAR is simply obtained by an equation-by-equation regression analysis. In this stationary framework, it can be argued that the static VAR regression estimator of Φ is also a consistent estimator of Φ^c in a time-varying VAR model as discussed above. Hence an estimate of Φ^c is available and maximum likelihood estimation only needs to apply to the remaining unknown coefficients in the (sparse, selection) matrices Φ_i^f, for $i = 1, \ldots, r$, and the $r \times r$ autoregressive coefficient matrix φ which we typically assume to be a diagonal matrix.

2.2 Time-varying variance matrix

Each Kalman filter step at time t requires a value for H_t. We use the score-driven approach to let the variance matrix H_t to change recursively over time. We define the $N^* \times 1$ vector $f_t^\sigma = \mathrm{vech}(H_t)$ with $N^* = N(N+1)/2$. Its dynamic specification is given by

$$f_{t+1}^\sigma = \omega + B\, f_t^\sigma + A\, s_t, \tag{7}$$

where ω is a constant vector, A and B are square coefficient matrices and s_t is the innovation vector that let f_t^σ change over time. The unknown coefficients in ω, A and B are placed in the parameter vector ψ. The distinguishing feature of the score-driven model is the definition of the innovation vector s_t as the scaled score vector of the predictive loglikelihood contribution at time t, that is $\ell_t \equiv \ell_t(\psi)$ in (6), with respect to f_t^σ. In particular, we have $s_t = S_t\, \nabla_t$ where S_t is the scaling matrix and ∇_t is the gradient vector. We first develop an expression for ∇_t and then an expression for the scaling matrix S_t. The transpose of the gradient vector is given by

$$\nabla_t' = \frac{\partial \ell_t}{\partial f_t^{\sigma\prime}} = \frac{\partial \ell_t}{\partial \mathrm{vec}(F_t)'} \cdot \frac{\partial \mathrm{vec}(F_t)}{\partial \mathrm{vech}(H_t)'} = \frac{\partial \ell_t}{\partial \mathrm{vec}(F_t)'} \cdot \frac{\partial \mathrm{vec}(H_t)}{\partial \mathrm{vech}(H_t)'}, \tag{8}$$

where the last equality holds since we have $F_t = Z_t P_t Z_t' + H_t$ for which Z_t does not depend on H_t and P_t is a function of $H_{p+1}, \ldots, H_{t-1}$, but not H_t. The two terms of (8) can be expressed as

$$\frac{\partial \ell_t}{\partial \mathrm{vec}(F_t)'} = \frac{1}{2}\left[\mathrm{vec}(v_t\, v_t')' - (\mathrm{vec}(F_t))'\right]\left(F_t^{-1} \otimes F_t^{-1}\right), \qquad \frac{\partial \mathrm{vec}(H_t)}{\partial \mathrm{vech}(H_t)'} = D_N,$$

where D_N is the $N^2 \times N^*$ duplication matrix, see Magnus and Neudecker (2007) for this result and for the definition of the duplication matrix. It follows that

$$\nabla_t = \frac{1}{2} D_N' \left(F_t^{-1} \otimes F_t^{-1}\right)\left(\mathrm{vec}(v_t\, v_t') - \mathrm{vec}(F_t)\right). \tag{9}$$

For many score-driven models, the inverse of the information matrix is taken as the scaling matrix S_t for the gradient vector. We derive an expression of the information matrix by the following steps,

$$\begin{aligned}
\mathcal{I}_t &= \mathbb{E}[\nabla_t\, \nabla_t' | \mathcal{F}_{t-1}] \\
&= \frac{1}{4} D_N' \left(F_t^{-1} \otimes F_t^{-1}\right) \mathbb{V}\mathrm{ar}\left[\mathrm{vec}(v_t\, v_t') - \mathrm{vec}(F_t) | \mathcal{F}_{t-1}\right]\left(F_t^{-1} \otimes F_t^{-1}\right) \\
&= \frac{1}{4} D_N' \left(F_t^{-1} \otimes F_t^{-1}\right)(\mathrm{I}_{N^2} + C_N) D_N \\
&= \frac{1}{2} D_N' \left(F_t^{-1} \otimes F_t^{-1}\right) D_N,
\end{aligned}$$

where we have exploited the properties $\mathbb{V}\mathrm{ar}[\mathrm{vec}(v_t\, v_t') - \mathrm{vec}(F_t)|\mathcal{F}_{t-1}] = (\mathrm{I}_{N^2} + C_N)\,(F_t \otimes F_t)$ and $(\mathrm{I}_{N^2} + C_N) D_N = 2\, D_N$, where C_N is the $N^2 \times N^2$ commutation matrix; see Chapter 3 in Magnus and Neudecker (2007) for definition and discussions. The inverse of the information matrix is used for the scaling of the score function and is given by

