A Flexible Bayesian MIDAS Approach Interpretable Nowcasting and Forecasting

Galina Potjagailo[1], David Kohns[2]

[1]Bank of England
[2]Heriot-Watt University, European Central Bank, Centre for Energy Economics Research and Policy (CEERP)

Conference on Real-Time Data Analysis, Methods, and Applications
Federal Reserve Bank of Cleveland
October 6, 2022

Disclaimer: The views expressed in this presentation are those of the authors, and not necessarily those of the Bank of England nor the European Central Bank. All errors and omissions are ours.

Motivation & What we do

- Nowcast and forecast models face challenges since Covid-19 shock
  - Large observations can distort parameters and increase uncertainty
  - Changing correlation structure among macro indicators
  - Survey indicators less informative for GDP during the pandemic
- Require flexible model features
  - Accounting for time-varying trends and stochastic volatilities (SV) beneficial in UC, DFM, VAR models (Stock and Watson, 2009; Clark, 2011; Antolin-Diaz et al., 2017)
  - Covid-19: Account for extreme observations via t-distr. errors or outliers (Carriero et al., 2021; Lenza and Primiceri, 2022; Antolin-Diaz et al., 2021)
- We combine such flexible features with
  - a multivariate Mixed Data Sampling (MIDAS) regression
  - a flexible group-shrinkage prior that allows for flexible variable selection and signal communication

Proposed Model

We propose the Trend-SV-t-BMIDAS model with three features

1. Time-varying unobserved components in the lower-frequency target variable (time-varying Trend, SV, t-distr. errors)
2. Timely information from high frequency indicators in a multivariate MIDAS block
3. Bayesian shrinkage via a group-global-local prior with three tiers of continuous shrinkage (overall, between indicators, and within lags of an indicator)

We enhance the prior with a sparsification step

This imposes variable selection and helps interpret signals over time via inclusion probabilities.

Links to the literature

- The model nests / compares to existing models when shutting down model features
  - multivariate BMIDAS with horseshoe (Kohns and Bhattacharjee, 2022) or spike-and-slab group shrinkage prior (Mogliani and Simoni, 2021) asymptotically
  - BMIDAS model with SV (Carriero et al., 2015)
  - Trend-SV-outl. DFM (Antolin-Diaz et al., 2021)

- MIDAS literature (Ghysels et al., 2007, 2020; Foroni et al., 2015)
- Global-local shrinkage priors (Polson and Scott, 2010; Polson et al., 2014; Carvalho et al., 2010) and group-lasso priors (Casella et al., 2010; Xu and Ghosh, 2015) and spike-and-slab (Ishwaran et al., 2005; Piironen et al., 2017)

Results

- For a nowcast application to UK GDP growth, we show that combining all features improves nowcasts relative to alternatives that shut down the Trend and/or SV-t, or use a prior without group shrinkage
- The shrinkage prior helps selecting a sparse group of the most informative indicators over nowcast periods (a few survey ind., then "hard" production and services ind.)
    - Covid-19 pandemic: move towards signals from indicators for services, away from production surveys, reflecting shifts in spending related to lockdowns
    - the model performs better during the Covid-19 period compared to "dense" specifications such as a DFM or model without group-shrinkage

Methodology

The Trend-BMIDAS-SV-t Model

$$y_t = \tau_t + \theta' Z_{t-h}^{(m)} + \sqrt{\lambda_t} e^{\frac{1}{2}(h_0 + w_h \tilde{h}_t)} \tilde{\epsilon}_t^y,$$
$$\tilde{\epsilon}_t^y \sim N(0,1), \ \lambda_t \sim IG(\nu/2, \nu/2) \tag{1}$$

$$\tau_t = \tau_{t-1} + e^{\frac{1}{2}(g_0 + w_g \tilde{g}_t)} \tilde{\epsilon}_t^\tau, \ \tilde{\epsilon}_t^\tau \sim N(0,1)$$
$$\tilde{h}_t = \tilde{h}_{t-1} + \tilde{\epsilon}_t^h, \ \tilde{\epsilon}_t^h \sim N(0,1) \tag{2}$$
$$\tilde{g}_t = \tilde{g}_{t-1} + \tilde{\epsilon}_t^g, \ \tilde{\epsilon}_t^g \sim N(0,1).$$

