Time-Varying Vector Autoregressive Models with Structural Dynamic Factors

Paolo Gorgi, Siem Jan Koopman, Julia Schaumburg

http://sjkoopman.net

Vrije Universiteit Amsterdam School of Business and Economics
CREATES, Aarhus University

ECB Workshop on "Advances in short-term forecasting"

29 September 2017

The VAR model

Consider the vector autoregressive (VAR) model of order p, the VAR(p) model:

$$y_t = \Phi_{[1]} y_{t-1} + \ldots + \Phi_{[p]} y_{t-p} + \varepsilon_t, \qquad \varepsilon_t \sim NID(0, H), \qquad t = p+1, \ldots, T,$$

where $\Phi_{[i]}$ is coefficient matrix, $i = 1, \ldots, p$, and H is variance matrix.

For parameter estimation, forecasting, impulse response analysis, etc., we refer to Hamilton (1994) and Lutkepohl (2005), amongst many others.

The VAR(p) model can be efficiently formulated as

$$y_t = \Phi Y_{t-1:p} + \varepsilon_t, \qquad \Phi = \left[\Phi_{[1]}, \ldots, \Phi_{[p]}\right], \qquad Y_{t-1:p} = \left(y_{t-1}', \ldots, y_{t-p}'\right)'$$

We assume that the initial observation set $\{y_1, \ldots, y_p\}$ is fixed and given.

Motivation for time-varying parameters in VAR model

In macroeconometrics, vector-autoregressive (VAR) models are often extended with time-varying parameters, we have

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

Earlier literature:

- Cogley/Sargent (2001 NBER): inflation-unemployment dynamics in changing monetary policy regimes
- Primiceri (2005 RES): role of monetary policy for macroeconomic performance in the 1970s and 1980s
- Canova/Ciccarelli (2004 JE, 2009 IER): Bayesian panel VAR models, multi-country analyses;
- Hubrich/Tetlow (2015 JME): amplification and feedback effects between financial sector shocks and economy in crises vs. normal times
- Prieto/Eickmeier/Marcellino (2016, JAE): role of financial shocks on macroeconomic variables during crisis 2008/09

Estimation for TV-VAR models

- The common approach to parameter estimation for VAR models with time-varying coefficients is based on Bayesian methods
- We develop an alternative methodology based on dynamic factors for VAR coefficient matrices and score-driven dynamics for the variance matrices:
  - flexible modeling setup: it allows for wide variety of empirical specifications;
  - simple, transparent and fast implementation: least squares methods and Kalman filter;
  - relatively easy for estimation, impulse response, analysis and forecasting.
- Our approach is explored in more generality by Delle Monache et al. (2016): *adaptive state space models*

Outline

- Introduction
- Econometric Model
- Simulations
- Empirical application: Macro-financial linkages in the U.S. economy
- Conclusion

Time-varying autoregressive coefficient matrix

- TV-VAR model :

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

where we assume that sequence of variance matrices $H_{p+1}, \ldots, H_T$ is known and fixed, for the moment.

- The time-varying VAR coefficient is a matrix function,

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

where the unobserved $r \times 1$ vector f_t has dynamic specification

$$f_{t+1} = \varphi f_t + \eta_t, \qquad \eta_t \sim N(0, \Sigma_\eta),$$

where φ is $r \times r$ diagonal matrix of coefficients and $\Sigma_\eta = \mathrm{I}_r - \varphi\varphi'$.

Linear Gaussian state space form

We have $y_t = \Phi_t Y_{t-1:p} + \varepsilon_t$ and $\Phi_t = \Phi(f_t) = \Phi^c + \Phi_1^f f_{1,t} + \cdots + \Phi_r^f f_{r,t}$.
Define $\tilde{y}_t = y_t - \Phi^c Y_{t-1:p}$ and consider the following equation equalities

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

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

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

where the properties of the disturbances ε_t and η_t are discussed above.