$$\mathcal{I}_t^{-1} = 2 D_n^{+} (F_t \otimes F_t) D_N^{+\prime}, \tag{10}$$

where $D_N^{+} = (D_N'\, D_N)^{-1} D_N'$ is the elimination matrix for symmetric matrices, see Magnus and Neudecker (1980). We set the scaling as $S_t = \mathcal{I}_t^{-1}$ and hence the scaled score $s_t = \mathcal{I}_t^{-1} \nabla_t$ becomes

$$\begin{aligned}
s_t &= D_N^{+} (F_t \otimes F_t) D_N^{+} D_N (F_t^{-1} \otimes F_t^{-1}) [\mathrm{vec}(v_t\, v_t') - \mathrm{vec}(F_t)] \\
&= D_N^{+} \left[\mathrm{vec}(v_t\, v_t') - \mathrm{vec}(F_t)\right] \\
&= \mathrm{vech}(v_t\, v_t') - \mathrm{vech}(F_t).
\end{aligned}$$

The derivation above relies on the same arguments as those used in Theorem 13 of Magnus and Neudecker (2007). For the score-driven update of the variance factors in f_t^σ of (7), we obtain

$$f_{t+1}^\sigma = \omega + A \, [\mathrm{vech}(v_t \, v_t') - \mathrm{vech}(F_t)] + B \, f_t^\sigma, \tag{11}$$

for $t = p+1, \ldots, T$.

The score updating function (11) can easily be incorporated in the Kalman filter (5) as follows:

$$\begin{array}{rclrclrcl}
v_t &=& \tilde{y}_t - Z_t a_t, & F_t &=& Z_t P_t Z_t' + H_t, & K_t &=& \varphi P_t Z_t' F_t^{-1}, \\
a_{t+1} &=& \varphi a_t + K_t v_t, & P_{t+1} &=& \varphi P_t (\varphi - K_t Z_t)' + \Sigma_\eta, & U_t &=& v_t \, v_t' - F_t, \\
f_{t+1}^\sigma &=& \omega + A \mathrm{vech}(U_t) + B f_t^\sigma, & H_{t+1} &=& \mathrm{unvech}(f_{t+1}^\sigma), & & &
\end{array}$$

for $t = p+1, \ldots, T$. The loglikelihood function (6) for a given parameter vector ψ can be computed in a similar way via equation (6). The additional unknown parameters in this development are vector ω and matrices A and B. In a stationary setting, we have $\mathbb{E}(f_t^\sigma) = (I - B)^{-1}\omega$ since $\mathbb{E}(U_t) = 0$. This "long-run" variance can be associated with the variance matrix of the static VAR as discussed at the beginning of this section. We may argue that the variance estimate of H, as obtained from the equation-by-equation regression computations, is the consistent estimate of $(I - B)^{-1}\omega$, after vectorisation vech. Hence, we have $\widehat{\omega} = (I - B)\widehat{H}$ where $\widehat{H}$ is the regression residual variance estimate of H in the static VAR model. Then the maximum likelihood estimation only needs to be extended for the parameters in the (sparse) matrices A and B.

2.3 Direct updating of time-varying variance matrix

For the purpose of obtaining a direct updating equation for the variance matrix H_t, we have that $\mathrm{vec}(H_t) = D_N \, \mathrm{vech}(H_t) = D_N \, f_t^\sigma$ and we obtain

$$\mathrm{vec}(H_{t+1}) = D_N \, \omega + D_N \, A \, D_N^+ [\mathrm{vec}(v_t \, v_t') - \mathrm{vec}(F_t)] + D_N \, B \, D_N^+ \mathrm{vec}(H_t),$$

for $t = p+1, \ldots, T$. When we specify the matrices A and B in (7) such that $D_N \, A \, D_N^+ = A^* \otimes A^*$ and $D_N \, B \, D_N^+ = B^* \otimes B^*$, we have

$$H_{t+1} = \Omega + A^* \, (v_t \, v_t' - F_t) \, A^{*\prime} + B^* \, H_t B^{*\prime}, \tag{12}$$

where Ω is a symmetric matrix such that $\text{vech}(\Omega) = \omega$. In case, the matrices A and B are diagonal, the matrices A^* and B^* are also diagonal with their diagonal elements equal to the square root values of the corresponding diagonal values of A and B, respectively. In case $A = a \cdot \mathrm{I}_{N^*}$ and $B = b \cdot \mathrm{I}_{N^*}$, the updating reduces simply to

