LINKÖPING UNIVERSITY

US Interest Rates: Are Relations Stable?

Sune Karlsson[1], Tamás Kiss[1], **Hoang Nguyen**[2] and Pär Österholm[1]

[1]School of Business, Örebro University
[2]Department of Management and Engineering, Linköping University

Friday 27th June, 2025

Introduction and motivation

- A lot has happened on interest rate markets (yield curve, spreads) during the past decades
    - Historically low interest rate levels after the GFC
    - Unconventional monetary policy tools (QE)
    - Tapering
    - Covid
    - Skyrocketing inflation and policy rates in 2022/23
- Most empirical models impose time-invariant relationships
    - Pooling data from very different periods
    - Need to assess whether this is tenable
- Large shocks more frequent than what a normal distribution suggests
    - Importance of allowing for fat-tailed/flexible distributions

Research question and contributions

Research question: Have interest rate markets changed from a modelling perspective?

Contributions:

1. Extend the standard BVAR model to allow for time-varying parameters in all or some equations and fat tails in the innovation distributions
2. Revisit Duffee (1998) by estimating a trivariate VAR with the short rate, term spread and corporate bond spread (credit spread) on US data up until early 2025

A short summary of results

- Time-variation matters mostly for the credit spread
    - Highest marginal likelihood for models with only time variation in the credit spread equation
    - helps improve point and density forecasts of the credit spread
- Short rate is best modelled with constant parameters and fat-tailed innovations

"Structural" form TVP-VAR

$$\boldsymbol{A}_t \boldsymbol{y}_t = \boldsymbol{b}_t + \boldsymbol{B}_{1t} \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{B}_{pt} \boldsymbol{y}_{t-p} + \boldsymbol{u}_t, \quad t = 1, \ldots, T,$$

- $\boldsymbol{A}_t$ lower triangular with ones on the diagonal
- Elements of $\boldsymbol{u}_t$ uncorrelated
- No restrictions on the reduced form and no structural interpretation
- Can be estimated equation by equation
- Time variation, parameters follow a random walk

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\theta)$$

where $\boldsymbol{\theta}_t$ collects the elements in $\boldsymbol{A}_t$, $\boldsymbol{b}_t$, $\boldsymbol{B}_{1t} \ldots \boldsymbol{B}_{pt}$

Stochastic volatility and fat tails

- Let

$$\boldsymbol{u}_t = \boldsymbol{W}_t^{1/2}\boldsymbol{H}_t^{1/2}\boldsymbol{\varepsilon}_t$$

  with $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \boldsymbol{I})$

- Logvolatilites follow a random walk

$$\boldsymbol{H}_t = \text{diag}(\exp(h_{1,t}), \ldots \exp(h_{n,t}))$$

$$h_{i,t} = h_{i,t-1} + \zeta_{i,t}, \zeta_{i,t} \sim \mathcal{N}(0, \sigma^2_{h,i})$$

- Mixing variables $w_{i,t} \sim \mathcal{IG}(\nu_i/2, \nu_i/2)$ with $\boldsymbol{W}_t = \text{diag}(w_{i,t})$
- $u_{i,t}$ independent, t-distributed with ν_i degrees of freedom and scale parameter $\exp(h_{i,t})$
- Orthogonal Student-t Stochastic Volatility, OT-SV

Order matters

- Reduced form errors $\boldsymbol{e}_t = \boldsymbol{A}_t^{-1}\boldsymbol{u}_t$ depend on order of variables
- $e_{i,t}$ linear combination of $u_{1,t}$ to $u_{i,t}$
  - Affects distribution
  - Affects volatilities
- Reduced form parameters $\boldsymbol{c}_t = \boldsymbol{A}_t^{-1}\boldsymbol{b}_t$, $\boldsymbol{C}_{i,t} = \boldsymbol{A}_t^{-1}\boldsymbol{B}_{i,t}$
- Implied prior on reduced form parameters nonlinear function of prior on structural form parameters and depends on order
- Implied prior on reduced form error covariance matrix $\boldsymbol{A}_t^{-1}\boldsymbol{W}_t^{1/2}\boldsymbol{H}_t\boldsymbol{W}_t^{1/2}\boldsymbol{A}_t^{-\top}$ not order invariant

