Forecasting with Bayesian non-parametric methods

Massimiliano MARCELLINO[1] *and* Michael PFARRHOFER[2]

(1) Bocconi University, IGIER & CEPR (2) University of Vienna & WU Vienna

12th ECB Conference on Forecasting Techniques
Frankfurt, June 12–13 2023

Introduction

- This paper is part of a research project on the use of classical and Bayesian non-parametric approaches to handle parameter time-variation and non-linearities in econometric models, possibly of large dimension.
- Non-parametric approaches have a long tradition in statistics and econometrics, but up to a few years ago their empirical application was uncommon, mainly due to computational issues, particularly so for large models.
- They are typically less efficient but more robust than parametric methods, requiring fewer (though still some) assumptions.
- Earlier papers on estimation, then also forecasting and structural analysis.

Introduction

Classical non-parametric approach (kernel based), relies on and extends work by Giraitis, Kapetanios and Yates:

- (2019) "Large Time-Varying Parameter VAR: A Non-Parametric Approach," with George Kapetanios and Fabrizio Venditti, Journal of Applied Econometrics, 34(7), 1027-1049.
- (2021) "Time-Varying Instrumental Variable Estimation," with Liudas Giraitis and George Kapetanios, Journal of Econometrics, 224(2), 394-415.
- (2022) "Estimation and inference in large heterogeneous panels with stochastic time-varying coefficients," with Yu Bai and George Kapetanios.
- (2023) "The time-varying 3PRF," with Yiannis Dendramis and George Kapetanios.
- (2023) "Time Varying IV-SVARs and the effects of monetary policy on financial variables," with Robin Braun and George Kapetanios.

Introduction

Bayesian non-parametric approach relies on various methods “imported” from statistics: BART, GP, BNN:

- (2023) “Tail forecasting with multivariate Bayesian additive regression trees,” with Todd Clark, Florian Huber, Gary Koop and Michael Pfarrhofer, International Economic Review, forthcoming.
- (2022) “Investigating growth at risk using a multi-country non-parametric quantile factor model,” with Todd Clark, Florian Huber, Gary Koop and Michael Pfarrhofer, arXiv 2110.03411.
- (2022) “Forecasting US Inflation Using Bayesian Nonparametric Models,” with Todd Clark, Florian Huber and Gary Koop, Cleveland FED WP WP 22-05.
- (2022) “Gaussian Process Vector Autoregressions and Macroeconomic Uncertainty,” with Niko Hauzenberger, Florian Huber and Nico Petz, CEPR WP 17646.

Introduction

- (202X) "Nonparametric methods for measuring asymmetries in monetary policy transmission," with Florian Huber, Michael Pfarrhofer and Anna Stelzer.
- (202X) "Unusual weather in unusual economic times," with Florian Huber, Tamás Krisztin and Leopold Ringwald.
- (202X) "Oil in the tails," with Christiane Baumeister and Florian Huber.
- (2023) "Enhanced Bayesian neural networks for macroeconomics and finance," with Niko Hauzenberger, Florian Huber and Karin Kleiber.

Introduction

- In this paper we review specification and estimation of Bayesian nonparametric models for forecasting (possibly large sets of) macroeconomic and financial variables, focusing on Bayesian Additive Regression Trees (BART) and Gaussian Processes (GP). This involves introducing flexible specifications for both the conditional mean functions and variances.
- We then apply various versions of these models for point, density and tail forecasting using a real time dataset for the euro area (Banbura *et al.*, 2021).
- The performance is compared with that of several variants of Bayesian VARs to assess the relevance of accounting for non-standard, general forms of nonlinearities and parameter time variation.

Econometric framework

- We work with general multivariate models of the form:

$$\boldsymbol{y}_t = F(\boldsymbol{x}_t) + \boldsymbol{\epsilon}_t \tag{1}$$

  $\boldsymbol{y}_t$ is an $n \times 1$ vector, and $\boldsymbol{x}_t = (\boldsymbol{y}'_{t-1}, \ldots, \boldsymbol{y}'_{t-p})'$ is $k = np \times 1$ collecting p lags

  - $F(\boldsymbol{x}_t) = (f_1(\boldsymbol{x}_t), \ldots, f_n(\boldsymbol{x}_t))'$ are (possibly unknown) conditional mean functions $f_i(\boldsymbol{x}_t) : \mathbb{R}^k \rightarrow \mathbb{R}$ for $i = 1, \ldots, n$, such that $F(\boldsymbol{x}_t) : \mathbb{R}^k \rightarrow \mathbb{R}^n$
  - $\boldsymbol{\epsilon}_t$ reflects the unpredictable component; it is specified such that we can estimate our model equation-by-equation

- Models will be distinguished with respect to how we treat the function $F(\boldsymbol{x}_t)$ and what we assume about the error term $\boldsymbol{\epsilon}_t$

Benchmark vector autoregression (VAR)

- The simplest case is to restrict $F(\boldsymbol{x}_t)$ to linearity, i.e., $F(\boldsymbol{x}_t) = \boldsymbol{A}_t \boldsymbol{x}_t$:

$$\boldsymbol{y}_t = \boldsymbol{A}_t \boldsymbol{x}_t + \boldsymbol{\epsilon}_t$$

  $\boldsymbol{A}_t$ are $n \times k$ (possibly time-varying) VAR coefficients

- `VAR`: $F(\boldsymbol{x}_t) = \boldsymbol{A}\boldsymbol{x}_t$, constant for all $t = 1, \ldots, T$,
- `TVP`: $F(\boldsymbol{x}_t) = \boldsymbol{A}_t \boldsymbol{x}_t$ combined with state equation for $\boldsymbol{a}_t = \text{vec}(\boldsymbol{A}_t)$

$$\boldsymbol{a}_t = \boldsymbol{a}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\boldsymbol{0}_{nk}, \boldsymbol{\Theta})$$

  with a diagonal matrix $\boldsymbol{\Theta}$ collecting state innovation variances — these govern the amount of time-variation

Bayesian additive regression trees (BART)