$$H_{t+1} = \Omega + a\ (v_t\, v_t' - F_t) + b\, H_t.$$

This time-varying variance matrix updating equation can even more conveniently be incorporated within the Kalman filter. Assume that the parameter vector ψ is given, such that the coefficient matrices Φ^c, Φ_i^f, φ, Ω, A, and B are known, for $i = 1, \ldots, p$, we are then able to simply add the updating for H_t into the Kalman filter equations of (5), that is

$$\begin{aligned} v_t &= \tilde{y}_t - Z_t a_t, & F_t &= Z_t P_t Z_t' + H_t, \\ & & K_t &= \varphi P_t Z_t' F_t^{-1}, \\ a_{t+1} &= \varphi a_t + K_t v_t, & P_{t+1} &= \varphi P_t (\varphi - K_t Z_t)' + \Sigma_\eta, \\ & & H_{t+1} &= \Omega + A^*\,(v_t\, v_t' - F_t)\, A^{*\prime} + B^*\, H_t B^{*\prime}, \end{aligned} \tag{13}$$

On the basis of this Kalman filter, the loglikelihood function $\ell(\psi)$ can be computed as in (6) and the unknown parameter vector ψ can be estimated via the numerical maximisation of $\ell(\psi)$ with respect to ψ. Furthermore, from the predicted estimates a_t, we obtain the estimated paths for the time-varying Φ_t in (2) while simultaneously we obtain the time-varying variances H_t.

3 Monte Carlo study

To verify whether the proposed methodology is capable to identify the underlying dynamic processes and the unknown parameters in finite samples, we have designed and carried out the following Monte Carlo study. We simulate a vector time series from a VAR(p) model with $p = 1$, a time-varying coefficient matrix Φ_t based on a single factor f_t, and a time-varying variance matrix H_t. The data generation process is given by

$$y_t = \Phi_t y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim NID(0, H_t), \qquad t = 1, \ldots, T,$$

with

$$\Phi_t = \Phi^f f_t, \qquad f_{t+1} = \varphi f_t + \eta_t, \qquad \eta_t \sim NID(0, 1-\varphi^2).$$

with the initial observation set to $y_0 = 0$, the initial factor has value $f_1 = 0.95$ and with the parameter values set to $\Phi^f_{ii} = 1$, $\Phi^f_{i,j} = 0$, $\varphi = 0.95$, for $i, j = 1, \ldots, N$ and $i \neq j$. The time-varying variance matrix is set equal to $H_t = H f^\sigma_t$ where the diagonal values of H are set to 1, the off-diagonal values of H are set to 0.1, and the time-varying f^σ_t is some deterministic pattern. We consider the following two possible patterns for f^σ_t:

$$\begin{aligned} &\text{the sine function:} \qquad && f^\sigma_t = 1 + 0.95\cos(2\pi t/150), \\ &\text{the step function:} \qquad && f^\sigma_t = 1.5 - I(t > T/2), \end{aligned}$$

for $t = 1, \ldots, T$, and where $I(\cdot)$ denotes the indicator function. The sample sizes in the simulation study are given by $N = 5, 7$ and $T = 250, 500$. In Figure 1 we present two examples of simulations of data sets with $N = 5$ and $T = 500$, one for the sine and one for step variance function. The impact of the variance factors is clearly visible: we can observe the two distinctive patterns of heteroskedasticity in the time series. Furthermore, the common factor in the autoregressive coefficients leads to co-movements in the time series.

Given a simulated time series vector, we apply our methodology in a straightforward manner. We use the amended Kalman filter equations (13) for the computation of the loglikelihood function (6); the maximum likelihood estimates for φ and Φ^f are obtained by its numerical maximisation. The final step is to construct the time-varying Φ_t and H_t with ψ replaced by its maximum likelihood estimate. In the Monte Carlo study we repeat this process 200 times and record all parameter estimates and extracted paths for Φ_t and H_t. We compute the mean squared error (MSE) for the 200 individual parameter estimates and the 200 paths of the time-varying elements of Φ_t and H_t. In case of the parameters, we consider the difference between the true parameter and its estimates. In case of the time-varying elements, we consider the difference between the true path and the filtered estimates, and the MSE over the time dimension. The results are presented in Table 1.