- τ_t: time-varying trend; τ_t and y_t are lower frequency (quarterly)
- $Z_t^{(m)} = (z_{1,t}^{(m)}, \cdots, z_{K,t}^{(m)})'$, i.e. m indicators (monthly) between $t-1$ and t for each indicator K.
  - θ: $(p_k + 1) * K$ parameters that link higher and lower frequency observations.
  - MIDAS with Almon lag polynomial restrictions, but U-MIDAS works too Almon
- h_t, g_t: SVs for observation and trend follow non-centered random walks (Frühwirth-Schnatter and Wagner, 2010)
- λ_t: enforces a ν-degrees of freedom t-distribution, fat-tailed SV

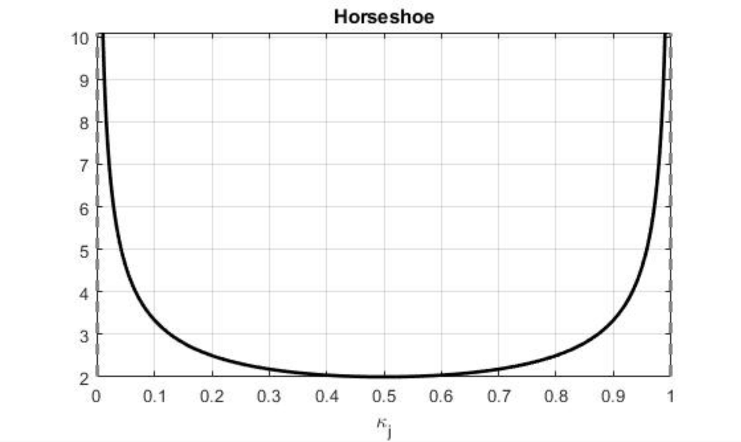
Group-shrinkage prior on multivariate MIDAS component

- GIGG (Group-Inverse-Gamma-Gamma) prior (Boss et al., 2021) on MIDAS
  - Global-local prior
  - Accounts for group-shrinkage + correlation within higher frequency lags
- Each group g has $p_g + 1$ parameters to estimate

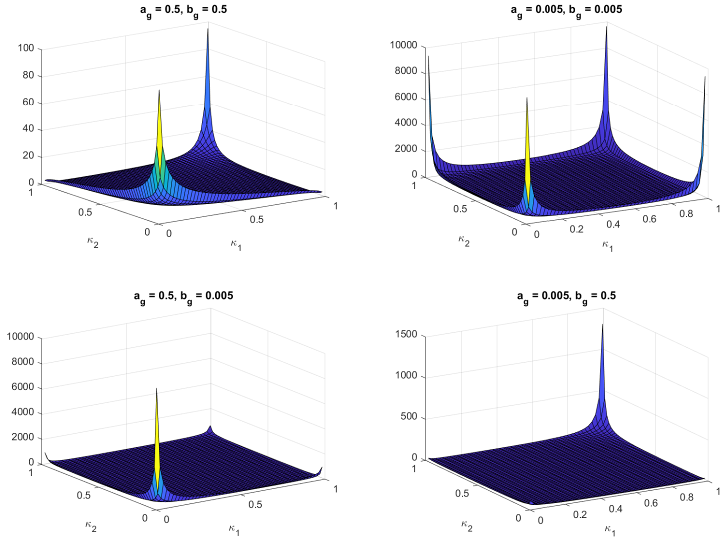
$$\begin{aligned} \theta_{g,i} &\sim N(0, \vartheta^2 \gamma_g^2 \varphi_{g,i}^2), \quad \forall i \in \{0, \cdots, p_g + 1\} \\ \vartheta &\sim C_+(0,1), \quad \gamma_g^2 | a_g \sim G(a_g, 1), \quad \varphi_{g,i}^2 | b_g \sim IG(b_g, 1), \end{aligned} \tag{3}$$