Hybrid TVP-VAR

- Follow Chan and Eisenstat (2018)
- Rewrite the equations of the TVP-VAR as

$$y_{i,t} = \tilde{\boldsymbol{x}}_t \boldsymbol{\beta}_{i,t} + \boldsymbol{z}_{i,t} \boldsymbol{\alpha}_{i,t} + \sqrt{w_{i,t}} \exp(h_{i,t}/2) \varepsilon_{i,t}, \quad i = 1, \ldots, k,$$

with $\tilde{\boldsymbol{x}}_t = (1, \boldsymbol{y}_{t-1}^{'}, \ldots, \boldsymbol{y}_{t-p}^{'})$, and $\boldsymbol{z}_{i,t} = (-y_{1,t}, \ldots, -y_{i-1,t})$

- Collect the parameters $\boldsymbol{\theta}_{it} = (\boldsymbol{\beta}_{i,t}^{'}, \boldsymbol{\alpha}_{i,t}^{'})^{'}$ and let $\boldsymbol{x}_{i,t} = (\tilde{\boldsymbol{x}}_t, \boldsymbol{z}_{i,t})$, the TVP VAR equation evolves as

$$\begin{aligned} y_{i,t} &= \boldsymbol{x}_{i,t} \boldsymbol{\theta}_{i,t} + \sqrt{w_{i,t}} \exp(h_{i,t}/2) \varepsilon_{i,t}, \\ \boldsymbol{\theta}_{i,t} &= \boldsymbol{\theta}_{i,t-1} + \boldsymbol{\eta}_{i,t}, \quad \boldsymbol{\eta}_{i,t} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\theta,i}), \\ h_{i,t} &= h_{i,t-1} + \zeta_{it} \quad \zeta_{i,t} \sim \mathcal{N}(0, \sigma_{h,i}^2). \end{aligned}$$

- "Easy" equation by equation estimation
- Can shut down time varying parameters in individual equations

Priors Posteriors Marginal likelihood

Empirical illustration

- The data is collected at monthly frequency from 1953M4 - 2025M2 from the Federal Reserve Bank of St. Louis

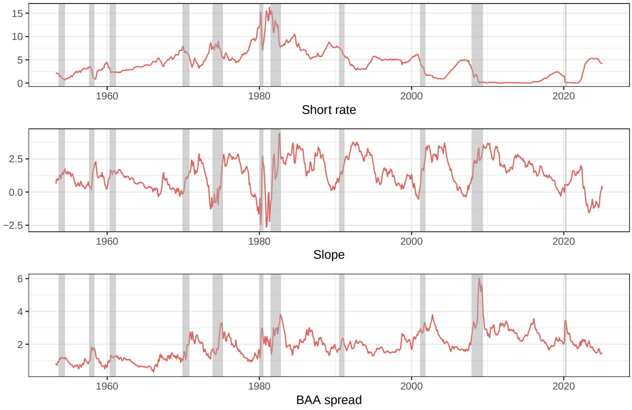

Figure: Data

Insample marginal likelihood

16 VAR models (p=3 lags) w/wo SV: Gaussian, Student-t and hybrid.

Table: Log marginal likelihood for VAR models with stochastic volatility

Model	Gaussian	Orthogonal Student's t	
	Log marginal likelihood	Log marginal likelihood	Degrees of freedom
000	1165.59	1173.51	(14.99; 26.64; 13.74)
100	1115.04	1122.54	(11.92; 25.22; 13.16)
010	1118.99	1126.50	(14.46; 26.11; 13.87)
001	1200.41	1206.92	(15.41; 26.66; 9.83)
110	1067.96	1075.45	(10.31; 29.23; 12.64)
101	1149.55	1153.51	(11.88; 26.53; 10.31)
011	1153.24	1158.26	(15.00; 25.24; 10.29)
111	1101.29	1105.07	(11.48; 29.19; 9.50)

- t-distribution preferred over Gaussian
- Model with TVP only in equation for BAA spread preferred

Posterior distributions

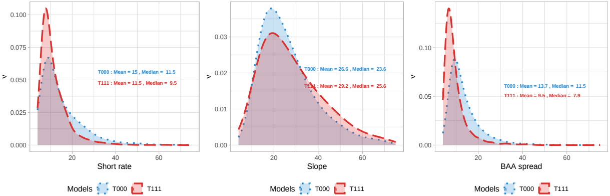

Figure: Posterior distribution of the degrees of freedom, ν_i, OT-SV model with and without TVP.