In Figure 2 we present the true processes for f_t and f^σ_t together with the median of their corresponding 200 extracted estimated paths, also with the 5% and 95% quantiles. It reflects the capabilities of our procedure to capture the underlying dynamics, both in the serial dependence as well as the variance (or volatility) factors. We believe that the procedure is rather successful in identifying the

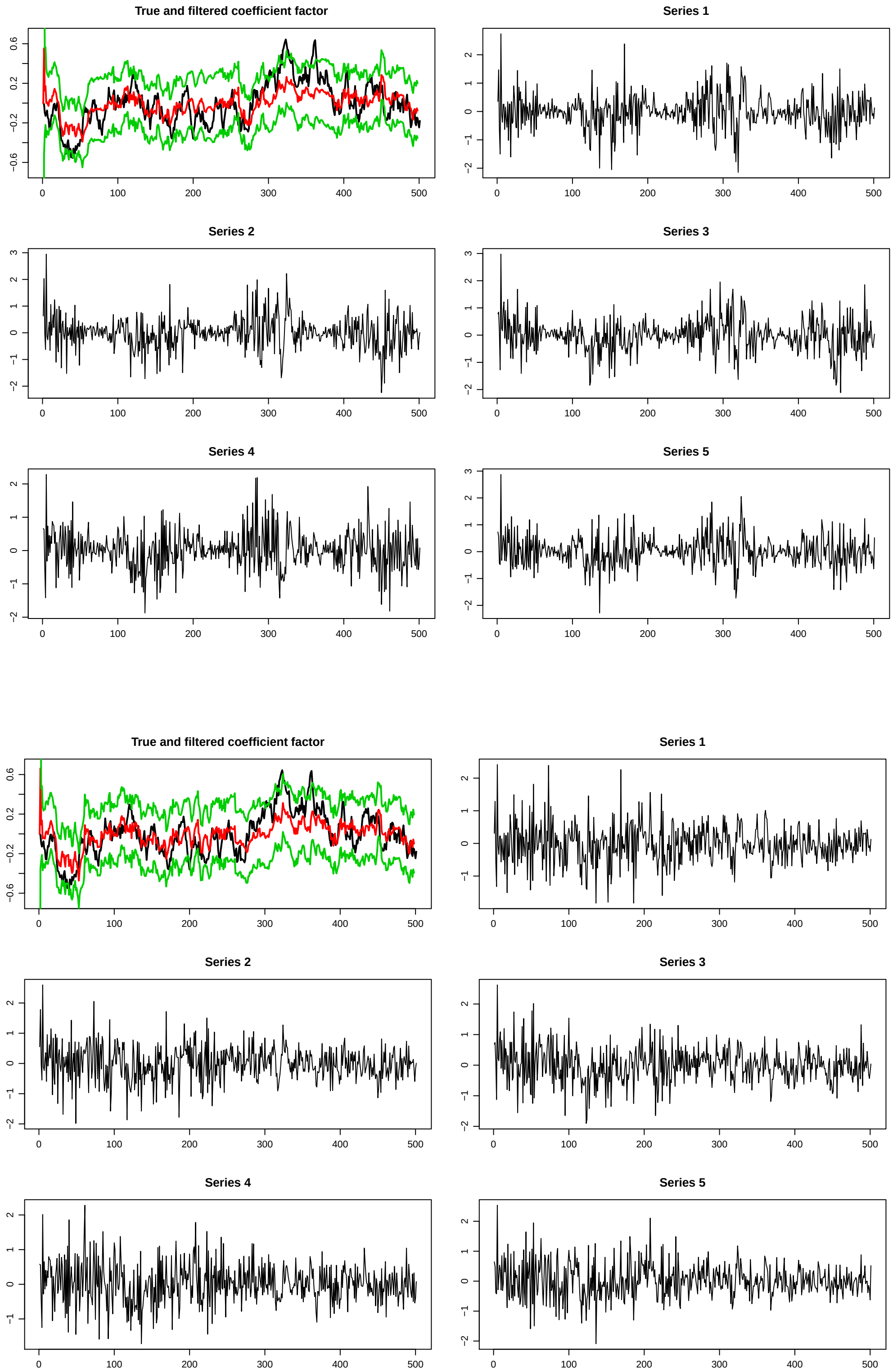

Figure 1: Two examples of simulated data sets with $N = 5$ from sine variance pattern (upper panel) and step variance pattern (lower panel)

Table 1: Simulation results

Mean squared errors for individual parameter estimates and time-varying factor estimates for f_t, over the 200 Monte Carlo replications. The cases of $N = 5, 7$ and $T = 250, 500$ are considered. Further, we consider two deterministic patterns of a time-varying variance: sine function and step function.

variance pattern "sine"				
	N=5		N=7	
	T=250	T=500	T=250	T=500
φ	0.01177	9e-04	0.00991	0.00048
Φ^f	0.41609	0.15935	0.33568	0.18428
f_t	0.06768	0.06491	0.06332	0.06485
variance pattern "step"				
	N=5		N=7	
	T=250	T=500	T=250	T=500
φ	0.01122	0.00076	0.00828	0.00063
Φ^f	0.24277	0.06938	0.20426	0.07465
f_t	0.07205	0.06799	0.06396	0.06871

dynamic structures accurately in a small-sample setting.