  ϑ controls the overall level of sparsity, γ_g^2 controls sparsity across groups g, $\varphi_{g,i}^2$ controls sparsity of group members i within g
  - $a_g < b_g \rightarrow$ stronger group-shrinkage, strong prior correlation among lags
  - $b_g < a_g \rightarrow$ stronger shrinkage within groups (individual lags selected)
  - At group-size 1, $a_g = b_g = 0.5$, give horseshoe prior (Carvalho et al., 2010)
- We set $g = k$, i.e. groups defined as lags of each indicator (can be extended to groupings across indicators if k large)

Univariate shrinkage with global-local prior



GIGG Prior Visualisation: Bi-variate shrinkage hyperparamers



Sparsification step for GIGG prior

With continuous priors the posteriors of lag groups remain non-zero with probability one (Hahn and Carvalho, 2015). This hampers the interpretability of results.

Final Step: decision theoretically motivated sparsification method to the posterior on θ_g

- Achieved by minimising a utility function over the Euclidean distance between a linear model that penalises group-size akin to Zou (2006) and our model's prediction:

$$\mathcal{L}(\tilde{Y}, \alpha) = \frac{1}{2}||Z^{(m)}\alpha - \tilde{Y}||_2^2 + \sum_{k=1}^{K} \phi_k ||\alpha_k||_2, \tag{4}$$

- penalisation term creates a soft-thresholding effect between [-ϕ_k, ϕ_k], forcing the coefficients on all group members to zero.

The relative frequency of lag-group k in the sparsified estimate $\alpha^{*(s)}$ over all Gibbs draws gives inclusion probabilities that inform on the relative impact of an indicator.

Other priors and estimation

Priors for latent states standard: $(\tau, \tilde{h}, \tilde{g})$ joint normal prior as in Chan and Jeliazkov (2009), McCausland et al. (2011) and Kim et al. (1998)

Estimation via Metropolis-within-Gibbs sampler M-H Gibbs

- Recursive sampling from conditional distributions: MIDAS parameters θ, GIGG hyperparameters, latent states $(\tau, \tilde{h}, \tilde{g})$ (non-recursively, as in Chan and Jeliazkov (2009)), λ_t, degrees of freedom ν
- sampling of ν requires Metropolis step
- 5000 burn-in iterations, retain further 5000 for inference

Empirical Application

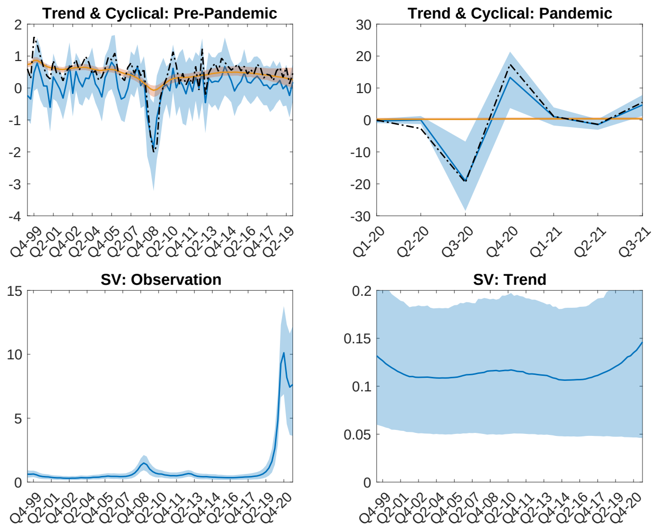
Nowcast quarterly UK GDP growth, 1999-2021

