# Large Mixed-Frequency VARs with a Parsimonious Time-Varying Parameter Structure

Thomas B. Götz and Klemens Hauzenberger (Deutsche Bundesbank)

Workshop on Forecasting

Frankfurt am Main, 8 September, 2017

## Introduction

To simultaneously consider mixed-frequency time series, their joint dynamics, and possible structural changes, we introduce a time-varying parameter mixed-frequency VAR (MF-VAR). To keep our approach from becoming too complex, we **implement time variation parsimoniously**: only the intercepts and a common factor in the error variances vary over time. We can therefore estimate moderately large systems in a reasonable amount of time making our modifications appealing for practical use. We further improve upon **computational efficiency** by exploiting gains from using sparse and block-banded matrix algebra as, for instance, in Chan (2015). Furthermore, we complement the standard Minnesota prior setting to long-run priors as in Giannone et al. (2016). For eleven U.S. variables, we examine the performance of our model (and two intermediate variants), **first and foremost with respect to GDP forecasting**, and compare them to the time-constant MF-VAR of Schorfheide and Song (2015). Our results demonstrate the feasibility and usefulness of our method.

## Motivating Example – Restricting Time Variation

- Our model lies "in-between" the time-constant MF-VAR of Schorfheide and Song (2015) and the fully fledged time-varying model of Cimadomo and D'Agostino (2015), thereby handling the trade-off between model flexibility and complexity
- For an $n = 3$-dimensional (GDP, CPI, fed funds rate) system we compare our model with a fully time-varying model $\Rightarrow n + 1$ instead of $n + n^2p + \frac{n(n+1)}{2}$ time-varying quantities ($p$ - lag length)

Table: Forecast Performance relative to a MF-VAR à la Primiceri (2005)

| Horizon | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rel. RMSE | 1.02 | 0.94 | 0.84 | 1.01 | 0.95 | 0.86 | 0.95 | 0.91 | 0.86 | 0.92 | 0.89 | 0.86 |
| Rel. logPL | -0.64 | -0.19 | 0.04 | 0.19 | 0.18 | 0.23 | 0.18 | 0.25 | 0.14 | 0.11 | 0.16 | -0.04 |

Note: -1 is a backcast, 0 and 1 are nowcasts, the remaining horizons refer to forecasts. More on forecast horizons later.

- Rarely worse performance than the fully fledged MF-VAR $\Rightarrow$ time variation in the data seems to be sufficiently covered by intercepts and (common) stochastic volatility

Table: Running Time relative to a MF-VAR à la Primiceri (2005)

| $p$ \ $n$ | 3 | 4 | 6 |
|---|---|---|---|
| 3 | 42.9 | 34.4 | 10.6 |
| 6 | 33.9 | 16.4 | |
| 12 | 25.0 | | |

- Blank entries indicate the unrestricted model to run into computational problems $\Rightarrow$ moderately large systems like we have in mind ($n = 10$ to $15$) become practically infeasible to run

## Our Modeling Framework – Two Blocks

- Distinguish between a monthly and a "quarterly" part:

$$\begin{bmatrix} x_{m,t} \\ x_{q,t} \end{bmatrix} = \begin{bmatrix} \Phi_{mm} & \Phi_{mq} & \Phi_{mc} \\ \Phi_{qm} & \Phi_{qq} & \Phi_{qc} \end{bmatrix} \begin{bmatrix} X_{m,t-1} \\ X_{q,t-1} \\ 1 \end{bmatrix} + \begin{bmatrix} \Phi_{mc,t} \\ \Phi_{qc,t} \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ u_{q,t} \end{bmatrix}$$

$$\begin{bmatrix} u_{m,t} \\ u_{q,t} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{mm,t} & \Sigma_{mq,t} \\ \Sigma_{qm,t} & \Sigma_{qq,t} \end{bmatrix} \right)$$