- Sum-of-tree model introduced in Chipman *et al.* (2010)
- Equation-specific functions $f_i(\boldsymbol{x}_t)$ are approximated using BART:

$$f_i(\boldsymbol{x}_t) \approx \sum_{s=1}^{S} l_{is}(\boldsymbol{x}_t | \mathcal{T}_{is}, \boldsymbol{\mu}_{is})$$

  $l_{is}(\boldsymbol{x}_t | \mathcal{T}_{is}, \boldsymbol{\mu}_{is})$ are individual tree functions (explaining small fractions)
  - Tree structures $\mathcal{T}_{is}$ and terminal nodes $\boldsymbol{\mu}_{is}$
  - BART approximates the function $f_i(\boldsymbol{x}_t)$ summing over S trees

- A simple example with scalar dependent $\boldsymbol{y} = (y_1, \ldots, y_T)'$, a single predictor $\boldsymbol{x} = (x_1, \ldots, x_T)'$ and shocks $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_T)'$

Bayesian additive regression trees (BART)

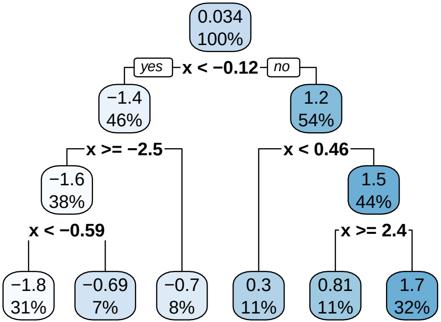
- Considering a single tree for full data matrices:

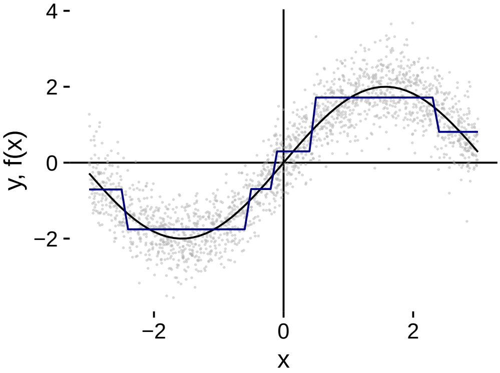
$$\boldsymbol{y} = l(\boldsymbol{x}|\mathcal{T}, \boldsymbol{\mu}) + \boldsymbol{\epsilon}, \quad E(\boldsymbol{y}|\boldsymbol{x}) = l(\boldsymbol{x}|\mathcal{T}, \boldsymbol{\mu}) = \sum_{k=1}^{\#\mathrm{TN}} \mu_k \mathbb{I}(\boldsymbol{x} \in \mathcal{S}_k)$$

- The tree is defined by sets $\mathcal{T} = \{\mathcal{S}_k\}_{k=1}^{\#\mathrm{TN}}$
- Sets $\mathcal{S}_k$ formed by partitioning $\boldsymbol{x}$ via splitting rules: $\{\boldsymbol{x} \leq c\}$ or $\{\boldsymbol{x} > c\}$
- Values of the function over t are associated terminal nodes μ_k

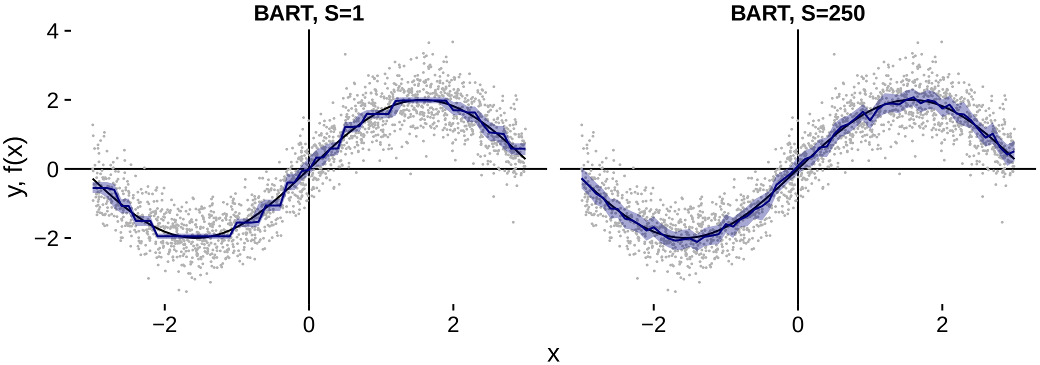
- BART sums over many such trees, and we have many predictors in practice — splitting rules (variables/thresholds) estimated alongside terminal nodes

Fitting data using a single regression tree





Fitting data using BART



Gaussian process (GP) regression

- GP prior on the conditional mean functions (see, e.g., Williams and Rasmussen, 2006):

$$f_i(\boldsymbol{x}_t) \sim \mathcal{GP}\left(0, \mathcal{K}_{\boldsymbol{\vartheta}_i}(\boldsymbol{x}_t, \boldsymbol{x}_t)\right)$$

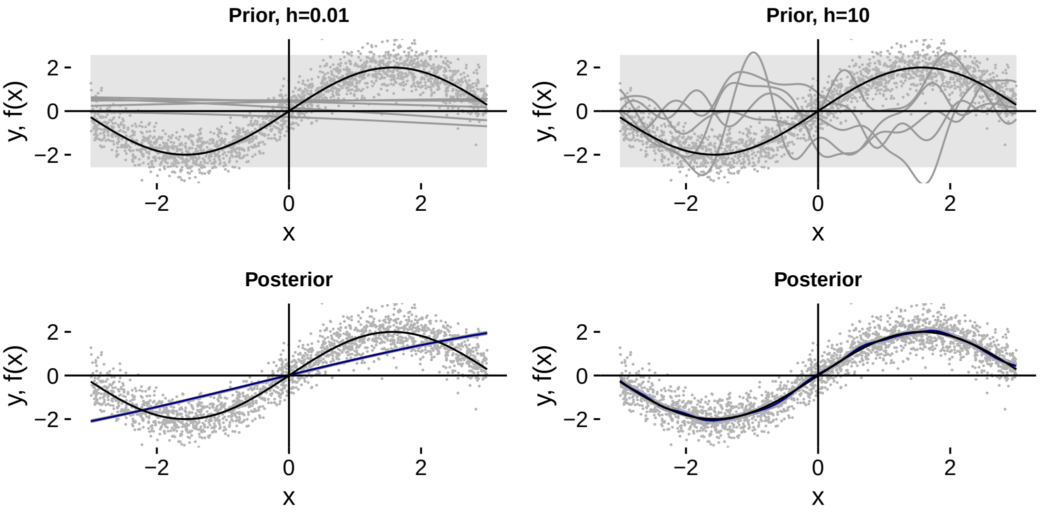
- Using $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T)'$, this prior becomes a (finite) multivariate Gaussian:

$$\boldsymbol{f}_i \sim \mathcal{N}\left(\boldsymbol{0}_T, \mathcal{K}_{\boldsymbol{\vartheta}_i}(\boldsymbol{X}, \boldsymbol{X}')\right), \quad \boldsymbol{f}_i = (f_i(\boldsymbol{x}_1), \ldots, f_i(\boldsymbol{x}_T))'$$

- Kernel $\mathcal{K}_{\boldsymbol{\vartheta}_i}(\boldsymbol{X}, \boldsymbol{X}')$ with typical $(t, \tilde{t})$ element $\mathcal{K}_{\boldsymbol{\vartheta}_i}(\boldsymbol{x}_t, \boldsymbol{x}_{\tilde{t}})$
  - Distance based, $d_{t\tilde{t}} = ||\boldsymbol{x}_t - \boldsymbol{x}_{\tilde{t}}||^2$ capturing similarities in input space
  - Hyperparameters $\boldsymbol{\vartheta}_i = (\xi_i, l_i)'$ regulate properties of permitted functions
- We use a squared exponential Kernel (others are available)

$$\mathcal{K}_{\boldsymbol{\vartheta}_i}\left(\boldsymbol{x}_t, \boldsymbol{x}_{\tilde{t}}\right) = \xi_i \times \exp\left(-l_i d_{t\tilde{t}}/2\right)$$

Gaussian process (GP) regression



Conditional variances

- Few common sources drive the reduced form shocks (i.e., primitive shocks)
- Variants of factor stochastic volatility (FSV,g Aguilar and West, 2000)

$$\boldsymbol{\epsilon}_t = \boldsymbol{L}\mathfrak{F}_t + \boldsymbol{\eta}_t$$

  $\mathfrak{F}_t$ are $q \times 1$ latent factors linked to observed space by $n \times q$ loadings matrix $\boldsymbol{L}$ plus idiosyncratic noise $\boldsymbol{\eta}_t = (\eta_{1t}, \ldots, \eta_{nt})'$

- We assume heteroskedastic factors

$$\mathfrak{F}_t \sim \mathcal{N}(\mathbf{0}_q, \boldsymbol{\Omega}_t), \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{H}_t),$$
$$\boldsymbol{\Omega}_t = \text{diag}\left(\exp(\omega_{1t}), \ldots, \exp(\omega_{qt})\right), \quad \boldsymbol{H}_t = \text{diag}\left(\exp(h_{1t}), \ldots, \exp(h_{nt})\right)$$

Conditional variances

- Stochastic volatility (SV) processes capturing dynamics of volatility

$$\omega_{it} = \phi_{i\omega}\omega_{it-1} + \varsigma_{i\omega}\xi_{it}, \quad \xi_{it} \sim \mathcal{N}(0,1), \quad \text{for } i = 1, \ldots, q,$$
$$h_{jt} = \mu_{jh} + \phi_{jh}(h_{jt-1} - \mu_{jh}) + \varsigma_{jh}\xi_{jt}, \quad \xi_{jt} \sim \mathcal{N}(0,1), \quad \text{for } j = 1, \ldots, n.$$

- Note that $\text{Var}(\boldsymbol{\epsilon}_t) = \text{Var}(\boldsymbol{L}\mathfrak{F}_t + \boldsymbol{\eta}_t) = \boldsymbol{L}\boldsymbol{\Omega}_t\boldsymbol{L}' + \boldsymbol{H}_t$
- Order-invariance of the multivariate system (see also Chan *et al.*, 2022)

- Conditionally independent equations $i = 1, \ldots, n$:

$$\tilde{\boldsymbol{y}}_t = \boldsymbol{y}_t - \boldsymbol{L}\mathfrak{F}_t, \quad \tilde{y}_{it} = f_i(\boldsymbol{x}_t) + \eta_{it}, \quad \eta_{it} \sim \mathcal{N}\left(0, \exp(h_{it})\right)$$

Summary of the models

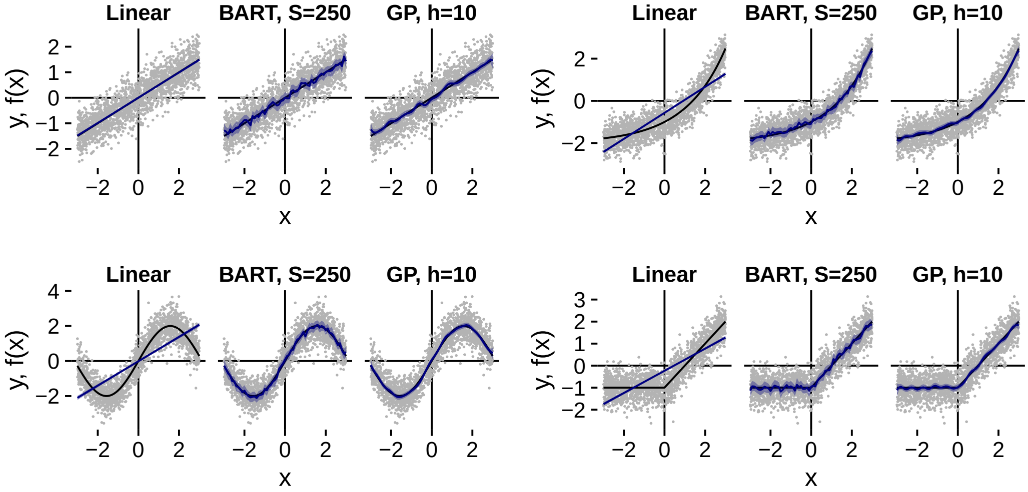
- General multivariate models of the form:

$$\boldsymbol{y}_t = F(\boldsymbol{x}_t) + \boldsymbol{L}\mathfrak{F}_t + \boldsymbol{\eta}_t, \quad \mathfrak{F}_t \sim \mathcal{N}(\mathbf{0}_q, \boldsymbol{\Omega}_t), \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{H}_t)$$