Kalman filter

Prediction error is defined as $v_t = y_t - E(y_t|\mathcal{F}_{t-1};\psi)$:

- $\mathcal{F}_{t-1}$ is set of all past information, including past observations;
- ψ is the parameter vector that collects all unknown coefficients in Φ^c, Φ_1^f, ..., Φ_r^f, H_{p+1}, ..., H_T, φ;
- when model is correct, the sequence $\{v_{p+1}, \ldots, v_T\}$ is serially uncorrelated;
- variance matrix of the prediction error is $F_t = Var(v_t|\mathcal{F}_{t-1};\psi) = Var(v_t;\psi)$.

For a given vector ψ, the Kalman filter is given by

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

with $a_t = E(f_t|\mathcal{F}_{t-1};\psi)$ and variance matrix $P_t = Var(f_t - a_t|\mathcal{F}_{t-1};\psi)$, for $t = p+1, \ldots, T$. Loglikelihood function is

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

Parameter estimation (1)

For model

$$y_t = \Phi_t Y_{t-1:p} + \varepsilon_t, \qquad \Phi_t = \Phi(f_t) = \Phi^c + \Phi_1^f f_{1,t} + \cdots + \Phi_r^f f_{r,t}, \qquad f_{t+1} = \varphi f_t + \eta_t,$$

parameter estimation concentrates on φ, Φ^c and Φ_i^f, for $i = 1, \ldots, r$.

MLE is maximisation of $\ell(\psi)$ wrt ψ, heavy task.

Our strategy :
Step 1: Use economic information (if available) to restrict entries of Φ^c and Φ_i^f.

Step 2: Obtain estimates of Φ^c via least squares method on static VAR.

Step 3: Only place coefficients of φ and Φ_i^f in ψ.

Step 4: Estimate this ψ by MLE using the Kalman filter.

Least squares estimate of Φ from static VAR is consistent estimate of Φ^c.
Notice that $E(f_t) = 0$.
MLE is for a small dimension of ψ.

Time-varying variance matrix

Each Kalman filter step at time t requires a value for H_t.

We use score-driven approach of Creal et al. (2013) to let variance matrix H_t change recursively over time.

We have $N^* \times 1$ vector $f_t^\sigma = \mathrm{vech}(H_t)$ with $N^* = N(N+1)/2$ and dynamic specification

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

where ω is constant vector, A and B are square coefficient matrices and s_t is innovation vector.

Distinguishing feature of score-driven model is definition of s_t as the scaled score vector of $\ell_t \equiv \ell_t(\psi)$ with respect to f_t^σ.

We have $s_t = S_t \nabla_t$ where S_t is scaling matrix and ∇_t is gradient vector.

Score-driven model for time-varying variance matrix

The transpose of the gradient vector is given by

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

last equality holds since $F_t = Z_t P_t Z_t' + H_t$ and Z_t does not depend on H_t and P_t is function of $H_{p+1}, \ldots, H_{t-1}$, but not H_t.

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

where D_N is the $N^2 \times N^*$ duplication matrix, see Magnus and Neudecker (2007).

It follows that

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

Score-driven model for time-varying variance matrix

The inverse of the information matrix is taken as scaling matrix S_t for gradient vector.

Information matrix:

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

since $Var[\mathrm{vec}(v_t\, v_t') - \mathrm{vec}(F_t) | \mathcal{F}_{t-1}] = (\mathrm{I}_{N^2} + C_N)\,(F_t \otimes F_t)$ and $(\mathrm{I}_{N^2} + C_N) D_N = 2\, D_N$, where C_N is the $N^2 \times N^2$ commutation matrix.

The inverse of the information matrix:

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

where $D_N^+ = (D_N'\, D_N)^{-1} D_N'$ is the elimination matrix for symmetric matrices.

Score-driven model for time-varying variance matrix

We set the scaling as $S_t = \mathcal{I}_t^{-1}$. The scaled score $s_t = \mathcal{I}_t^{-1}\nabla_t$ becomes

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

For the score-driven update of the variance factors in f_t^σ, we obtain

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

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

The score updating function can easily be incorporated in the Kalman filter:

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

Direct updating for time-varying variance matrix

We have $\mathrm{vec}(H_t) = D_N\,\mathrm{vech}(H_t) = D_N\, f_t^{\sigma}$ 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 $D_N\,A\,D_N^{+} = A^{*} \otimes A^{*}$ and $D_N\,B\,D_N^{+} = B^{*} \otimes B^{*}$, we have

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

with $\mathrm{vech}(\Omega) = \omega$.