- $X_{m,t} = (x'_{m,t}, \ldots, x'_{m,t-p-1})'$ and $X_{q,t} = (x'_{q,t}, \ldots, x'_{q,t-p-1})'$ are $pn_m \times 1$ and $pn_q \times 1$ vectors (potentially latent observations; observable ones denoted $Y_{m,t}$ and $Y_{q,t}$)
- Compactly (i.e., stacked over $t$), $X = \Phi Z + \Phi_c^T + U$ with $u = \mathrm{vec}(U) \sim N(0, \Omega \otimes \Psi)$, where $\Psi$ governs "usual" cross-sectional and $\Omega$ serial structures (Chan, 2015) $\Rightarrow$ this **Kronecker structure** proves quite useful for computational efficiency
- **Block I** (dealing with the "incomplete" data given Block II):
  - For the balanced part of the sample, "reduced" state vector, $S_t = (x'_{q,t}, \ldots, x'_{q,t-p})'$, noting that $Y_{m,t} = X_{m,t}$:

$$S_t = \begin{bmatrix} \Phi_{qq} & 0 \\ I_{pn_q} & 0 \end{bmatrix} S_{t-1} + \begin{bmatrix} \Phi_{qm} \\ 0 \end{bmatrix} Y_{m,t-1} + \begin{bmatrix} \Phi_{qc,t} \\ 0 \end{bmatrix} + \begin{bmatrix} u_{q,t} \\ 0 \end{bmatrix} \text{ (state equation)}$$

$$\begin{bmatrix} y_{m,t} \\ y_{q,t} \end{bmatrix} = \begin{bmatrix} 0 & \Phi_{mq,1} & \Phi_{mq,2} & \Phi_{mq,3:p} \\ 1/3\, I_{n_q} & 1/3\, I_{n_q} & 1/3\, I_{n_q} & 0 \end{bmatrix} S_t + \begin{bmatrix} \Phi_{mm} \\ 0 \end{bmatrix} Y_{m,t-1} + \begin{bmatrix} \Phi_{mc,t} \\ 0 \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ 0 \end{bmatrix} \text{ (measurement equation)}$$

  - For the ragged edge (unbalanced) part, $S_t = (x'_t, \ldots, x'_{t-p})'$, where $x_t = (x'_{m,t}, x'_{q,t})'$

$$S_t = \begin{bmatrix} \Phi_{mm,1} & \Phi_{mq,1} & \cdots & \Phi_{mm,p} & \Phi_{mq,p} & 0 \\ \Phi_{qm,1} & \Phi_{qq,1} & \cdots & \Phi_{qm,p} & \Phi_{qq,p} & 0 \\ & & I_{p(n_m+n_q)} & & & 0 \end{bmatrix} S_{t-1} + \begin{bmatrix} \Phi_{mc} + \Phi_{mc,t} \\ \Phi_{qc} + \Phi_{qc,t} \\ 0 \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ u_{q,t} \\ 0 \end{bmatrix} \text{ (state equation)}$$

$$W_t^* \begin{bmatrix} y_{m,t} \\ y_{q,t} \end{bmatrix} = W_t^* \begin{bmatrix} I_{n_m} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3\, I_{n_q} & 0 & 1/3\, I_{n_q} & 0 & 1/3\, I_{n_q} & 0 \end{bmatrix} S_t \text{ (measurement equation)}$$

- **Block II** (estimating the VAR parameters given Block I)
  - **Time-varying intercepts** evolve as random walks: $(\Phi_{mc,t}, \Phi_{qc,t})' \equiv \Phi_{c,t} = \Phi_{c,t-1} + \nu_t, \quad \nu_t \sim N(0, Q)$
  - **Common stochastic volatilities** evolve as AR(1) process: $h_t = \rho h_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_h^2)$ and resulting model is nested in $u$ by setting $\Omega = \mathrm{diag}(\exp(h_1), \ldots, \exp(h_T))$