- We consider four options for $F(\boldsymbol{x}_t) = (f_1(\boldsymbol{x}_t), \ldots, f_n(\boldsymbol{x}_t))'$
  1. **Linear/TVP**: $F(\boldsymbol{x}_t) = \boldsymbol{A}_t\boldsymbol{x}_t$, horseshoe shrinkage
  2. **BART**: $f_i(\boldsymbol{x}_t) \approx \sum_{s=1}^{S} l_{is}(\boldsymbol{x}_t|\mathcal{T}_{is}, \boldsymbol{\mu}_{is})$
  3. **GP**: $f_i(\boldsymbol{x}_t) \sim \mathcal{GP}\left(0, \mathcal{K}_{\boldsymbol{\theta}_i}(\boldsymbol{x}_t, \boldsymbol{x}_t)\right)$

- Homoskedastic by assuming $\boldsymbol{H}_t = \boldsymbol{H}$ and $\boldsymbol{\Omega}_t = \boldsymbol{I}_q$, labeled "`hom`"
- Heteroskedastic case (stochastic volatility) indicated by "`sv`," e.g., "`Linear sv`" is a standard BVAR with SV

How do the competitors fit data?



Sketch of the estimation algorithm

Standard MCMC algorithm (Gibbs sampling, Metropolis-Hastings steps); conditional updates for ...

1. $F(\boldsymbol{x}_t)$ equation-by-equation conditional on the factors;
2. any hyperparameters related to $F(\boldsymbol{x}_t)$, e.g., shrinkage for linear/TVP, Kernel hyperparameters for GP, etc.;
3. $\boldsymbol{L}$ and $\mathfrak{F}_t$ conditional on each other, $\mathfrak{F}_t$ in one block;
4. log-volatility processes in $\boldsymbol{H}_t$ and $\boldsymbol{\Omega}_t$ conditional on $\boldsymbol{\eta}_t$ and $\boldsymbol{f}_t$;
5. missing values in $\boldsymbol{y}_t$ due to the release calendar
6. iterative multi-step-ahead forecasts via Monte Carlo to account for all sources of uncertainty, yields $\boldsymbol{y}_{T+h}$

Details on step 1/2: $F(x_t)$ for BVAR/TVP

- Global-local shrinkage on the elements in A; or in non-centered TVP parameterization, on initial state A_0 and square root of the state innovation variances $\sqrt{\mathbf{\Theta}}$
- For some generic coefficient β_j, this prior is given by:

$$\beta_j \sim \mathcal{N}(0, \lambda_j^2 \tau^2), \quad \lambda_j \sim \mathcal{C}^+(0,1), \quad \tau \sim \mathcal{C}^+(0,1)$$

- Coefficients and state innovation variances can be sampled equation-by-equation (as can the TVPs via standard methods)
  - Conditionally conjugate posteriors for all required quantities
  - Global/local shrinkage factors can be updated in one block

Details on step 1: $F(\boldsymbol{x}_t)$ for BART

- Tree generating stochastic process that serves as a prior
  - Probability that a node at depth $d = 1, \ldots,$ is nonterminal: $\alpha/(1+d)^{\beta}$ for $\alpha \in (0, 1)$ and $\beta \in \mathbb{R}^{+}$ (we use default values of Chipman *et al.*, 2010)
  - Discrete uniform prior on variables that appear in splitting rules
  - Discrete uniform prior (on range of splitting variable) for thresholds
- Independent Gaussian priors on the terminal node parameters for tree $s = 1, \ldots, S$, in equation $i = 1, \ldots, n$:

$$\mu_{is,k} \sim \mathcal{N}(0, \phi_i), \text{ for } k = 1, \ldots, \#\mathrm{TN}$$

- Informativeness of this prior is set in a data-based fashion:

$$\sqrt{\phi_i} = \frac{\max(\boldsymbol{y}_i) - \min(\boldsymbol{y}_i)}{2\gamma\sqrt{S}}$$

Details on step 1: $F(\boldsymbol{x}_t)$ for BART

- Individual trees sampled conditional on remaining $S-1$ trees, marginal of the terminal node parameters $\boldsymbol{\mu}_{is}$, see Chipman *et al.* (2010)
  - Keeps dimensionality of the problem fixed
  - Metropolis-Hastings updates of trees, with four transitions: (1) grow or (2) prune tree, (3) change splitting rule, (4) swap splitting rule
- Gaussian posteriors of terminal nodes (conditional on tree structures)

Details on step 1/2: $F(x_t)$ for GP

- Conditional on the FSV-related quantities, we have:

$$\tilde{\boldsymbol{y}}_i = \boldsymbol{f}_i + \boldsymbol{\eta}_i, \quad \boldsymbol{f}_i \sim \mathcal{N}(\boldsymbol{0}_T, \mathcal{K}_{\boldsymbol{\vartheta}_i}), \quad \boldsymbol{\eta}_i \sim \mathcal{N}(\boldsymbol{0}_T, \boldsymbol{H}_i)$$

  where all quantities are for equation $i = 1, \ldots, n$, and stacked over t, and $\boldsymbol{H}_i = \mathrm{diag}(\exp(h_{i1}), \ldots, \exp(h_{iT}))$
- Posterior of $\boldsymbol{f}_i | \tilde{\boldsymbol{y}}_i \sim \mathcal{N}(\overline{\boldsymbol{f}}_i, \overline{\boldsymbol{V}}_i)$ is a multivariate Gaussian

$$\overline{\boldsymbol{f}}_i = \mathcal{K}_{\boldsymbol{\vartheta}_i}(\mathcal{K}_{\boldsymbol{\vartheta}_i} + \boldsymbol{H}_i)^{-1}\tilde{\boldsymbol{y}}_i$$
$$\overline{\boldsymbol{V}}_i = \mathcal{K}_{\boldsymbol{\vartheta}_i} - \mathcal{K}_{\boldsymbol{\vartheta}_i}(\mathcal{K}_{\boldsymbol{\vartheta}_i} + \boldsymbol{H}_i)^{-1}\mathcal{K}_{\boldsymbol{\vartheta}_i}$$