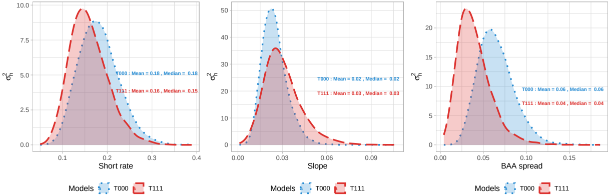

Figure: Posterior distribution of $\sigma^2_{h,i}$ of the OT-SV model models with and without TVP.

Time-varying parameters

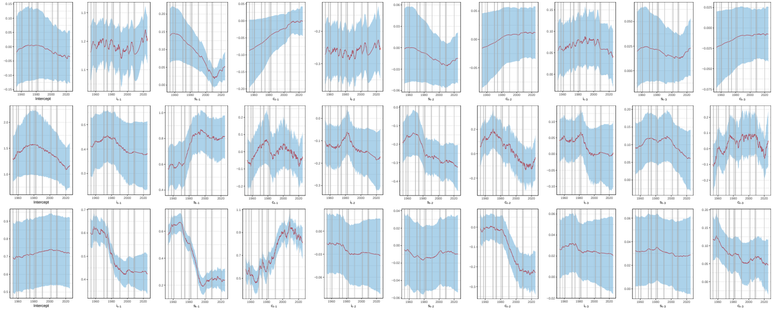

Figure: The time varying parameter $\boldsymbol{\theta}$ of the VAR(3)-OT-SV models with TVP.

Posterior distributions

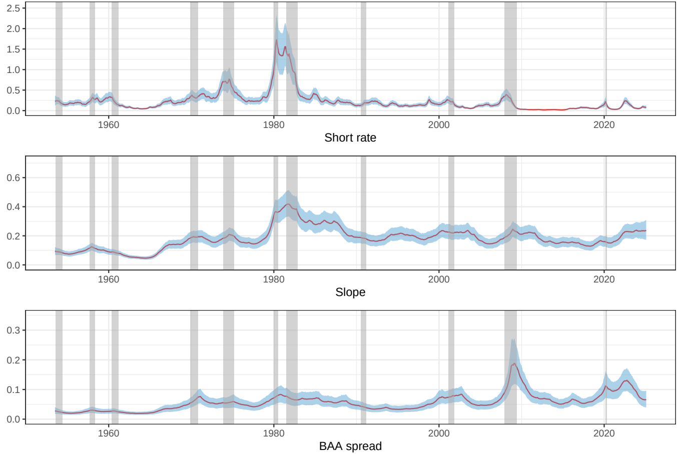

Figure: Stochastic volatility estimates T001 model

The red solid line gives the point estimate under the assumption of an Orthogonal Student's t-distribution. The coloured bands bands show the 80% credible interval.

Impulse responses

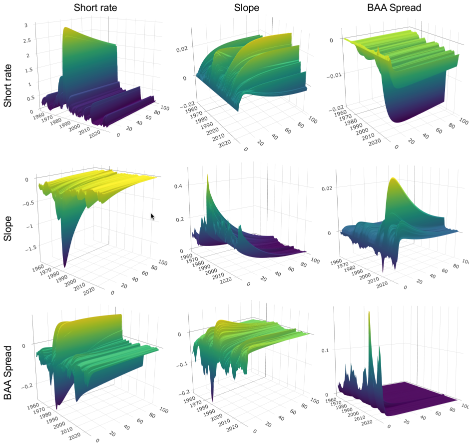

Figure: The impulse response functions of the T001 model. Standard deviation shocks

Impulses are shown in the column headers, response variables are given by the rows. The size of the impulse is one standard deviation. The two horizontal axes represent the date of the shock (left horizontal axis) and the horizon in months (right horizontal axis), respectively.

Forecast performance

- 279 recursive forecasts.
- $h = 1, \ldots 24$ steps ahead forecasts.
- Estimation samples end in 1999M12 to 2023M2.
- Evaluated using Mean Squared Forecast errors and Log Predictive Scores

Forecast performance - point forecasts

- Short rate
  - No gain from TVP
  - t-distribution improves
- Slope
  - No gain from TVP
  - Weak improvement with t-distribution
- BAA spread
  - TVP in spread equation improves
  - No clear gain from t-distribution