## Priors and Hyperparameters

- **Minnesota** (natural conjugate $MNIW$) **priors** for $\Phi$ and $\Psi$ implemented as in ...
  - Banbura et al., 2010 (including prior on sum of coefficients)
  - Giannone et al., 2016 (priors for the long run)
- An initial state $S_0$ is extracted by running the SSM (in a time-constant fashion) over a pre-sample (4 years)
- Prior for $\Phi_{c,1}$ is $N(0, Q_0^{-1})$; $Q \sim IG\left(Q_0, k_Q^2 \cdot Q_0 \cdot \Psi\right)$, where $k_Q$ controls the amount of time variation a-priori allowed for and $Q_0$ corresponds to the length of the pre-sample (as in Primiceri, 2005)
- $\rho \sim N(0.9, k_V^2)\, \mathbb{I}(|\rho| < 1)$; $\sigma_h^2 \sim IG(5, k_S(5-1))$ as in Chan (2015)
- We determine suitable values for the hyperparameters $\theta = (\lambda, k_Q, k_V, k_S)'$ by conducting a **preliminary analysis on a grid of values**
  - Choose the grid point that led to the **best GDP forecast performance** (this is our main goal) over the last $H$ months
  - In our application, we set $H = 48$ (i.e., 4 years), fix $k_V = 0.2$ as well as $k_S = 0.01$ (see Chan, 2015) and let

$$k_Q = \left\{\sqrt{0.001}, \sqrt{0.01}\right\}$$
$$\lambda = \{0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 1, 5\}$$

  - $\lambda$ controls the tightness of the Minnesota prior, i.e., the relative importance of the prior beliefs w.r.t. the information in the data

## Posterior Analysis

- Conditional posterior distributions for $S^T$, $Q$, $\rho$, $\sigma_h^2$ and $h^T$ are fairly standard
- We sample $\Psi$ marginally, and then $\mathrm{vec}(\Phi)|\Psi$:

$$\Psi|\Phi_c^T, h^T, S_q^T, S^T, y^T \sim IW\left(\hat{\Psi}, \nu_0 + T\right) \text{ and } \mathrm{vec}(\Phi)|\Psi, \Phi_c^T, h^T, S_q^T, S^T, y^T \sim N\left(\mathrm{vec}(\hat{\Phi}), K_\Phi^{-1} \otimes \Psi\right)$$

- Conventionally drawing $\Phi$ involves (i) the inverse of the $np + 1$ square matrix $K_\Phi$, and (ii) the Choleski decomp. of the $(np+1)n$ square matrix $K_\Phi^{-1} \otimes \Psi$
- Using $D \sim MN(0_{m\times n}, I_m \otimes I_n) \Rightarrow C_K'^{-1} D C_\Psi \sim MN\left(0_{m\times n}, (C_K'^{-1} I_m C_K^{-1}) \otimes (C_\Psi' I_n C_\Psi)\right)$, we can **draw $\Phi$ more efficiently** as

$$\Phi = \hat{\Phi} + C_K'^{-1} D C_\Psi,$$

where $\hat{\Phi} = (C_K C_K')^{-1} (V_{\Phi_0}^{-1}\Phi_0 + Z\Omega^{-1}X') = C_K' \backslash \left(C_K \backslash \left(V_{\Phi_0}^{-1}\Phi_0 + Z\Omega^{-1}\tilde{X}'\right)\right)$

- For $\Phi_c^T$, rewrite $\Phi_{c,t} = \Phi_{c,t-1} + \nu_t$ by stacking over $t$

$$\underbrace{\begin{bmatrix} I_n & & & \\ -I_n & I_n & & \\ & \ddots & \ddots & \\ & & -I_n & I_n \end{bmatrix}}_{H} \mathrm{vec}(\Phi_c^T) = \mathrm{vec}(\nu^T) \left[\sim N(0, I_T \otimes Q)\right],$$

and premultiply by $H^{-1}$ to get the prior $\mathrm{vec}(\Phi_c^T)|Q \sim N\left(0, (H'(I_T \otimes Q)^{-1}H)^{-1}\right)$ leading to the posterior $\Phi_c^T|\Phi, \Psi, h^T, Q, S_q^T, S^T, y^T \sim N(\hat{\Phi}_c^T, P_{\Phi_c}^{-1})$; a similar "trick" as for $K_\Phi$ and $\Psi$ then applies to $P_{\Phi_c}$ enabling us to **efficiently compute** $C_{P_\Phi}$ and, thereby, $\hat{\Phi}_c^T$