- Predictive distribution takes a similar form

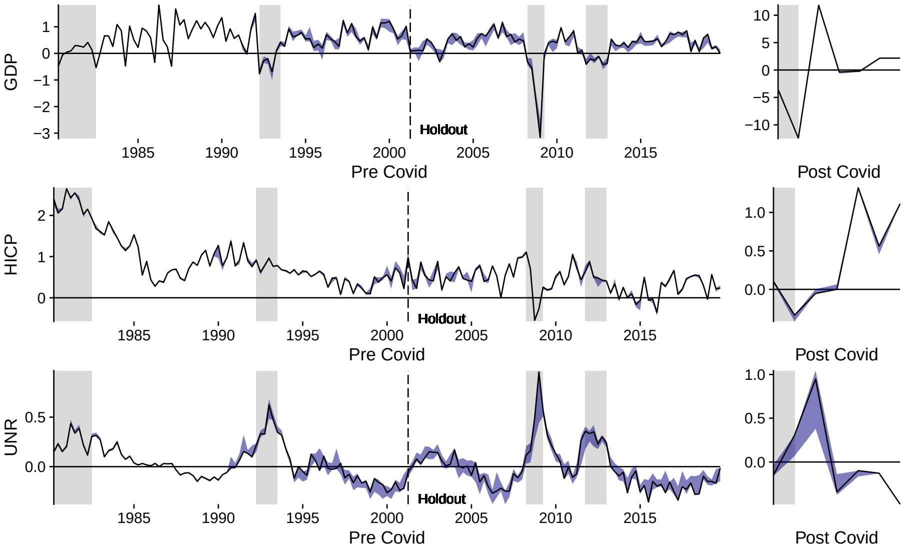
Details on step 1/2: $F(\boldsymbol{x}_t)$ for GP

- Conditional likelihood $\tilde{\boldsymbol{y}}_i \sim \mathcal{N}(\mathbf{0}_T, \mathcal{K}_{\boldsymbol{\vartheta}_i} + \boldsymbol{H}_i)$ is used to update hyperparameters $\boldsymbol{\vartheta}_i = (\xi_i, h_i)'$
- Metropolis-Hastings updates for $\xi_i \sim \mathcal{G}(a_\xi, b_\xi)$ and $l_i \sim \mathcal{G}(a_l, b_l)$
  - Specific to equation i, i.e., flexibility regarding type of dependent
  - Asymmetric log-normal proposals, straightforward acceptance probabilities

Real time forecasts for the EA

- Series start in 1980Q2 and we use $p = 4$ lags with small $(n = 3)$ information sets
- History of vintages from 2001Q1 to 2021Q3
  - First evaluation period 2001Q2
  - Final observation 2021Q3
- Evaluated using final vintage data
- *Target variables*: QoQ GDP growth (`GDP`), QoQ headline inflation (`HICP`), unemployment rate (`UNR`, in differences)
- Conditional nowcast scheme to sample ragged edges online
- *Benchmark*: Linear BVAR with homoskedastic errors

Real time vintages (dataset and size of revisions)



Predictive loss functions

- Let $y^{(r)}_{i,t+h}$ be the realization and $y^{(fp)}_{i,t+h}$ is the pth quantile of the predictive distribution
- Quantile score (QS) as in Giacomini and Komunjer (2005)

$$\mathrm{QS}(p)_{i,t+h} = \left(y^{(r)}_{i,t+h} - y^{(fp)}_{i,t+h}\right) \times \left(p - \mathbb{I}\left(y^{(r)}_{i,t+h} < y^{(fp)}_{i,t+h}\right)\right),$$

  e.g., QS(0.5) is the (half) mean absolute error (point forecast metric)

- Quantile-weighted continuous ranked probability score (CRPS), as in Gneiting and Ranjan (2011)

$$\mathrm{CRPS}_{i,t+h}(\mathfrak{w}_p) = \int_0^1 \mathfrak{w}_p \, \mathrm{QS}(p)_{i,t+h} \, \mathrm{d}p,$$

- Weights $\mathfrak{w}_p$ to target different parts of the distribution

Estimation times

- `Linear-hom` takes about 30 seconds for 10k posterior and predictive draws (Macbook Air M1, 2020)

Specification	**Relative time**
Linear hom	1.00
Linear sv	1.09
BART hom	1.18
BART sv	1.28
TVP sv	1.84
GP hom	2.39
GP sv	2.49