Forecast performance - point forecasts

Table: MSFEs and relative MSFEs

	1M	2M	3M	6M	12M	24M
(a) 3m Tbill						
G000	0.027	0.088	0.172	0.573	1.825	4.753
G001	0.996	0.997	0.999	1.007	1.000	0.998
G011	0.998	1.001	1.007	1.022	0.978	0.938
G111	1.086	1.107	1.091	1.040	1.060	1.136
T000	0.998	0.998	0.998	0.994	0.992*	0.993*
T001	0.994	0.994	0.997	1.002	0.991	0.989
T011	0.997	1.003	1.009	1.021	0.969	0.928
T111	1.055	1.065	1.050	0.999	1.029	1.117

The first line of each panel reports the MSFE of the benchmark Gaussian VAR model without time-varying parameters during January 2000 to March 2023 (279 estimations). The relative performance is computed as the ratio of the MSFE of alternative specifications over the benchmark; entries less than 1 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Forecast performance - point forecasts

Table: MSFEs and relative MSFEs

	1M	2M	3M	6M	12M	24M
(b) Slope						
G000	0.052	0.142	0.226	0.470	0.878	1.406
G001	0.998	0.999	1.009	1.057	1.164	1.369
G011	0.998	0.970	0.903*	0.827*	1.062	1.370
G111	1.003	0.978	0.915	0.848	1.197	1.671
T000	1.003	1.007	1.006	0.999	0.994	0.990**
T001	1.005	1.009	1.020	1.062	1.161	1.358
T011	1.002	0.971	0.902*	0.825*	1.053	1.356
T111	1.000	0.968	0.901	0.826	1.180	1.661

The first line of each panel reports the MSFE of the benchmark Gaussian VAR model without time-varying parameters during January 2000 to March 2023 (279 estimations). The relative performance is computed as the ratio of the MSFE of alternative specifications over the benchmark; entries less than 1 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Forecast performance - point forecasts

Table: MSFEs and relative MSFEs

	1M	2M	3M	6M	12M	24M
(c) BAA spread						
G000	0.040	0.118	0.202	0.436	0.836	1.326
G001	0.932	0.922	0.897	0.860	0.795*	0.781*
G011	0.945	0.939	0.915	0.871	0.782*	0.762*
G111	0.957	0.964	0.949	0.911	0.834	0.853
T000	1.001	1.002	0.999	0.999	0.998	1.001
T001	0.933	0.933	0.899	0.852*	0.784*	0.774*
T011	0.947	0.949	0.916	0.874	0.781*	0.762*
T111	0.956	0.966	0.939	0.892	0.810	0.838

The first line of each panel reports the MSFE of the benchmark Gaussian VAR model without time-varying parameters during January 2000 to March 2023 (279 estimations). The relative performance is computed as the ratio of the MSFE of alternative specifications over the benchmark; entries less than 1 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Forecast performance - density forecasts

- Short rate
  - No gain from TVP
  - No gain from t-distribution
- Slope
  - Weak gain from TVP in slope equation
  - t-distribution improves
- BAA spread
  - TVP in spread equation improves
  - No clear gain from t-distributon

Forecast performance - density forecasts

Table: LPS and difference in LPS relative to the Gaussian VAR model with constant parameters

	1M	2M	3M	6M	12M	24M
(a) TBILL 3M						
G000	0.969	0.304	-0.089	-0.931	-1.915	-2.626
G001	-0.001	-0.000	-0.012	-0.028	-0.036	-0.013
G011	-0.004	-0.010	-0.024	-0.037	0.012	0.013
G111	-0.062	-0.062	-0.067	-0.070	-0.057	0.007
T000	0.009*	-0.002	-0.012	-0.034	-0.133	-0.062
T001	0.000	-0.008	-0.023	-0.075	-0.020	-0.079
T011	0.000	-0.011	-0.033	-0.055	-0.234	-0.061
T111	-0.048	-0.048	-0.066	-0.060	-0.274	-0.069

The first line of each panel reports the LPS of the benchmark Gaussian VAR model without time-varying parameters during January 2000 to March 2023 (279 estimations). The difference in LPS is computed as the LPS of the alternative specifications minus the LPS of the benchmark model; entries greater than 0 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Forecast performance - density forecasts

Table: LPS and difference in LPS relative to the Gaussian VAR model with constant parameters

	1M	2M	3M	6M	12M	24M
(b) Slope						
G000	0.104	-0.399	-0.657	-1.093	-1.474	-1.785
G001	-0.001	0.000	-0.005	-0.033	-0.096	-0.200
G011	0.018	0.046*	0.087**	0.138**	0.020	-0.130
G111	0.023	0.051*	0.098**	0.168**	0.048	-0.121
T000	0.005*	0.003	-0.000	0.001	-0.005	-0.007
T001	0.004	0.001	-0.008	-0.037	-0.100	-0.211
T011	0.021	0.049*	0.091**	0.143**	0.030	-0.124
T111	0.024	0.054*	0.103**	0.175**	0.054	-0.118