- Setup
  - In-Sample Start: Q1 1999, Nowcast Start: Q1 2011
  - Nowcast End: "pre-pandemic" Q4 2019, "including pandemic" Q3 2021
- Monthly indicators
  - indices of services and production, trade
  - surveys (CBI, PMI, GfK)
  - labour market (unemployment rate, employment, vacancies, hours)
  - mortgage approvals, VISA consumer spending
- Nowcast evaluation
  - pseudo-real-time calendar: 20 nowcasts per quarter around data releases calendar
  - each nowcast has new information set, uses the latest 6 monthly observations (3 contemporaneous, 3 lagged) of that indicator available at time of nowcast
- Metrics
  - Point: Root-mean-squared forecast error (RMSFE)
  - Density: Average cumulative rank probability score (CRPS)

Results

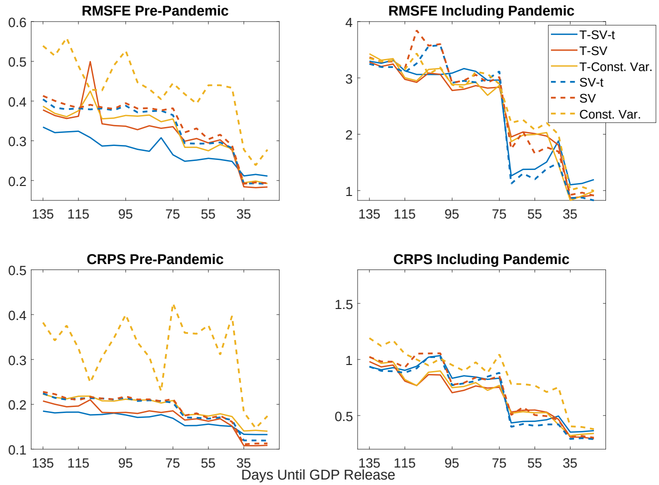
- Trend-SVt-BMIDAS model gives posterior estimates for time-varying trend, cyclical component, and stochastic volatilities of GDP growth and trend
    - we show them estimated with data covering full sample period
    - sensitivity of posterior estimates to other model specifications (no Trend, const. Var.) and over the data release cycle - Trend-SVt model most robust
- Nowcast evaluation compared to model alternatives
    - Shutting down time-varying components (No Trend, SV without t-distr. errors, homoskedastic model)
    - No group shrinkage: horseshoe prior
    - Alternatives to BMIDAS: Combined univariate MIDAS, mixed-frequency DFM (both with time-varying trend and SV-t)

Posterior estimates for trend, cyclical component and stochastic volatilities.