## Empirical Analysis – Data & Setup

- Three models to be compared with the benchmark MF-VAR of Schorfheide and Song (2015):
  - The "full" one including time-varying intercepts (TVi) & common stochastic volatility (CSV), labelled **TVi-MF-VAR-CSV**
  - The "intermediate" models including only one of the two features, i.e., **TVi-MF-VAR** and **MF-CSV-CSV**

Table: Data and Stylized Release Calendar

| Series ID | Transf. | Pub-Lag | Description |
|---|---|---|---|
| GDP | Log | 1 Qrt | Real gross domestic product, sa |
| INV | Log | 1 Qrt | Real gross private domestic investment, sa |
| GOV | Log | 1 Qrt | Real government consumption expenditures and gross investment, sa |
| UNR | 1/100 | 1 Mth | Civilian unemployment rate, sa |
| HRS | Log | 1 Mth | Index of aggregate weekly hours, sa |
| CPI | Log | 1 Mth | Consumer price index for all urban consumers, sa |
| IPI | Log | 1 Mth | Industrial production index, sa |
| PCE | Log | 1 Mth | Personal consumption expenditures index, sa |
| FF | 1/100 | ./. | Effective federal funds rate, nsa |
| TB | 1/100 | ./. | 10-year treasury constant maturity rate, nsa |
| SP500 | Log | ./. | S&P 500 stock index, adjusted close price, nsa |

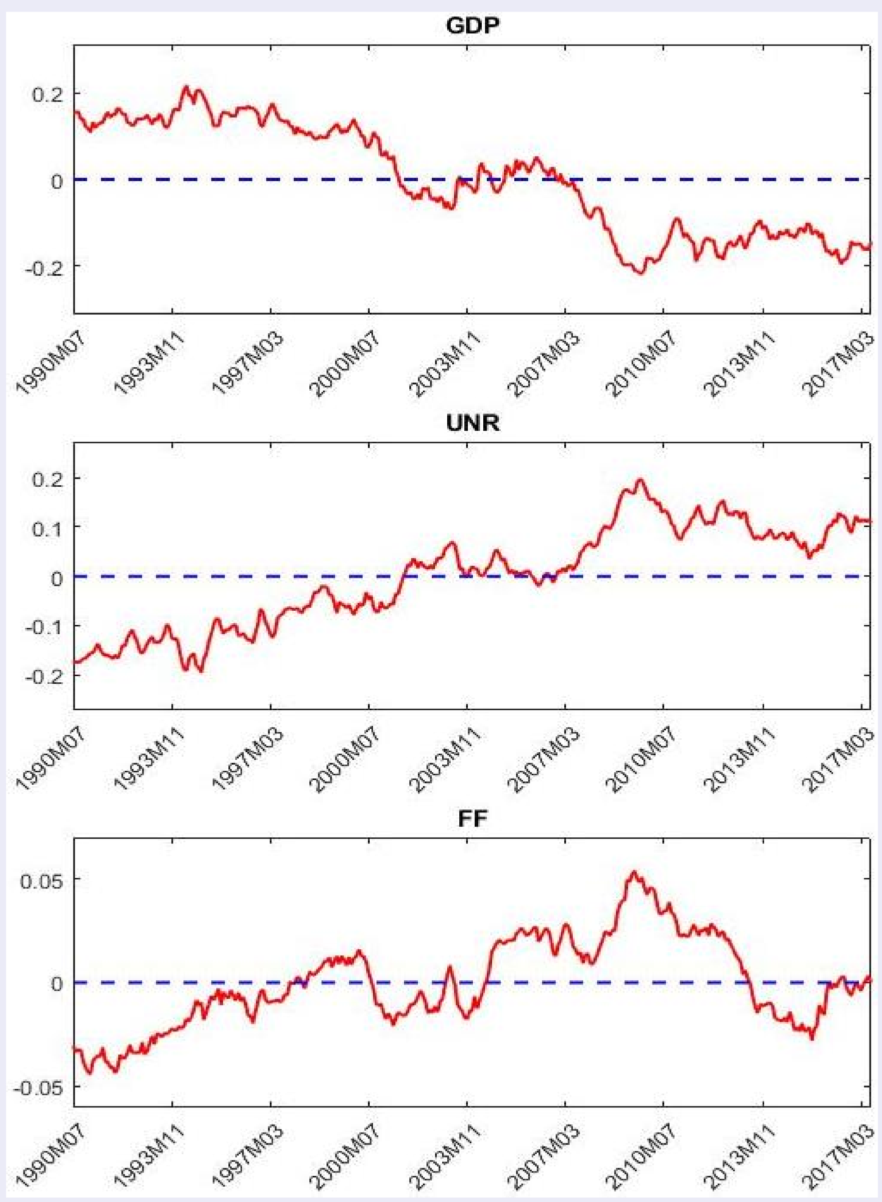
Note: GDP, INV and GOV are available on a quarterly basis, whereby second releases are considered; Data downloaded on the June 1st, 2017.