The first line of each panel reports the LPS of the benchmark Gaussian VAR model without time-varying parameters during January 2000 to March 2023 (279 estimations). The difference in LPS is computed as the LPS of the alternative specifications minus the LPS of the benchmark model; entries greater than 0 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Forecast performance - density forecasts

Table: LPS and difference in LPS relative to the Gaussian VAR model with constant parameters

	1M	2M	3M	6M	12M	24M
(c) BAA spread						
G000	0.408	-0.160	-0.472	-0.985	-1.509	-1.879
G001	0.056**	0.052	0.074	0.207**	0.415***	0.563***
G011	0.035	-0.005	0.005	0.207**	0.418***	0.545***
G111	0.041	-0.006	0.020	0.209**	0.412**	0.533***
T000	0.014**	0.011**	0.003	0.000	0.003	-0.027
T001	0.079***	0.043	0.081	0.214**	0.429***	0.585***
T011	0.051*	0.046	0.080	0.211**	0.424***	0.549***
T111	0.078***	0.044	0.043	0.211**	0.421**	0.538***

The first line of each panel reports the LPS of the benchmark Gaussian VAR model without time-varying parameters during January 2000 to March 2023 (279 estimations). The difference in LPS is computed as the LPS of the alternative specifications minus the LPS of the benchmark model; entries greater than 0 indicate that the given model is better. ***,**,* denote that the corresponding model significantly outperforms the benchmark at 1%, 5%, 10% level based on the one-sided Diebold and Mariano (1995) test where the standard errors of the test statistics are computed with the Newey–West estimator (Clark, 2011).

Forecast performance - density forecasts

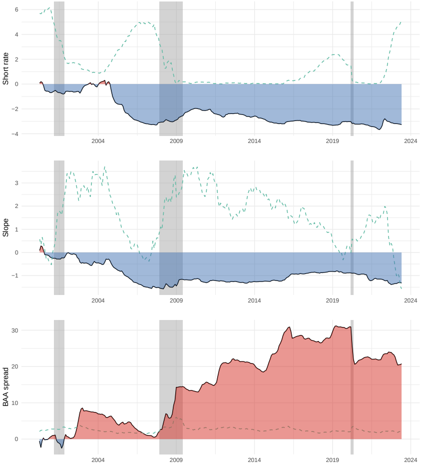

Figure: Cumulative log Bayes factors of the predictive density for 3 month ahead forecasts between the G001 and the G000 models.

Positive values (red) means G001 predicts better and negative values (blue) means that G000 model does better. The dashed lines illustrate the scale values of the original variables. See Geweke and Amisano (2010) for details.

Forecast performance - density forecasts

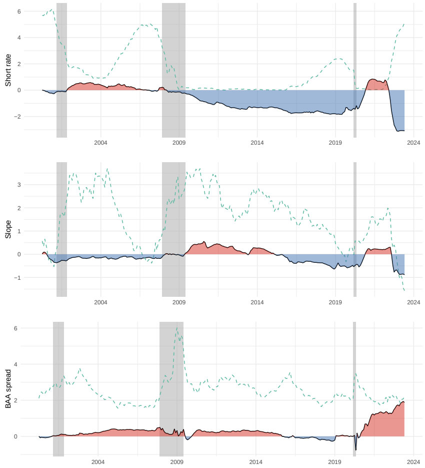

Figure: Cumulative log Bayes factors of the predictive density for 3 month ahead forecasts between the T001 and the G001 models.

Positive values (red) means T001 predicts better and negative values (blue) means that G001 model does better. The dashed lines illustrate the scale values of the original variables. See Geweke and Amisano (2010) for details.

Conclusions

- The interest rate markets have not been stable over time
- Both in sample and out of sample evidence favours time variation in the equation for the BAA spread
- The Short rate and Slope equations appear to have been stable over time
- The in sample evidence favours fat tails in the form of a *t*-distribution
- Little out of sample evidence in favour of fat tails
- An increase in the BAA spread decreases the three-month Treasury bill rate.
- An increase in the three-month Treasury bill rate decreases the BAA spread

Thank you

J. C. Chan and E. Eisenstat. Comparing hybrid time-varying parameter VARs. *Economics Letters*, 171:1–5, 2018.

T. E. Clark. Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. *Journal of Business and Economic Statistics*, 29(3):327–341, 2011.