4 Empirical illustration: Dynamic macro-financial linkages

We estimate a six-dimensional time-varying parameter VAR. The data set is the same as used in Prieto et al. (2016).[1] It contains two macroeconomic variables, nominal GDP growth and inflation (the GDP deflator), as well as four financial variables: real house price inflation, the Baa-Aaa spread of corporate bonds, real stock price inflation, and the federal funds rate as monetary policy variable. The sample ranges from 1958Q1 until 2012Q2. Plots of the data are shown in Figure 3.

Our most general empirical specification is a VAR(2) model with two factors for the coefficient dynamics and a time-varying covariance matrix:

$$
\begin{aligned}
y_t &= \Phi_{1t} y_{t-1} + \Phi_{2t} y_{t-2} + \varepsilon_t \quad \varepsilon_t \sim N(0, H_t), && (14)\\
\Phi_{jt} &= \Phi_j^c + \Phi_{j,1}^f f_{t,1} + \Phi_{j,2}^f f_{t,2}, \quad j = 1, 2, &&\\
f_{t+1,j} &= \varphi_j f_{t,j} + \eta_{t,j} \quad \eta_{t,j} \sim N(0, 1 - \varphi_j^2), \quad j = 1, 2 && (15)
\end{aligned}
$$

where we assume that Φ_1^c and Φ_2^c are full matrices, $\Phi_{1,1}^f$ and $\Phi_{2,1}^f$ are diagonal matrices and $\Phi_{1,2}^f$ and $\Phi_{2,2}^f$ have zero entries except for the four coefficients that measure the impact of the financial

[1] We obtain their data from the website of the Journal of Applied Econometrics; see the online appendix to Prieto et al. (2016) for a description.