- $p = 6$ monthly lags; 1986-1989 used as pre-sample; generally, 10000 draws as burn-in, 2000 retained; forecasts assumed to be computed **at the end of each month**

## Full Sample (1990:M1-2017:M7) Estimation

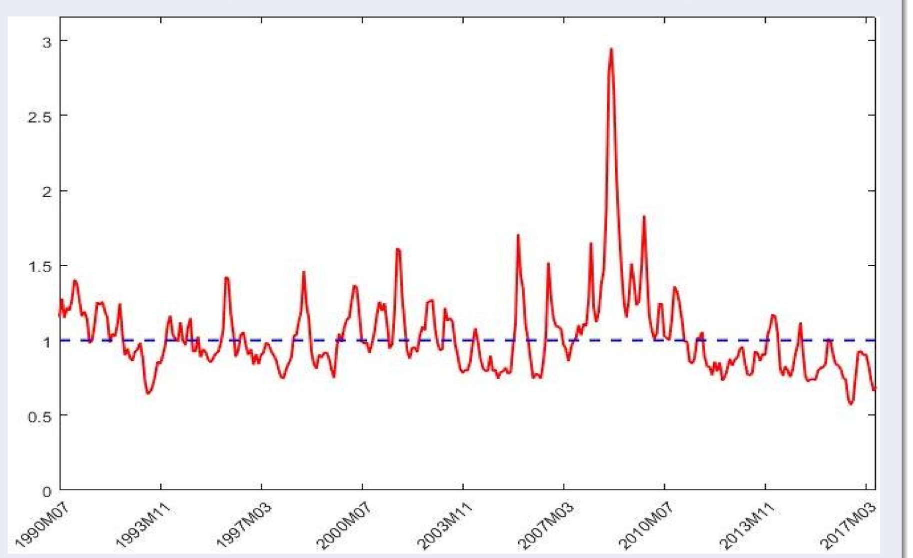
- **Focus on GDP**; also consider the monthly targets **UNR** and **FF** whenever possible
- Conditional means: not fluctuations in the (long-run) growth rates; rather, **changes in the intercept's contribution between two points in time**