F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):134–144, 1995.

G. R. Duffee. The relation between treasury yields and corporate bond yield spreads. *The Journal of Finance*, 53(6):2225–2241, 1998.

J. Geweke. Bayesian treatment of the independent Student-t linear model. *Journal of Applied Econometrics*, 8(S1):S19–S40, 1993.

J. Geweke and G. Amisano. Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230, 2010.

G. Kastner and S. Frühwirth-Schnatter. Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics and Data Analysis*, 76:408–423, 2014.

G. Koop and D. Korobilis. *Bayesian multivariate time series methods for empirical macroeconomics*. Now Publishers Inc, 2010.

G. O. Roberts and J. S. Rosenthal. Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009.

Y. Yu and X.-L. Meng. To center or not to center: That is not the question—an ancillarity–sufficiency interweaving strategy (ASIS) for boosting MCMC efficiency. *Journal of Computational and Graphical Statistics*, 20(3):531–570, 2011.

Prior distributions

We denote $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_{1:k,0}, \boldsymbol{\Sigma}_{\theta,1:k}, \boldsymbol{\Sigma}_{h,1:k}, \boldsymbol{\nu}, \widetilde{\boldsymbol{\theta}}_{1:k,1:T}, \mathbf{h}_{0:T}, \mathbf{w}_{1:T}\}$ as the parameter set of the HTVP-VAR model with orthogonal Student SV innovations.

- $\boldsymbol{\theta}_{i,0}$ follows a Minnesota prior with the overall shrinkage $l_1 = 0.2$ and the cross-variable shrinkage $l_2 = 0.5$, see Koop and Korobilis (2010).
- The prior for the degree of freedom is $\nu_i \sim \mathcal{G}(2, 0.1)$ such that the prior mean of the degrees of freedom is 20.
- The prior for the mixing variables are based on the model assumption $w_{i,t}|\nu_i \sim \mathcal{IG}(\frac{\nu_i}{2}, \frac{\nu_i}{2})$.
- The prior for the variance of shock to the volatility and time-varying parameters are $\sigma^2_{h,1:k} \sim \mathcal{G}(\frac{1}{2}, \frac{1}{2\mathbf{V}_h})$ and $\sigma^2_{\theta,1:k} \sim \mathcal{G}(\frac{1}{2}, \frac{1}{2\mathbf{V}_\theta})$ where $\mathbf{V}_h = \mathbf{V}_\theta = 1$.
- The priors of the initial volatility and $\mathbf{h}_0$ and initial regression coefficients $\boldsymbol{\theta}_{i,0}$ are $\mathbf{h}_0 \sim N(\mathbf{a}_{h_0}, \mathbf{V}_{h_0})$ where $\mathbf{a}_{h_0} = \log \hat{\Sigma}_{OLS}$ is the estimated residual variance of a univariate AR(p) model using the ordinary least square method, and $\mathbf{V}_{h_0} = 4\mathbf{I}_k$, see Clark (2011);

Back

MCMC preliminaries

Non-centered parameterization (NCP)

- For the time varying parameters rewrite the equation as

$$y_{i,t} = \boldsymbol{x}_{i,t}\boldsymbol{\theta}_{i,0} + \boldsymbol{x}_{i,t}\boldsymbol{\Sigma}_{\theta,i}^{1/2}\tilde{\boldsymbol{\theta}}_{i,t} + \sqrt{w_{i,t}}\exp(h_{i,t}/2)\varepsilon_{i,t}$$

for $\tilde{\boldsymbol{\theta}}_{i,t} = \boldsymbol{\Sigma}_{\theta,i}^{-1/2}\left(\boldsymbol{\theta}_{i,t} - \boldsymbol{\theta}_{i,0}\right) = \tilde{\boldsymbol{\theta}}_{i,t-1} + \tilde{\boldsymbol{\eta}}_t$ with $\tilde{\boldsymbol{\eta}}_t \sim \mathcal{N}(0, \boldsymbol{I})$