In case $A = a \cdot \mathrm{I}_{N^*}$ and $B = b \cdot \mathrm{I}_{N^*}$, the updating reduces simply to

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

This time-varying variance matrix updating equation can even more conveniently be incorporated within the Kalman filter:

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

Parameter estimation (2)

For model $y_t = \Phi_t Y_{t-1:p} + \varepsilon_t$ with

$$\varepsilon_t \sim NID(0, H_t) \qquad f_t^{\sigma} = \text{vech}(H_t), \qquad f_{t+1}^{\sigma} = \omega + B\, f_t^{\sigma} + A\, s_t,$$

additional parameter estimation concentrates on ω, A and B.

MLE is maximisation of $\ell(\psi)$ wrt ψ, remains light task.

Our strategy :
Step 4: Notice that under stationarity, $E(f_t^{\sigma}) = (I - B)^{-1}\omega$.

Step 5: Hence least squares estimate of H in static VAR is consistent estimate of $\text{unvech}[(I - B)^{-1}\omega]$.

Step 6: Only place coefficients of A and B in ψ.

Step 7: Estimate slightly extended ψ by MLE using the Kalman filter with f_t^{σ} or direct H_t updating.

MLE is for a small dimension of ψ.

Outline

- Introduction
- Econometric Model
- **Simulations**
- Empirical application: Macro-financial linkages in the U.S. economy
- Conclusion

Simulation setup

- Zero-mean VAR(1) model with time-varying coefficient matrix and scalar factor f_t:

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

with

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

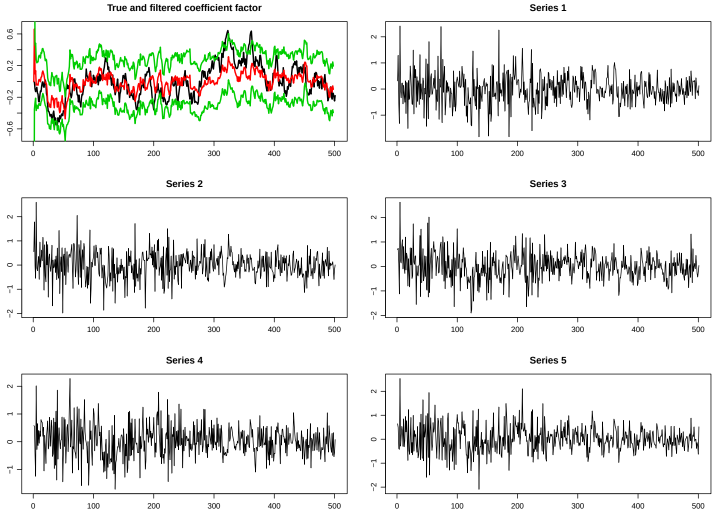
- Time-varying H_t : step function or sine function:

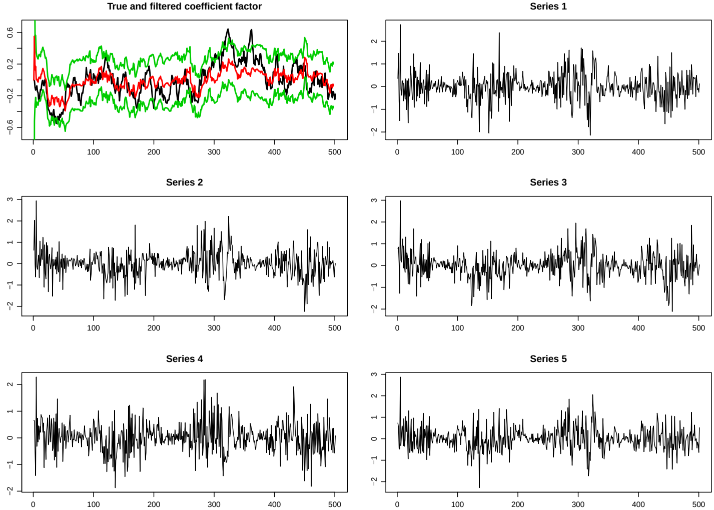
$$\begin{aligned} &\text{the sine function:} \qquad f_t^\sigma = 1 + 0.95\cos(2\pi t/150), \\ &\text{the step function:} \qquad f_t^\sigma = 1.5 - I(t > T/2). \end{aligned}$$