Forecast results

$h = 1$

		None					Left					Right					QS50				
		Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion
GDP	TVP SV	0.92	0.99	0.88	1.00	0.86	0.98	1.02	0.96	1.01	0.95	0.85	0.95	0.80	0.98	0.79	0.92	1.00	0.87	0.99	0.87
GDP	Linear SV	**0.82**	**0.98**	**0.72**	1.01	**0.68**	0.86	1.02	0.77	1.03	**0.68**	**0.78**	**0.95**	**0.69**	0.98	**0.69**	**0.80**	0.99	**0.68**	1.00	**0.67**
GDP	GP SV	0.83	0.99	0.74	0.99	0.72	0.86	1.00	0.78	0.99	0.71	0.80	0.98	0.70	0.97	0.72	0.81	**0.98**	0.71	0.98	0.70
GDP	GP hom	0.82	0.99	0.72	**0.95**	0.72	**0.85**	1.01	**0.76**	**0.96**	0.72	0.79	0.97	0.70	**0.94**	0.73	0.81	0.98	0.71	**0.94**	0.73
GDP	BART SV	0.83	1.00	0.74	1.01	0.71	0.87	1.02	0.78	1.02	0.71	0.80	0.99	0.70	0.99	0.72	0.81	0.99	0.70	1.00	0.68
GDP	BART hom	0.94	1.00	0.90	0.99	0.90	0.94	**1.00**	0.90	0.99	0.88	0.94	1.00	0.91	0.99	0.92	0.93	0.99	0.89	0.99	0.89
HICP	TVP SV	0.99	1.02	0.90	1.00	0.99	1.01	1.03	0.94	1.01	1.02	0.98	1.01	0.87	1.00	0.97	0.99	1.00	0.91	0.99	0.99
HICP	Linear SV	0.99	1.00	0.92	1.00	0.98	1.00	1.01	0.93	1.00	1.00	0.97	**0.99**	0.90	1.00	0.97	0.98	0.99	0.93	1.00	0.98
HICP	GP SV	0.92	1.00	0.60	0.99	0.90	**0.93**	**0.98**	0.70	**0.97**	0.92	0.91	1.03	**0.50**	1.02	0.88	0.92	0.97	0.66	**0.98**	0.90
HICP	GP hom	**0.91**	**0.99**	**0.59**	**0.99**	**0.89**	0.93	0.99	**0.67**	0.99	**0.91**	**0.90**	1.01	0.50	1.00	**0.87**	**0.91**	**0.96**	**0.66**	0.98	**0.89**
HICP	BART SV	0.97	1.03	0.72	1.02	0.95	0.98	1.02	0.78	1.00	0.97	0.96	1.04	0.67	1.04	0.94	0.95	1.00	0.77	1.03	0.93
HICP	BART hom	0.97	1.02	0.76	1.01	0.95	0.96	1.00	0.78	0.99	0.96	0.97	1.04	0.73	1.04	0.96	0.97	1.01	0.80	1.01	0.96
UNR	TVP SV	0.95	0.99	0.82	0.97	0.94	0.93	1.00	0.75	0.97	**0.92**	0.96	0.99	0.89	0.97	0.96	0.93	0.97	0.78	0.96	0.91
UNR	Linear SV	0.98	1.03	0.83	0.98	0.97	0.97	1.04	0.76	0.97	0.97	0.99	1.02	0.89	0.99	0.98	0.96	1.01	0.80	0.97	0.96
UNR	GP SV	0.91	0.94	0.81	0.85	0.93	0.92	0.98	0.72	0.87	0.94	0.90	0.91	0.88	0.84	0.93	0.89	0.93	0.78	0.84	0.92
UNR	GP hom	0.90	0.94	**0.78**	**0.83**	0.93	**0.91**	0.97	**0.71**	**0.84**	0.93	0.90	0.92	**0.85**	**0.83**	0.93	0.89	0.92	**0.78**	**0.81**	0.92
UNR	BART SV	**0.89**	**0.90**	0.87	0.87	**0.91**	0.91	0.93	0.83	0.87	0.92	**0.89**	**0.88**	0.91	0.87	**0.90**	**0.87**	**0.88**	0.84	0.86	**0.88**
UNR	BART hom	0.92	0.90	0.97	0.87	0.94	0.94	**0.92**	0.99	0.88	0.96	0.91	0.89	0.96	0.87	0.93	0.90	0.89	0.94	0.87	0.92