- For all time periods

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\theta}_{i,0} + \boldsymbol{Z}_i\widetilde{\boldsymbol{\theta}}_i + \boldsymbol{W}_i^{1/2}\boldsymbol{H}_i^{1/2}\boldsymbol{\varepsilon}_i$$

where $\boldsymbol{X}_i = (\boldsymbol{x}_{i,1}', \ldots, \boldsymbol{x}_{i,T}')'$, $\boldsymbol{Z}_i = \text{diag}(\boldsymbol{x}_{i,1}\boldsymbol{\Sigma}_{\theta,i}^{1/2}, \ldots, \boldsymbol{x}_{i,T}\boldsymbol{\Sigma}_{\theta,i}^{1/2})$, $\widetilde{\boldsymbol{\theta}}_i = (\widetilde{\boldsymbol{\theta}}_{i,1}', \ldots \widetilde{\boldsymbol{\theta}}_{i,T}')'$, $\boldsymbol{W}_i = \text{diag}(\boldsymbol{w}_i)$, $\boldsymbol{H}_i = \text{diag}(\exp(\boldsymbol{h}_i))$,

- For the logvolatilities let $\tilde{h}_{i,t} = h_{i,t}/\sigma_{h,i}$ we then have

$$\tilde{h}_{i,t} = \tilde{h}_{i,t-1} + \tilde{\zeta}_t, \quad \tilde{\zeta}_t \sim \mathcal{N}(0, 1)$$

Back

MCMC overview

For $i = 1, \ldots, n$

1. Sample $(\widetilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i})$ using the ancillarity-sufficiency interweaving strategy (ASIS) by proposed Yu and Meng (2011),
   1. Draw $\widetilde{\boldsymbol{\theta}}_i$ in the non-centered parameterization (NCP) setting.
   2. Draw $\boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i}$ in the NCP setting.
   3. Move $\boldsymbol{\theta}_{i,t} = \boldsymbol{\theta}_{i,0} + \boldsymbol{\Sigma}_{\theta,i}^{1/2}\widetilde{\boldsymbol{\theta}}_{i,t}$ in the centered parameterization (CP) setting.
   4. Draw $\boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i}$ in the CP setting.
   5. Move $\widetilde{\boldsymbol{\theta}}_{i,t} = \boldsymbol{\Sigma}_{\theta,i}^{-1/2}(\boldsymbol{\theta}_{i,t} - \boldsymbol{\theta}_{i,0})$ in the NCP setting.
2. Sample $\boldsymbol{h}_i, \sigma^2_{h,i}$ using the ASIS proposed by Kastner and Frühwirth-Schnatter (2014)
3. Sample $w_{i,t}$ for $t = 1, \ldots, T$ from Inverse Gamma distributions following Geweke (1993).
4. Sample ν_i, using an adaptive random walk Metropolis-Hastings algorithm, see Roberts and Rosenthal (2009).

Back

MCMC details

Sampling $(\widetilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i})$ using ASIS

1. Draw $\widetilde{\boldsymbol{\theta}}_i | \boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i}$ from the normal conjugate posterior in the NCP setting using

$$\boldsymbol{y}_i - \boldsymbol{X}_i \boldsymbol{\theta}_{i,0} = \boldsymbol{Z}_i \widetilde{\boldsymbol{\theta}}_i + \boldsymbol{W}_i^{1/2} \boldsymbol{H}_i^{1/2} \boldsymbol{\varepsilon}_i,$$

2. Draw $\boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i} | \widetilde{\boldsymbol{\theta}}_i$ from the normal conjugate posterior in the NCP setting using

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\theta}_{i,0} + \left(\boldsymbol{X}_i \odot \widetilde{\boldsymbol{\theta}}_i \right) \boldsymbol{\sigma}_{\theta,i} + \boldsymbol{W}_i^{1/2} \boldsymbol{H}_i^{1/2} \boldsymbol{\varepsilon}_i = \widetilde{\boldsymbol{X}}_i \boldsymbol{b}_i + \boldsymbol{W}_i^{1/2} \boldsymbol{H}_i^{1/2} \boldsymbol{\varepsilon}_i,$$

   where $\boldsymbol{b}_i = \left(\boldsymbol{\theta}'_{i,0}, \boldsymbol{\sigma}'_{\theta,i} \right)'$.