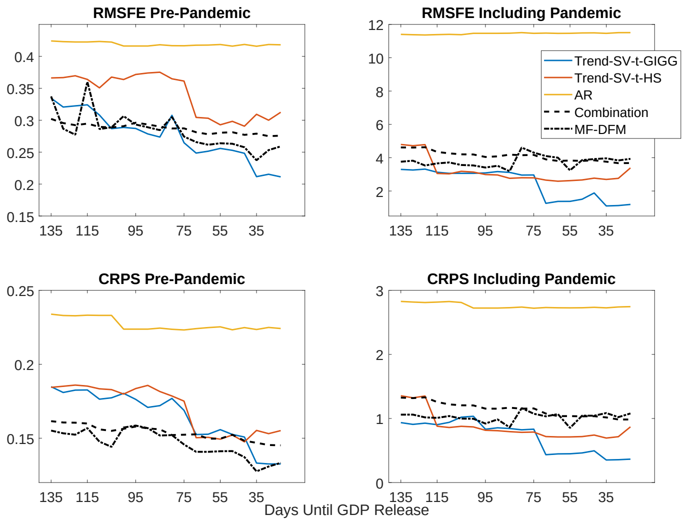


sensitivity

Shutting down time-var. Trend and/or SV-t



Performance against other prior and alternatives to multivariate MIDAS

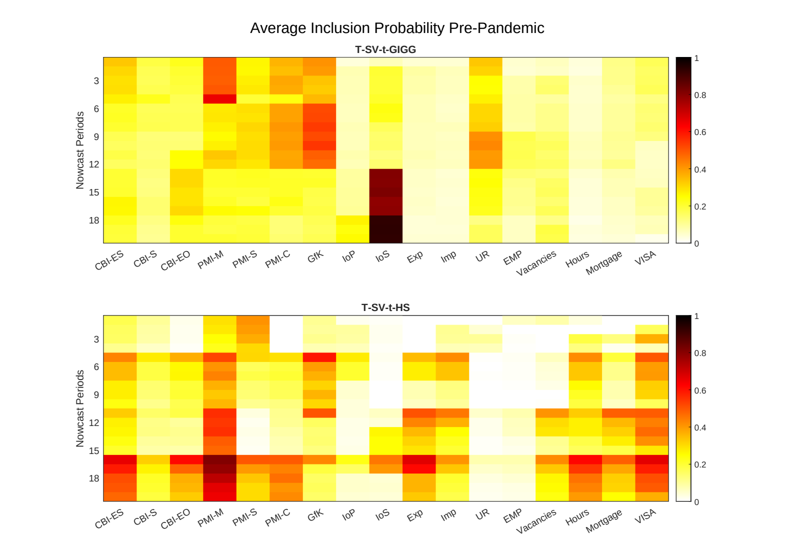


Which signals does the model rely upon?

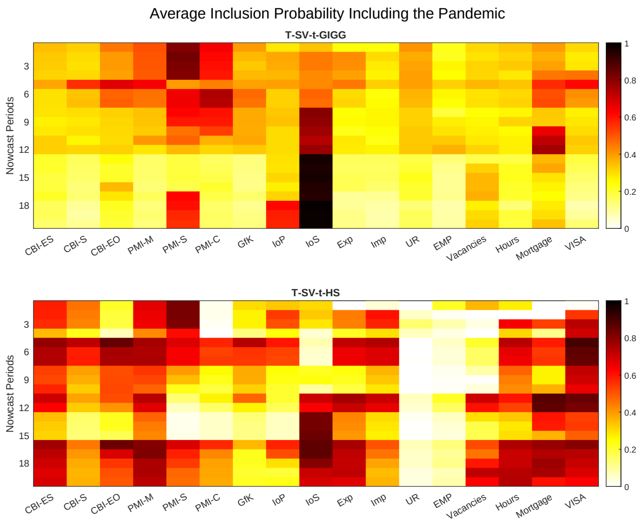
- Results so far:
  - Trend-SV-t BMIDAS performs competitively against various alternatives where some of the model features are shut down
  - Strong performance for late nowcast periods when "hard" indicators available
  - During the Covid-19 pandemic period, the model picks up the recovery earlier and updates nowcasts about the initial trough earlier on

- Which signals does the model exploit over the data release cycle?
  - The sparsification step on the GIGG prior allows us to derive inclusion probabilities for each indicator over time and over the data release cycle
  - We show inclusion probabilities prior to the Covid-19 pandemic and including the pandemic, compared to prior without group shrinkage (Horseshoe prior)

Average posterior inclusion probabilities, until 2019



Average posterior inclusion probabilities, including pandemic



Conclusions

- We have brought time-varying trends, volatilities and accounts for large observations to the Bayesian MIDAS framework.
  - A flexible group-shrinkage prior regularises the MIDAS.
  - Sparsification step allows for variable selection which facilitates interpretability.
- Results bring new impetus to the debate on density vs sparsity in macroeconomic forecasting (Giannone et al., 2021)
  - a sparse specification performs strongly: grouping + time variation important
  - during Covid-19: signals mainly from service indicators + mortgage approvals
- Framework can be applied to a range of time series exercises where group shrinkage can be relevant, can be useful for disaggregated data (e.g. prices, labour market data).