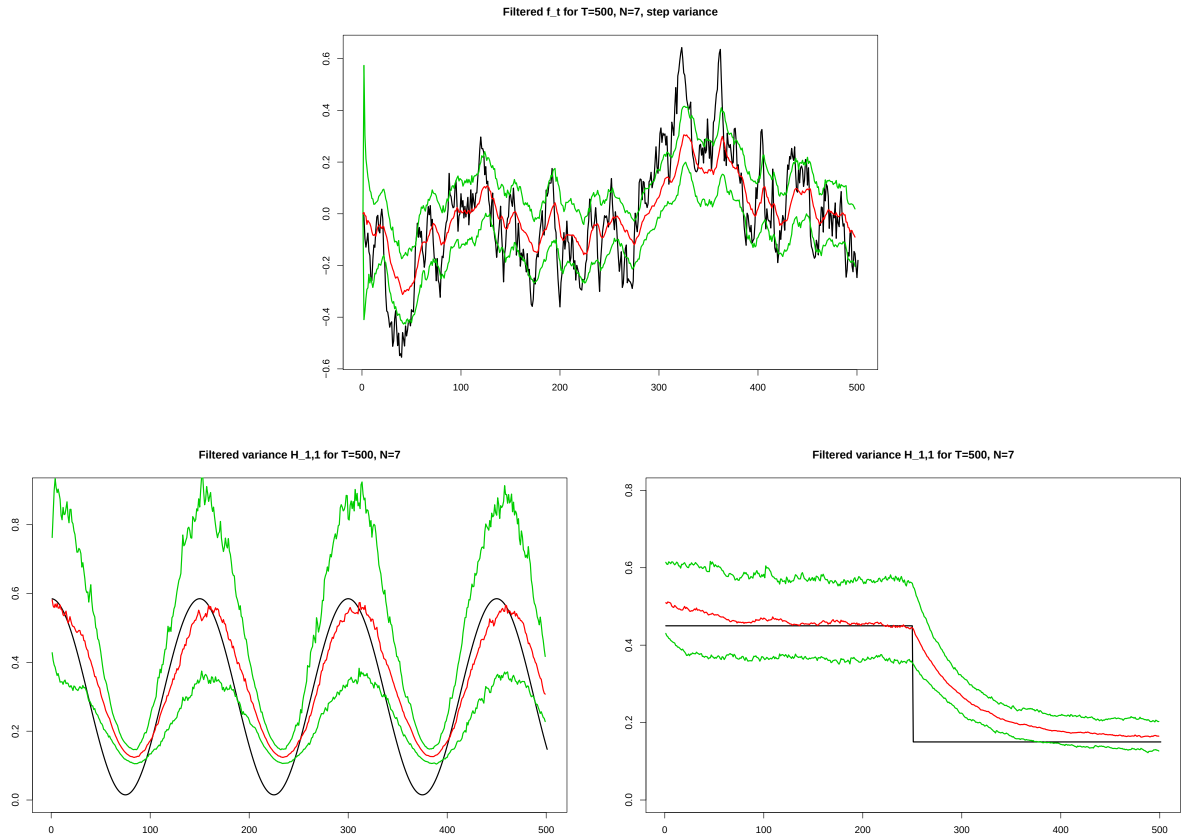

Figure 2: Upper panel: True AR(1) process for the factor driving the time-variation in the coefficient matrices, with median filtered path (red) and empirical 5% and 95% quantiles (green) over 200 simulation replications. Lower panel: True variance patterns (sine and step) with median filtered path (red) and empirical 5% and 95% quantiles (green) over 200 simulation replications.

variables on GDP growth. Consequently, $f_{t,1}$ captures the changing persistence in the six variables and $f_{t,2}$ indicates how financial-macro spillovers vary over time.

We also consider restricted versions of (14)-(15). Table 2 lists a goodness-of-fit comparison of several model specifications. Allowing for score-driven dynamics in the covariance matrix of the disturbances clearly improves the model fit. We find that in terms of the Akaike information criterion (AICc), the VAR model with one lag and time-variation in the coefficient matrices according to the two factors fits the data best.

Figure 4 shows a plot of the two VAR coefficient factors implied by the VAR model specification (4) in Table 2. Table 3 contains the parameter estimates of the final model.

Finally, Figure 5 shows plots of the estimated time-varying variances in our VAR model.

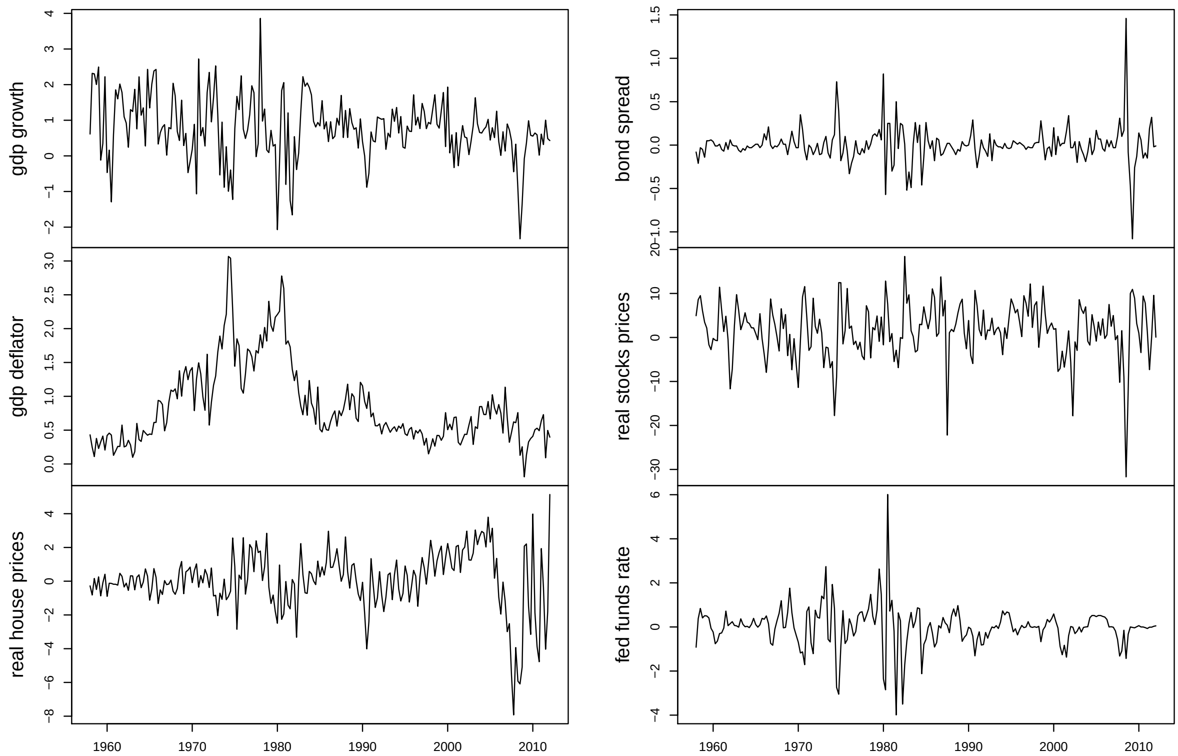

Figure 3: Plots of the macro and financial variables that enter the time-varying parameter VAR model.