Figure: Time-Varying Intercepts



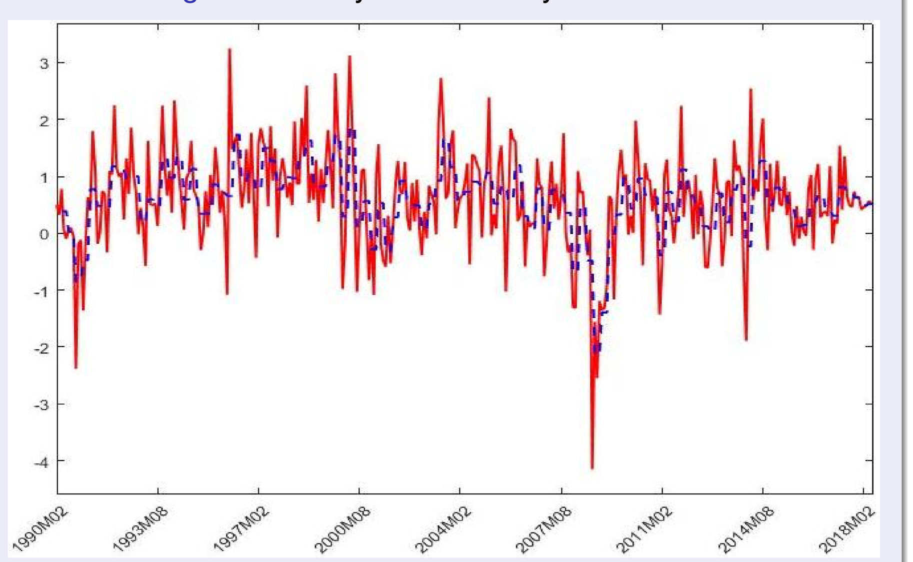

Note: Demeaned medians of the time-varying intercepts; multiplied by 100

Figure: Common Stochastic Volatility



Note: Medians of the common stochastic volatility in standard deviations

Figure: Monthly vs. Quarterly GDP Growth



Note: Growth rates of the medians of monthly GDP (including forecasts); scaled by 3 to make them comparable to quarter-on-quarter rates

- Much more movement in the monthly series due to pronounced (common) stochastic volatility
- Computation time: 41min vs. 26min.

## Forecast Evaluation Exercise

- **Increasing sequence of est. samples**: 1990:M1-2006:M5, . . ., 1990:M1-2017:M7; **pseudo real-time**; up to twelve forecasts for each indicator; laws of motion for TVi and CSV imposed over forecast horizon
- Forecast horizon: amount of months between a forecast is made and the end of the reference period
  - Thus, for quarterly series $h_Q = -1, 0, \ldots, 10$ with $h_Q = -1$ a backcast, $h_Q = 0, 1, 2$ nowcasts and $h_Q \geq 3$ forecasts
  - For the monthly series, publication delays are shorter s.t. $h_M \geq 0$ for some series (e.g., IPI) or $h_M \geq 1$ for others (e.g., FF)
- Relative (to MF-VAR) RMSE's and logPL's considered to **assess point and density forecast accuracy**
- Only results with standard Minnesota priors displayed (work somewhat better than long-run priors)

Table: Forecast performance relative to the MF-VAR of Schorfheide and Song (2015)