3. Move $\boldsymbol{\theta}_{i,t} = \boldsymbol{\theta}_{i,0} + \boldsymbol{\Sigma}_{\theta,i}^{1/2} \widetilde{\boldsymbol{\theta}}_{i,t}$ in the CP setting.
4. Draw $\boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i} | \boldsymbol{\theta}_{i,1:T}$ in the CP setting.
   - $\pi(\sigma^2_{\theta,i,j} | \boldsymbol{\theta}_{i,0:T})$ is generalized inverse Gaussian distribution (GIG)
   - $\pi(\boldsymbol{\theta}_{i,0} | \boldsymbol{\theta}_{i,1})$ is normal using $\boldsymbol{\theta}_{i,1} = \boldsymbol{\theta}_{i,0} + \boldsymbol{\eta}_{i,1}$
5. Move $\widetilde{\boldsymbol{\theta}}_{i,t} = \boldsymbol{\Sigma}_{\theta,i}^{-1/2} (\boldsymbol{\theta}_{i,t} - \boldsymbol{\theta}_{i,0})$ in the NCP setting.

Back

Marginal likelihood and model choice

Need to estimate

$$p(\boldsymbol{Y}) = \int p(\boldsymbol{Y}|\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})d\boldsymbol{\Theta} = \prod_{i=1}^{n} \int p(\boldsymbol{y}_i|\boldsymbol{\Theta}_i)\pi(\boldsymbol{\Theta}_i)d\boldsymbol{\Theta}_i$$
$$= \prod_{i=1}^{n} \int p(\boldsymbol{y}_i|\boldsymbol{\Theta}_{i,1}, \boldsymbol{\Theta}_{i,2})\pi(\boldsymbol{\Theta}_{i,2}|\boldsymbol{\Theta}_{i,1})d\boldsymbol{\Theta}_{i,2}\pi(\boldsymbol{\Theta}_{i,1})d\boldsymbol{\Theta}_{i,1}$$
$$= \prod_{i=1}^{n} \int p(\boldsymbol{y}_i|\boldsymbol{\Theta}_{i,1})\pi(\boldsymbol{\Theta}_{i,1})d\boldsymbol{\Theta}_{i,1}$$

for the parameters $\boldsymbol{\Theta}_{i,1} = \{\boldsymbol{\theta}_{i,0}, \boldsymbol{\Sigma}_{\theta,i}, \sigma^2_{h,i}, \nu_i\}$ and the latent states $\boldsymbol{\Theta}_{i,2} = \{\widetilde{\boldsymbol{\theta}}_i, \boldsymbol{h}_i, \boldsymbol{w}_i\}$.

Back

Marginal likelihood and model choice

Estimating $p(\boldsymbol{y}_i|\boldsymbol{\Theta}_{i,1})$

- $\widetilde{\boldsymbol{\theta}}_i$ can be integrated out analytically, leaving $p(\boldsymbol{y}_i|\boldsymbol{h}_i, \boldsymbol{w}_i, \boldsymbol{\Theta}_{i,1})$
- $p(\boldsymbol{y}_i|\boldsymbol{\Theta}_{i,1})$ is estimated using importance sampling

$$\hat{p}(\boldsymbol{y}_i|\boldsymbol{\Theta}_{i,1}) = \frac{1}{R}\sum_{r=1}^{R} \frac{p(\boldsymbol{y}_i|\boldsymbol{h}_i^{(r)}, \boldsymbol{w}_i^{(r)}, \boldsymbol{\Theta}_{i,1})\pi(\boldsymbol{h}_i^{(r)}, \boldsymbol{w}_i^{(r)})}{f(\boldsymbol{h}_i^{(r)}, \boldsymbol{w}_i^{(r)})}$$

for the importance function $f(\boldsymbol{h}_i, \boldsymbol{w}_i) = f(\boldsymbol{h}_i)f(\boldsymbol{w}_i|\boldsymbol{h}_i)$

- $f(\boldsymbol{h}_i)$ multivariate normal and $f(w_{i,t}|h_{i,t})$ inverse Gamma

Estimating $p(\boldsymbol{y}_i)$

- Outer importance sampling loop

$$\hat{p}(\boldsymbol{y}_i) = \frac{1}{R}\sum_{r=1}^{R} \frac{\hat{p}(\boldsymbol{y}_i|\boldsymbol{\Theta}_{i,1}^{(r)})\pi(\boldsymbol{\Theta}_{i,1}^{(r)})}{f(\boldsymbol{\Theta}_{i,1}^{(r)})}$$

for the importance function $f(\boldsymbol{\Theta}_{i,1}) = f(\boldsymbol{\theta}_{i,0})f(\sigma_{h,i}^2)f(\nu_i)$

- $f(\boldsymbol{\theta}_{i,0})$ multivariate normal, $f(\sigma_{h,i}^2)$ and $f(\nu_i)$ Gamma distributions

Back