- $N = 5, 7;\ T = 250, 500$
- Parameter values:

$$\Sigma_\varepsilon = \mathrm{I}_N,\ \Phi^c_{ii} = 0.5,\ \Phi^c_{i,j} = -0.1,\ \Phi^f_{ii} = 0.2,\ \Phi^f_{i,j} = -0.1,\ \varphi = 0.6,$$

for $i, j = 1, \dots, N$ and $i \neq j$.

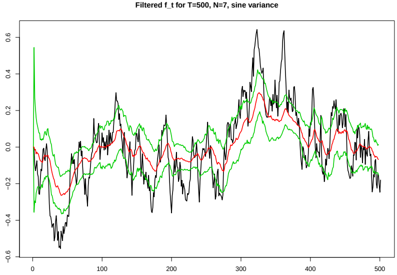

True and filtered coefficient factor
Series 1
Series 2
Series 3
Series 4
Series 5
0 100 200 300 400 500
−0.6 −0.2 0.2 0.6
−2 −1 0 1 2

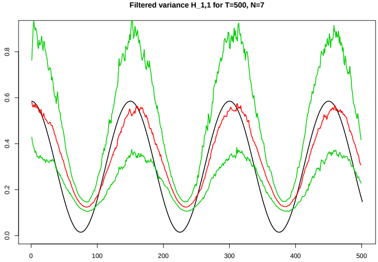

True and filtered coefficient factor
Series 1
Series 2
Series 3
Series 4
Series 5
0 100 200 300 400 500
−0.6 −0.2 0.2 0.6
−2 −1 0 1 2 3

Simulations: Mean Squared Errors

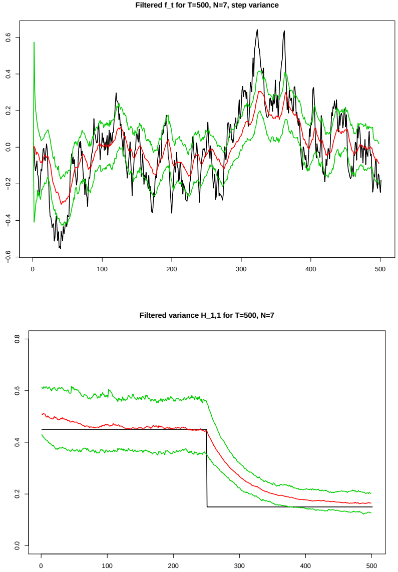
variance pattern "sine"				
	N=5		N=7	
	T=250	T=500	T=250	T=500
φ	0.01177	9e-04	0.00991	0.00048
Φ^f	0.41609	0.15935	0.33568	0.18428
f_t	0.06768	0.06491	0.06332	0.06485
variance pattern "step"				
	N=5		N=7	
	T=250	T=500	T=250	T=500
φ	0.01122	0.00076	0.00828	0.00063
Φ^f	0.24277	0.06938	0.20426	0.07465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