Table 2: Specification

Comparison of the goodness-of-fit of eight empirical specifications: Static vs dynamic covariance matrices, static coefficient matrices vs coefficient matrices that are driven by one or two factors, respectively. We consider one and two lags for the VAR. Before estimation, the data are centered and scaled by their unconditional variances.

	one lag				two lags			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
H	✓				✓			
H_t		✓	✓	✓		✓	✓	✓
$f_{t,1}$			✓	✓			✓	✓
$f_{t,2}$				✓				✓
Φ^c	✓	✓	✓	✓	✓	✓	✓	✓
# parameters	36	40	47	52	72	76	89	98
LogLik	-1496.2	-1437.2	-1429.5	**-1414.5**	-1437.5	-1393.7	-1356.3	**-1348.5**
AICc	3079.2	2973.1	2979.7	**2966.6**	3092.1	3022.9	**3016.8**	3057.5

5 Conclusion

Table 3: Parameter estimates

Parameter estimates for the VAR model with one lag, two factors for the coefficient matrices, and a time-varying covariance matrix. For the parameters ω_1, ω_2, A, B, φ_1, and φ_2, we use transform for numerical stability within the estimation. Therefore, we report both the unconstrained estimates, and the constrained estimates with their standard errors in brackets. The nonzero entries in the loading matrices Φ_1^f and Φ_2^f are not constrained in the estimation and are therefore reported directly with their standard errors.

ω_1	0.6964	-0.3618	$\Phi^f_{1,11}$	-1.1224	$\Phi^f_{2,13}$	-0.0830
		(1.0357)		(0.2342)		(1.2251)
ω_2	-0.0163	-0.0467	$\Phi^f_{1,22}$	-0.5944	$\Phi^f_{2,14}$	-0.1355
		(1.1812)		(0.3513)		(1.5007)
A	0.3255	-0.7284	$\Phi^f_{1,33}$	0.5313	$\Phi^f_{2,15}$	0.2935
		(2.2364)		(0.4356)		(1.3403)
B	0.9171	2.4032	$\Phi^f_{1,44}$	0.0149	$\Phi^f_{2,16}$	0.1391
		(1.2238)		(0.5280)		(0.9997)
φ_1	0.3554	-0.5952	$\Phi^f_{1,55}$	0.0319		
		(0.2646)		(0.4720)		
φ_2	0.9197	2.4380	$\Phi^f_{1,66}$	0.9570		
		(0.0970)		(0.2649)		