Forecast results

$h = 4$

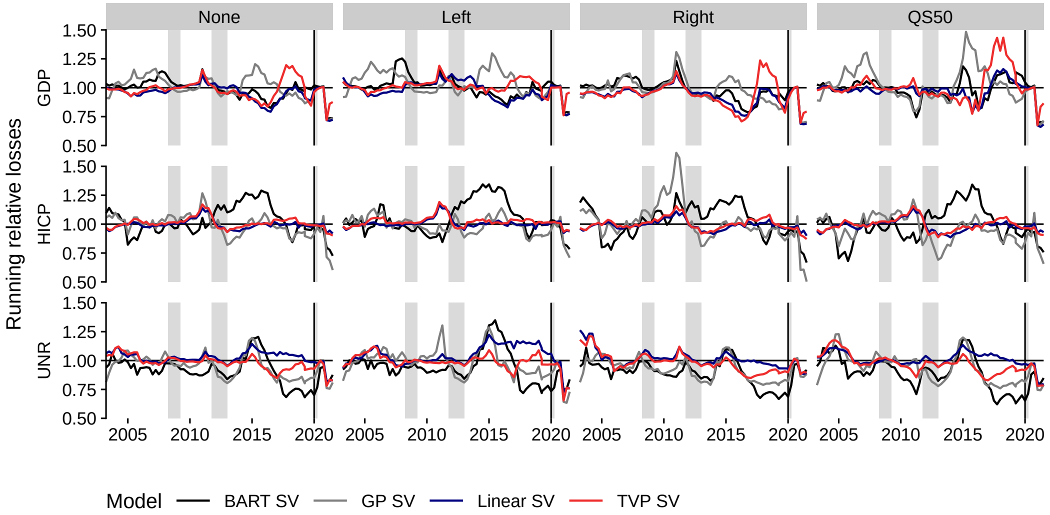
		None					Left					Right					QS50				
		Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion
GDP	TVP SV	0.99	0.98	1.00	0.98	1.00	1.00	1.01	1.00	0.99	1.02	0.98	0.94	1.01	0.96	0.99	0.99	0.99	0.99	0.97	1.01
GDP	Linear SV	0.99	**0.97**	1.01	0.99	**0.99**	1.00	1.00	1.02	1.00	1.01	0.98	**0.94**	1.01	0.97	0.99	0.99	0.98	1.00	0.98	1.00
GDP	GP SV	0.99	0.98	1.00	0.97	1.00	0.99	**0.99**	1.00	0.96	1.03	0.99	0.97	1.00	0.97	0.99	0.98	0.97	0.99	0.97	0.99
GDP	GP hom	**0.98**	0.98	**0.97**	**0.94**	1.00	**0.98**	1.01	**0.95**	**0.93**	1.03	**0.98**	0.96	**0.99**	**0.96**	**0.98**	**0.97**	**0.97**	**0.98**	**0.95**	**0.98**
GDP	BART SV	1.00	0.99	1.01	1.00	1.00	1.02	1.01	1.03	1.01	1.03	0.99	0.96	1.01	0.99	0.99	0.99	0.98	1.00	0.99	0.99
GDP	BART hom	1.00	0.99	1.00	0.99	1.00	1.00	1.00	1.00	0.99	1.01	0.99	0.98	1.01	0.99	0.99	0.99	0.98	1.00	0.99	0.99
HICP	TVP SV	1.01	1.01	0.99	1.00	1.01	1.02	1.02	0.99	0.99	1.03	0.99	1.00	0.99	1.01	0.99	1.00	1.01	0.99	1.01	1.00
HICP	Linear SV	1.00	0.99	1.00	1.00	0.99	1.00	1.00	1.01	0.99	1.00	0.99	0.99	1.00	1.01	0.98	0.99	0.99	1.00	1.00	0.99
HICP	GP SV	0.93	0.94	0.82	0.95	0.92	0.92	0.92	0.85	0.93	0.91	0.94	0.98	0.78	0.98	0.94	0.90	0.91	0.86	0.95	0.89
HICP	GP hom	**0.90**	**0.92**	**0.81**	0.93	**0.89**	**0.89**	**0.90**	**0.84**	0.92	**0.89**	**0.91**	**0.94**	**0.78**	0.95	**0.90**	**0.89**	**0.89**	**0.86**	0.91	**0.88**
HICP	BART SV	0.93	0.93	0.91	**0.89**	0.94	0.93	0.93	0.93	**0.89**	0.94	0.93	0.94	0.90	**0.88**	0.95	0.92	0.92	0.92	**0.90**	0.93
HICP	BART hom	0.94	0.94	0.91	0.91	0.94	0.93	0.94	0.91	0.91	0.94	0.94	0.95	0.90	0.92	0.95	0.93	0.93	0.92	0.90	0.94
UNR	TVP SV	1.00	1.01	**0.95**	0.99	1.01	1.00	1.02	**0.91**	1.01	1.00	1.01	1.01	**0.98**	0.97	1.02	1.00	1.01	**0.94**	1.00	1.00
UNR	Linear SV	1.03	1.03	1.03	0.97	1.05	1.04	1.04	1.06	0.99	1.07	1.01	1.01	1.01	0.96	1.04	1.03	1.03	1.03	0.99	1.05
UNR	GP SV	**0.97**	**0.96**	1.03	**0.91**	1.01	1.00	**0.99**	1.07	0.94	1.03	**0.95**	**0.94**	1.02	**0.88**	**0.99**	**0.98**	**0.97**	1.01	**0.94**	**1.00**
UNR	GP hom	1.01	1.00	1.08	0.96	1.04	1.04	1.02	1.13	0.97	1.07	0.99	0.98	1.05	0.96	1.02	1.01	1.00	1.09	0.97	1.03
UNR	BART SV	0.99	0.98	1.03	0.93	1.02	1.01	1.00	1.07	**0.94**	1.04	0.97	0.97	1.01	0.93	1.00	0.99	0.99	1.00	0.94	1.01
UNR	BART hom	1.00	0.99	1.07	0.95	1.03	1.02	1.00	1.11	0.94	1.05	0.99	0.98	1.04	0.95	1.01	1.01	1.00	1.07	0.96	1.03