Thank you

Thank you

Contact: galina.potjagailo@bankofengland.co.uk

Appendix

Almon lag polynomial restricted MIDAS

- U-MIDAS (Foroni et al., 2015) involves many parameters and can lead to erratic weight profiles
- Restrict coefficients via Almon lag-polynomials on θ_i: assuming a $p_k << L_k$ polynomial process of the coefficients across high-frequency observations

Almon Lag MIDAS

Assume lags $i = 0, \cdots, L$ can be represented by a 3rd degree polynomial, then each HF parameter process, θ_i can be written as:

$$\theta_i = \beta_0 + \beta_1 i + \beta_2 i^2 + \beta_3 i^3 \tag{5}$$

We add economically relevant restrictions (Smith and Giles, 1976)

$$\begin{aligned} \theta'_L &= 0 \\ \theta_L &= 0 \end{aligned} \tag{6}$$

<u>But</u>: Smoothness of Almon-polynomial increases parameter correlation back

Metropolis-within-Gibbs sampling algorithm

1. Sample $\theta|\bullet \sim p(\theta|\mathbf{y}, \bullet)$
2. Sample hyper-parameters $\vartheta, \gamma_k^2, \varphi_{kj}^2, \nu_p$ in one block
   2.1 $\vartheta^2 \sim p(\vartheta^2|\mathbf{y}, \bullet)$
   2.2 $\gamma_k^2 \sim 1/p(\gamma_k^{-2}|\mathbf{y}, \bullet)$
   2.3 $\varphi_{kj}^2 \sim p(\varphi_{kj}^2|\mathbf{y}, \bullet)$
3. sample $\tilde{\boldsymbol{\tau}} \sim p(\tilde{\boldsymbol{\tau}}|\mathbf{y}, \bullet)$ and $\tau_0 \sim p(\tau_0|\mathbf{y}, \bullet)$
4. sample $\tilde{\mathbf{h}} \sim p(\tilde{\mathbf{h}}|\mathbf{y}, \bullet)$, $h_0 \sim p(h_0|\mathbf{y}, \bullet)$ and $w_h \sim p(w_h|\mathbf{y}, \bullet)$
5. sample $\tilde{\mathbf{g}} \sim p(\tilde{\mathbf{g}}|\mathbf{y}, \bullet)$, $g_0 \sim p(h_0|\mathbf{y}, \bullet)$ and $\sim p(w_g|\mathbf{y}, \bullet)$
6. Sample $\{\lambda_t\}_{t=1}^T \sim p(\lambda_t|\mathbf{y}, \bullet)$
7. Sample $\nu \sim p(\nu|\mathbf{y}, \bullet)$ with a Metropolis step back

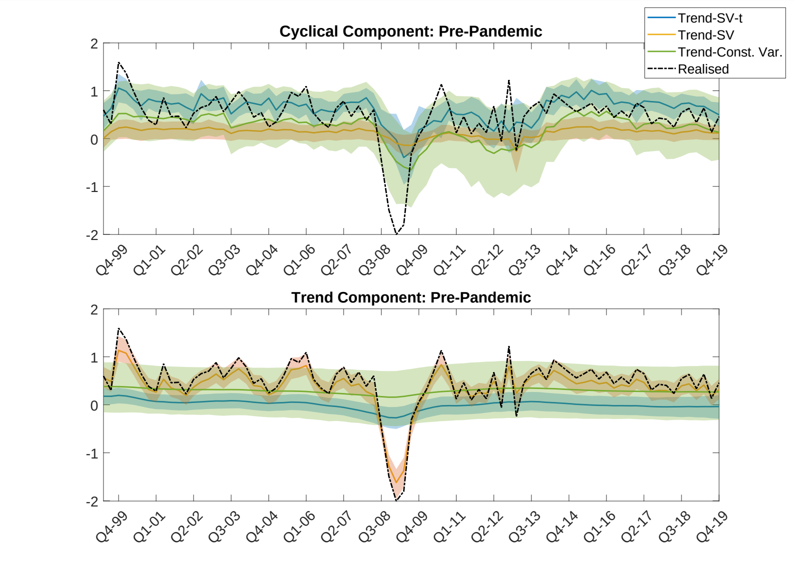
- sampling technique of Chan and Jeliazkov (2009) allows drawing steps 3.-5. in a non-recursive fashion which increases efficiency and can be sped up using sparse-matrices