Simulations with sine function for variance





Simulations with step function for variance



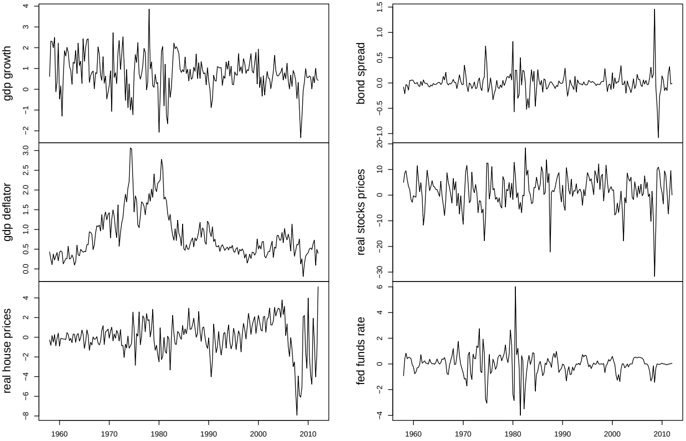
Outline

1. Introduction
2. Econometric Model
3. Simulations
4. **Empirical application: Macro-financial linkages in the U.S. economy**
5. Conclusion

Application: Macro-financial linkages

- Six-dimensional VAR with two lags, data from Prieto/Eickmeier/Marcellino (2016, JAE).
- Macroeconomic variables: Nominal GDP growth, inflation (GDP deflator).
- Financial variables: real house price inflation, corporate bond spread (Baa-Aaa), real stock price inflation, federal funds rate.
- Data transformations such that all time series are $I(0)$.
- Sample: 1958Q1 - 2012Q2.

Transformed data set



Empirical specification

VAR(2) model with two factors:

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

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 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.

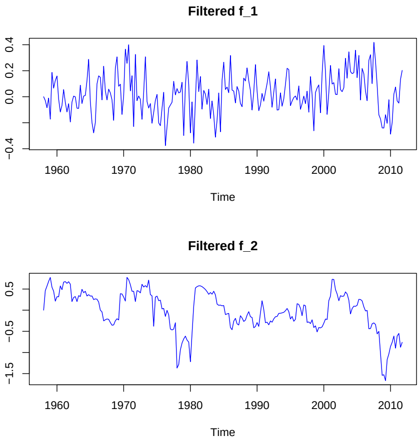
Model specifications

	one lag				two lags			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
H	✓				✓			
H_t		✓	✓	✓		✓	✓	✓
$f_{t,1}$			✓	✓			✓	✓
$f_{t,2}$				✓				✓
Φ^c	✓	✓	✓	✓	✓	✓	✓	✓
$\#\psi$	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

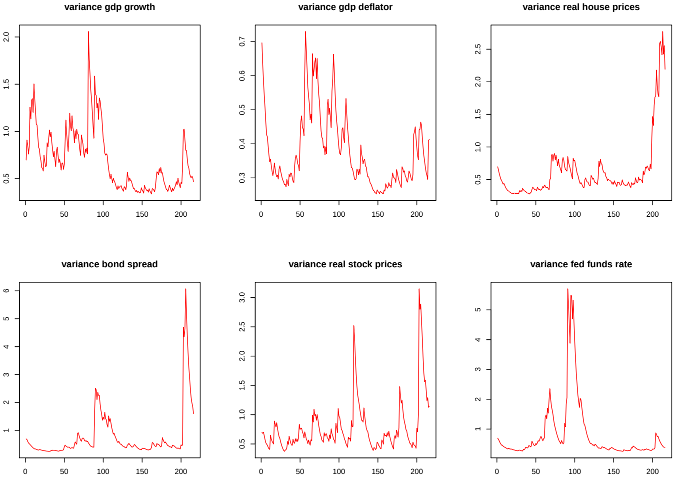
Some estimation results

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)		

Filtered factors $f_{t,1}$ and $f_{t,2}$



Time-varying variances



Conclusion

- New frequentist estimation method for VAR models with time-varying coefficient matrices.
- Simple, fast and transparent implementation, but highly flexible for empirical specifications.
- Simulations: Good performance in filtering dynamic factors and estimation of constant parameters.
- Empirics: Evidence for time-variation in financial-macro spillover coefficients.
- Future steps:
  - Impulse response functions.
  - Forecasting
  - Derivation of stability conditions, consistency and asymptotic theory.
  - Extension of empirical analysis to include formal significant tests.

Thank you.

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*.

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

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

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