Forecast results

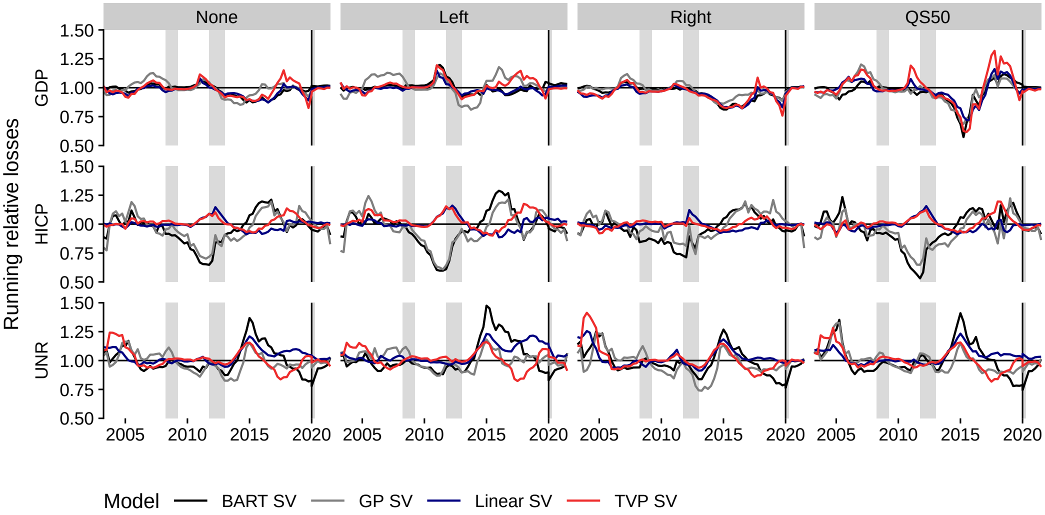
$h = 8$

		None					Left					Right					QS50				
		Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion	Full	Pre Covid	Post Covid	Recession	Expansion
GDP	TVP SV	0.99	0.97	0.99	**0.98**	1.00	**0.99**	0.98	0.99	0.97	1.01	0.99	0.96	1.00	**0.98**	0.99	0.98	0.98	**0.99**	**0.98**	0.99
	Linear SV	**0.98**	**0.96**	1.00	0.98	**0.99**	0.99	**0.96**	1.01	0.98	**0.99**	**0.98**	**0.96**	1.00	0.98	**0.98**	**0.98**	**0.97**	0.99	0.98	**0.98**
	GP SV	0.99	0.98	1.00	0.98	1.01	0.99	0.99	1.00	**0.97**	1.04	0.99	0.99	1.00	0.99	1.00	0.99	0.98	0.99	0.99	0.99
	GP hom	0.99	0.99	**0.99**	0.98	1.00	0.99	1.00	**0.98**	0.97	1.03	0.99	0.98	**1.00**	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	BART SV	0.99	0.97	1.01	0.99	1.00	1.00	0.98	1.01	0.98	1.03	0.99	0.96	1.01	0.99	0.99	0.99	0.98	1.00	0.99	0.99
	BART hom	0.99	0.98	1.00	0.99	0.99	0.99	0.99	1.00	0.99	1.00	0.99	0.97	1.00	0.99	0.99	0.99	0.98	1.00	0.99	0.99
HICP	TVP SV	0.99	0.99	0.99	0.99	0.99	1.00	1.00	0.97	0.98	1.00	0.99	0.98	1.01	0.99	0.99	0.99	0.99	1.00	1.00	0.99
	Linear SV	1.01	1.01	1.01	1.03	1.00	1.01	1.01	1.01	1.04	1.01	1.00	1.00	1.01	1.03	0.99	1.01	1.01	1.01	1.03	1.00
	GP SV	0.93	0.92	1.00	1.00	0.92	0.92	0.91	0.98	0.99	0.90	0.96	0.95	1.01	1.01	0.95	0.91	0.89	1.00	0.96	0.89
	GP hom	**0.92**	**0.91**	**0.97**	**0.97**	**0.91**	**0.91**	**0.91**	**0.91**	**0.96**	**0.89**	**0.94**	**0.92**	1.02	**0.98**	**0.93**	**0.90**	**0.88**	0.98	**0.94**	**0.89**
	BART SV	0.93	0.92	0.98	1.00	0.92	0.93	0.93	0.95	1.01	0.91	0.95	0.93	1.01	1.00	0.94	0.92	0.90	**0.97**	0.95	0.91
	BART hom	0.94	0.93	0.99	0.99	0.93	0.94	0.94	0.95	1.01	0.92	0.94	0.93	1.02	0.98	0.94	0.93	0.92	0.99	0.97	0.92
UNR	TVP SV	**0.99**	0.99	0.99	0.98	**0.99**	**1.00**	1.00	**1.00**	1.04	**0.98**	0.98	0.98	0.98	0.93	1.00	**0.99**	**0.99**	0.99	1.01	**0.98**
	Linear SV	0.99	**0.98**	1.01	0.95	1.01	1.00	**0.99**	1.02	1.00	1.00	0.97	**0.96**	1.01	0.91	1.01	1.00	1.00	1.02	0.98	1.02
	GP SV	1.00	1.00	0.99	**0.87**	1.07	1.03	1.03	1.01	0.92	1.08	**0.97**	0.97	0.97	**0.82**	1.06	1.02	1.02	**0.98**	**0.92**	1.07
	GP hom	1.02	1.03	**0.98**	0.92	1.07	1.05	1.05	1.02	0.96	1.09	1.00	1.01	**0.94**	0.89	1.06	1.04	1.05	1.00	0.97	1.07
	BART SV	1.01	1.01	1.02	0.89	1.08	1.05	1.05	1.03	**0.92**	1.11	0.98	0.98	1.01	0.85	1.06	1.02	1.02	1.01	0.92	1.07
	BART hom	1.03	1.03	1.02	0.91	1.09	1.06	1.06	1.05	0.93	1.12	1.00	1.00	0.99	0.90	1.06	1.03	1.03	1.02	0.93	1.08

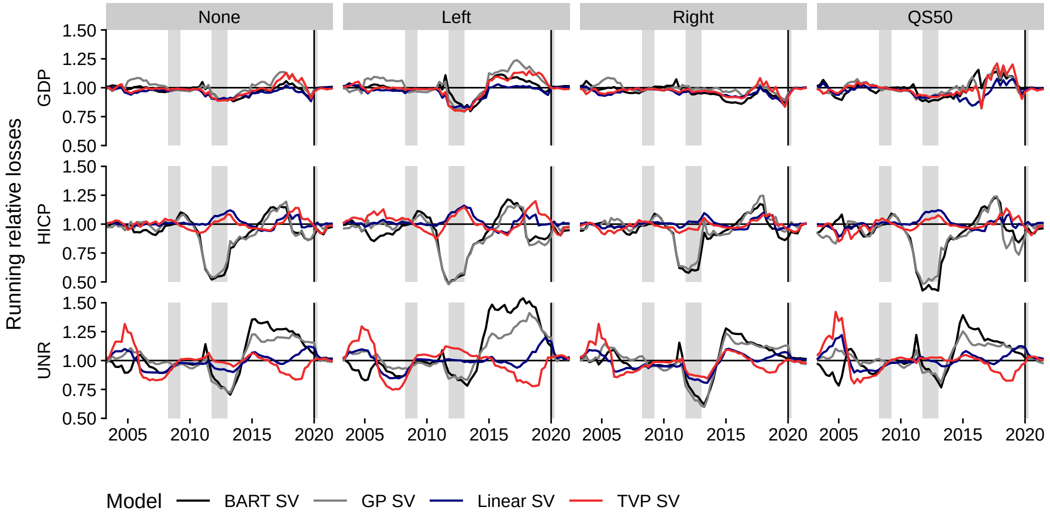
Running window (8 quarters) $h = 1$



Running window (8 quarters) $h = 4$



Running window (8 quarters) $h = 8$



Conclusions

- In this paper we review specification and estimation of Bayesian nonparametric models for forecasting (possibly large sets of) macroeconomic and financial variables, focusing BART and GP, possibly with SV.
- We apply various versions of these models for point, density and tail forecasting using a real time dataset for the euro area (Banbura *et al.*, 2021).
- The performance is compared with that of several variants of Bayesian VARs to assess the relevance of accounting for non-standard, general forms of nonlinearities and parameter time variation.

Conclusions

- For GDP, BVAR-SV overall best, except during recessions and in the left tail, where GP-hom best, with gains shrinking in h.
- For HICP, GP-hom overall best, with larger gains post-Covid and for h=1, but still present for other periods and horizons.
- For UNR, BART overall best, GP-hom during recessions and in the right tail (UNR up).
- For GP, and partly BART, often little or no gains from SV, much less than for BVAR.
- GP overall quite promising for forecasting during problematic periods.