Pseudo Real Time Calendar for UK Nowcast Application

Nowcast	Quarter	Days to GDP	Month	Timing within month	Release	Publication Lag
1		135	1	1st of month	PMIs	m-1
2		125	1	End of 2nd week	IoP, IoS, Ex, Im	m-2
3		120	1	3rd week	Labour market data	m-2
4		115	1	3rd Friday of month	Mortgage & Visa	m-1
5		110	1	End of 3rd week	CBIs & GfK	m
6	Reference	105	2	1st of month	PMIs	m-1
7	quarter	97	2	Mid of 2nd week	Quarterly GDP	q-1
8	(nowcast)	95	2	End of 2nd week	IoP, IoS, Ex, Im	m-2
9		90	2	3rd week	Labour market data	m-2
10		85	2	3rd Friday of month	Mortgage & Visa	m-1
11		80	2	End of 3rd week	CBIs & GfK	m
12		75	3	1st of month	PMIs	m-1
13		65	3	End of 2nd week	IoP, IoS, Ex, Im	m-2
14		60	3	3rd week	Labour market data	m-2
15		55	3	3rd Friday of month	Mortgage & Visa	m-1
16		50	3	End of 3rd week	CBIs & GfK	m
17		45	1	1st of month	PMIs	m-1
18	Subsequent	35	1	End of 2nd week	IoP, IoS, Ex, Im	m-2
19	quarter	30	1	3rd week	Labour market data	m-2
20	(backcast)	25	1	3rd Friday of month	Mortgage & Visa	m-1

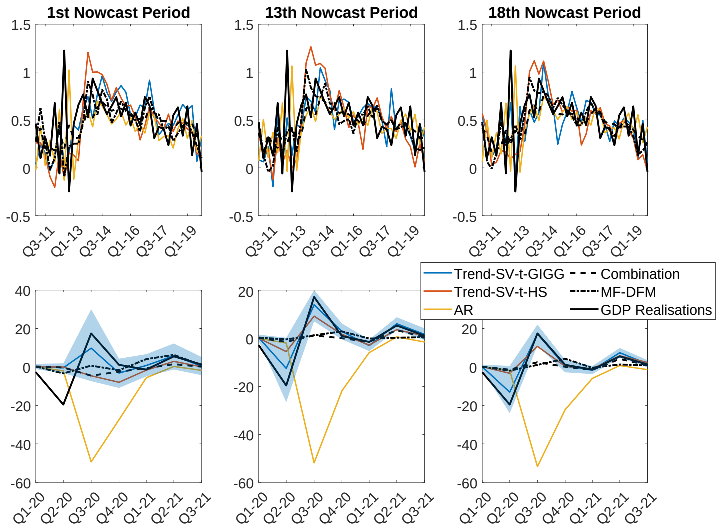
back

Trend and SV posterior estimates, alternative models.



back

Posterior mean and density nowcasts, selected nowcast periods.



References

Antolin-Diaz, J., Drechsel, T., and Petrella, I. (2017). Tracking the slowdown in long-run GDP growth. Review of Economics and Statistics, 99(2):343–356.