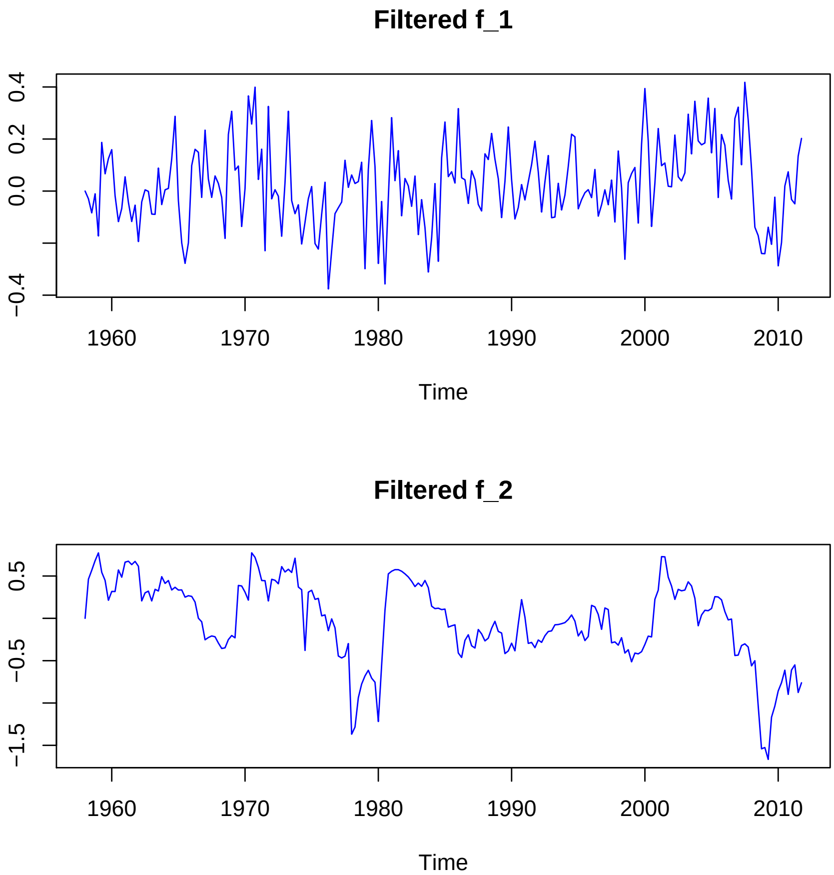

Figure 4: Filtered factors

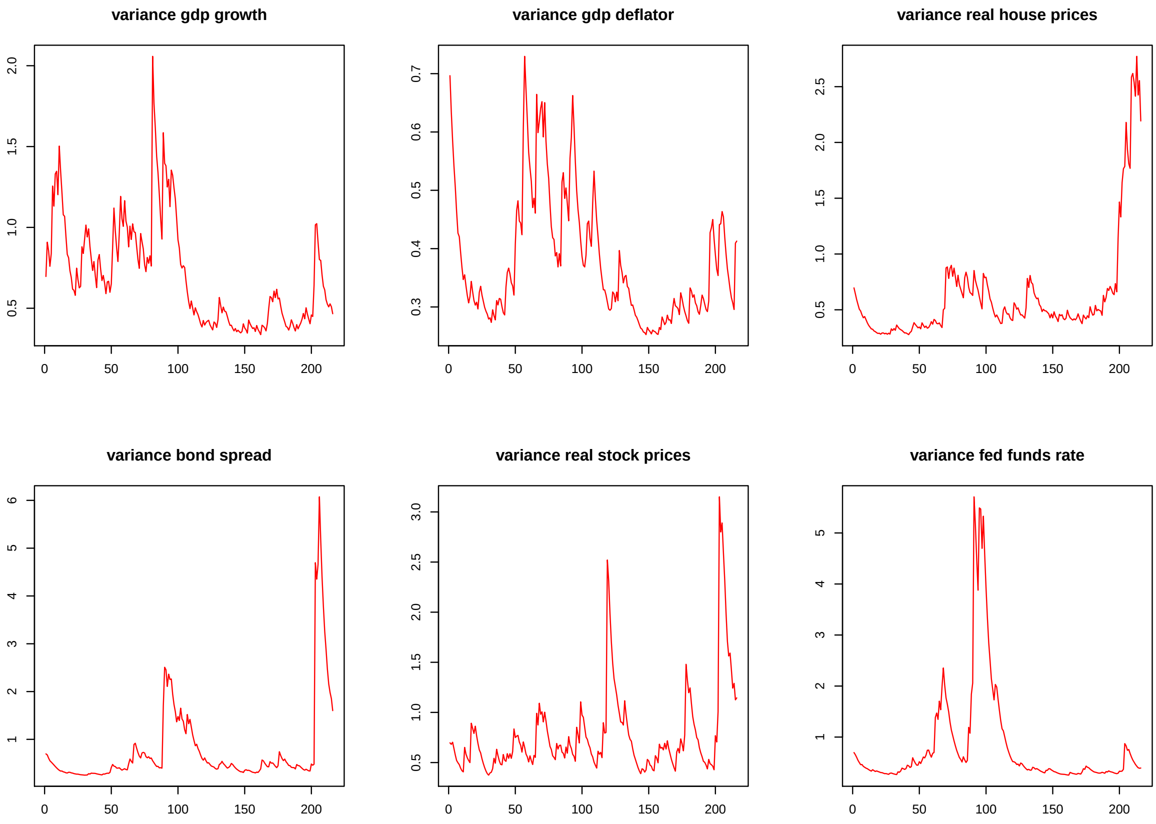

Figure 5: Filtered variaces

References

Canova, F. and Ciccarelli, M. (2004). Forecasting and turning point predictions in a bayesian panel var model. *Journal of Econometrics*, 120:327–359.

Canova, F. and Ciccarelli, M. (2009). Estimating multicountry var models. *International Economic Review*, 50:929–959.

Creal, D. D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28:777–795.

Delle Monache, D., Petrella, I., and Venditti, F. (2016). Adaptive state space models with applications to the business cycle and financial stress. *Working Paper*.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2nd edition.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press, Princeton.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails*. Cambridge University Press.

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin.

Magnus, J. and Neudecker, H. (1980). The elimination matrix: Some lemmas and applications. *SIAM Journal on Algebraic and Discrete Methods*, 1:422–449.

Magnus, J. and Neudecker, H. (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics 3rd edition*. Wiley Series in Probabiliy and Statistics.

Prieto, E., Eickmeier, S., and Marcellino, M. (2016). Time variation in macro-financial linkages. *Journal of Applied Econometrics*, 31:1215.123.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72:821852.