| | Horizon | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | GDP | | | | | |
| RMSE's | TVi-MF-VAR | 0.98 | 0.99 | 1.07 | 0.99 | 0.98 | 1.07 | 0.95 | 0.97 | 1.04 | 0.92 | 0.93 | 0.99 |
| | MF-VAR-CSV | 0.96 | 0.93 | 0.96 | 0.94 | 0.94 | 0.97 | 0.97 | 0.98 | 1.00 | 0.99 | 0.99 | 1.01 |
| | TVi-MF-VAR-CSV | 0.92 | 0.89 | 0.95 | 0.95 | 0.91 | 1.00 | 0.95 | 0.94 | 0.99 | 0.96 | 0.94 | 0.99 |
| logPL's | TVi-MF-VAR | 0.06 | 0.15 | -0.10 | 0.03 | 0.22 | 0.14 | 0.06 | 0.34 | 0.23 | 0.08 | 0.16 | 0.08 |
| | MF-VAR-CSV | 0.13 | 0.19 | 0.09 | 0.03 | 0.21 | 0.22 | 0.02 | 0.30 | 0.19 | 0.01 | 0.07 | 0.05 |
| | TVi-MF-VAR-CSV | 0.16 | 0.22 | 0.10 | 0.06 | 0.25 | 0.23 | 0.06 | 0.37 | 0.26 | 0.10 | 0.18 | 0.13 |
| | | | | | | | | UNR | | | | | |
| RMSE's | TVi-MF-VAR | | 0.92 | 0.89 | 0.87 | 0.85 | 0.83 | 0.85 | 0.86 | 0.87 | 0.87 | 0.87 | 0.87 |
| | MF-VAR-CSV | | 1.05 | 1.09 | 1.12 | 1.12 | 1.11 | 1.08 | 1.06 | 1.05 | 1.04 | 1.02 | 1.01 |
| | TVi-MF-VAR-CSV | | 0.98 | 0.97 | 0.98 | 0.96 | 0.93 | 0.91 | 0.90 | 0.88 | 0.87 | 0.86 | 0.85 |
| logPL's | TVi-MF-VAR | | 0.07 | 0.10 | 0.15 | 0.24 | 0.35 | 0.50 | 0.80 | 1.17 | 1.24 | 1.29 | 1.21 |
| | MF-VAR-CSV | | -0.02 | -0.01 | 0.03 | 0.13 | 0.25 | 0.47 | 0.65 | 0.94 | 0.71 | 0.82 | 0.75 |
| | TVi-MF-VAR-CSV | | 0.02 | 0.06 | 0.15 | 0.29 | 0.42 | 0.67 | 0.94 | 1.42 | 1.66 | 1.86 | 1.85 |
| | | | | | | | | FF | | | | | |
| RMSE's | TVi-MF-VAR | | | 0.94 | 0.95 | 0.95 | 0.96 | 0.96 | 0.95 | 0.95 | 0.95 | 0.96 | 0.97 |
| | MF-VAR-CSV | | | 1.01 | 0.99 | 0.99 | 0.98 | 0.97 | 0.96 | 0.96 | 0.96 | 0.96 | 0.97 |
| | TVi-MF-VAR-CSV | | | 0.95 | 0.93 | 0.93 | 0.93 | 0.92 | 0.91 | 0.91 | 0.91 | 0.92 | 0.93 |
| logPL's | TVi-MF-VAR | | | 0.09 | 0.11 | 0.08 | 0.09 | 0.14 | 0.08 | 0.13 | 0.06 | 0.14 | 0.14 |
| | MF-VAR-CSV | | | 0.34 | 0.40 | 0.36 | 0.07 | -0.24 | -0.26 | -0.39 | -0.45 | -0.53 | -0.50 |
| | TVi-MF-VAR-CSV | | | 0.37 | 0.41 | 0.55 | 0.48 | 0.52 | 0.50 | 0.50 | 0.42 | 0.46 | 0.53 |

## References

- Banbura, M., Giannone, D. and Reichlin, L. (2010), "Large Bayesian vector auto regressions". *Journal of Applied Econometrics*, 25: 71-92.
- Chan, J. C. (2015), "Large Bayesian VARs: a flexible Kronecker error covariance structure". Mimeo, Australian National University.
- Cimadomo, J. and D'Agostino, A. (2016), "Combining time variation and mixed frequencies: an analysis of government spending multipliers in Italy". *Journal of Applied Econometrics*, 31(7): 1276-1290.
- Giannone, D., Lenza, M. and Primiceri, G. E. (2016), "Priors for the long run". CEPR Discussion Paper No. 11261.
- Primiceri, G. E. (2005), "Time varying structural vector autoregressions and monetary policy", *Review of Economic Studies*, 72, 821-852.
- Schorfheide, F. and Song, D. (2015), "Real-time forecasting with a mixed-frequency VAR", *Journal of Business & Economic Statistics*, 33(3), 366-380.