Antolin-Diaz, J., Drechsel, T., and Petrella, I. (2021). Advances in nowcasting economic activity: Secular trends, large shocks and new data. CEPR Discussion Paper No. DP15926.

Boss, J., Datta, J., Wang, X., Park, S. K., Kang, J., and Mukherjee, B. (2021). Group Inverse-Gamma Gamma Shrinkage for Sparse Regression with Block-Correlated Predictors. arXiv preprint arXiv:2102.10670.

Carriero, A., Clark, T. E., and Marcellino, M. (2015). Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility. Journal of the Royal Statistical Society. Series A,(Statistics in Society), 178(4):837.

Carriero, A., Clark, T. E., Marcellino, M. G., and Mertens, E. (2021). Addressing COVID-19 outliers in BVARs with stochastic volatility. CEPR Discussion Paper No. DP15964.

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2010). The horseshoe estimator for sparse signals. Biometrika, 97(2):465–480.

Casella, G., Ghosh, M., Gill, J., and Kyung, M. (2010). Penalized regression, standard errors, and Bayesian lassos. Bayesian Analysis, 5(2):369–411.

Chan, J. C. and Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models. International Journal of Mathematical Modelling and Numerical Optimisation, 1(1-2):101–120.

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

Foroni, C., Marcellino, M., and Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. Journal of the Royal Statistical Society: Series A (Statistics in Society), 178(1):57–82.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for gaussian and partial non-gaussian state space models. Journal of Econometrics, 154(1):85–100.

Ghysels, E., Kvedaras, V., and Zemlys-Balevičius, V. (2020). Mixed data sampling (midas) regression models. In Handbook of Statistics, volume 42, pages 117–153. Elsevier.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). MIDAS regressions: Further results and new directions. Econometric reviews, 26(1):53–90.

Giannone, D., Lenza, M., and Primiceri, G. E. (2021). Economic predictions with big data: The illusion of sparsity. Econometrica, 89(5):2409–2437.

Hahn, P. R. and Carvalho, C. M. (2015). Decoupling shrinkage and selection in bayesian linear models: a posterior summary perspective. Journal of the American Statistical Association, 110(509):435–448.

Ishwaran, H., Rao, J. S., et al. (2005). Spike and slab variable selection: frequentist and Bayesian strategies. The Annals of Statistics, 33(2):730–773.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. The Review of Economic Studies, 65(3):361–393.

Kohns, D. and Bhattacharjee, A. (2022). Nowcasting growth using google trends data: A bayesian structural time series model. International Journal of Forecasting.

Lenza, M. and Primiceri, G. E. (2022). How to estimate a vector autoregression after March 2020. Journal of Applied Econometrics.

McCausland, W. J., Miller, S., and Pelletier, D. (2011). Simulation smoothing for state–space models: A computational efficiency analysis. Computational Statistics & Data Analysis, 55(1):199–212.

Mogliani, M. and Simoni, A. (2021). Bayesian MIDAS penalized regressions: estimation, selection, and prediction. Journal of Econometrics, 222(1):833–860.

Piironen, J., Vehtari, A., et al. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors. Electronic Journal of Statistics, 11(2):5018–5051.

Polson, N. G. and Scott, J. G. (2010). Shrink globally, act locally: Sparse Bayesian regularization and prediction. Bayesian Statistics, 9:501–538.

Polson, N. G., Scott, J. G., and Windle, J. (2014). The Bayesian bridge. Journal of the Royal Statistical Society: Series B: Statistical Methodology, pages 713–733.

Smith, R. G. and Giles, D. E. (1976). The Almon estimator: Methodology and users' guide. Reserve Bank of New Zealand.

Stock, J. H. and Watson, M. W. (2009). Forecasting in dynamic factor models subject to structural instability. The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry, 173:205.

Xu, X. and Ghosh, M. (2015). Bayesian variable selection and estimation for group lasso. Bayesian Analysis, 10(4):909–936.

Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American Statistical Association, 101(476):1418–1429.