Scenario Analysis with Multivariate Bayesian Machine Learning Models

Michael Pfarrhofer (*WU Vienna*) Anna Stelzer (*Oesterreichische Nationalbank*)

Introduction

Methods and algorithms for **scenario analysis** with **dynamic multivariate models** when the functional form of the conditional mean is **nonlinear and/or unknown**:

- Traditional parametric nonlinear models (thresholds, MS, TVPs);
- Bayesian Machine Learning [Clark et al., 2023, Huber et al., 2023]:

1. can **capture genuine nonlinearities** in data generating processes
2. offer **robustness against misspecification** in likelihood-based inference
3. thus often exhibit **good predictive performance**

Scenario analysis in our context broadly refers to **counterfactual experiments** with empirical multivariate models (in the spirit of SVARs):

- **Conditional forecasting** and **structural scenario analysis** (SSA), with restrictions on observables and/or structural shocks [Antolin-Diaz et al., 2021]
- Nonlinear **IRFs**, GIRFs [Koop et al., 1996], can estimate **dynamic** (*assumptions*: **causal**) **effects** — **interpretable ML** [Rambachan and Shephard, 2021]

Nonlinearity and Predictive distributions

Endogenous variables in $\boldsymbol{y}_t = (y_{1t}, \dots, y_{nt})'$ with lags $\boldsymbol{x}_t = (\boldsymbol{y}_{t-1}, \dots, \boldsymbol{y}_{t-p})'$

$$\boldsymbol{y}_t = \boldsymbol{F}(\boldsymbol{y}_{t-1}, \dots, \boldsymbol{y}_{t-p}) + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\boldsymbol{0}_n, \boldsymbol{\Sigma}_t) \tag{1}$$

Recursive substitution for $h = 1, 2, \dots, H$

$$\boldsymbol{y}_{t+h} = \tilde{\boldsymbol{F}}_{(h)}(\boldsymbol{x}_{t+1}, \boldsymbol{\epsilon}_{t+1}, \boldsymbol{\epsilon}_{t+2}, \dots, \boldsymbol{\epsilon}_{t+h}) = \boldsymbol{F}(\tilde{\boldsymbol{F}}_{(h-1)}(\boldsymbol{x}_{t+1}, \boldsymbol{\epsilon}_{t+1}, \dots, \boldsymbol{\epsilon}_{t+(h-1)})) + \boldsymbol{\epsilon}_{t+h}$$

No closed-form solutions (we provide alternative framework), because:

- $\tilde{\boldsymbol{F}}_{(h)}$ is not additively separable in all its arguments
- Functions of stochastic errors $\boldsymbol{\epsilon}_t$ yield unknown distributions/moments

Predictive Simulation

Let Ξ contain all coefficients and latent states needed to implement the desired model and $\mathcal{I}$ is the information set used to infer Ξ. At forecast origin τ:

$$p_\tau(\boldsymbol{y}_{\tau+1} \,|\, \mathcal{I}) = \int p_\tau(\boldsymbol{y}_{\tau+1} \,|\, \mathcal{I}, \Xi) p(\Xi \,|\, \mathcal{I}) d\Xi$$

$p_\tau(\boldsymbol{y}_{\tau+1} | \mathcal{I})$ generally does not take a known form, neither does the distribution of higher-order forecasts for $h \geq 2$ — simulate forward [Geweke and Amisano, 2010]

1-step ahead conditional under (1): $p_\tau(\boldsymbol{y}_{\tau+1} | \mathcal{I}, \Xi^{(m)}) = \mathcal{N}(\boldsymbol{F}^{(m)}(\boldsymbol{x}_{\tau+1}), \boldsymbol{\Sigma}^{(m)}_{\tau+1})$ allows conditioning recursively on draws indexed (m) for preceding horizons

$$p_\tau(\boldsymbol{y}_{\tau+h} \,|\, \mathcal{I}, \boldsymbol{y}^{(m)}_{\tau+1:\tau+h-1}, \Xi^{(m)}) = \mathcal{N}(\boldsymbol{F}(\boldsymbol{y}^{(m)}_{\tau+h-1}, \dots, \boldsymbol{y}^{(m)}_{\tau+1}, \boldsymbol{y}_\tau, \dots), \boldsymbol{\Sigma}^{(m)}_{\tau+h})$$

exploiting that joint predictive pdf factors into product of the conditional 1-step ahead densities, drawing in each MCMC sweep explores: (drop index τ)

$$p(\boldsymbol{y}_{\tau+1:\tau+h} \,|\, \mathcal{I}) = \int p(\boldsymbol{y}_{\tau+1} \,|\, \mathcal{I}, \Xi) \prod_{j=2}^{h} p(\boldsymbol{y}_{\tau+j} \,|\, \boldsymbol{y}_{\tau+1:\tau+j-1}, \mathcal{I}, \Xi) p(\Xi \,|\, \mathcal{I}) d\Xi \tag{2}$$

Conditional Forecast

Set of restrictions $\mathcal{C}_{1:H} = \{\mathcal{C}_1, \dots, \mathcal{C}_H\}$ over full forecast path — through additional conditioning argument in (2), but key issues:

- **Cannot work directly with joint distribution** to impose them across all horizons simultaneously (no closed-form solutions)
- **Decomposition not generally possible** with one-step ahead densities

Define $\mathcal{C}_h$ as stochastic restrictions [Andersson et al., 2010] of the form:

$$\boldsymbol{R}_h \boldsymbol{y}_{\tau+h} \sim \mathcal{N}(\boldsymbol{r}_h, \boldsymbol{\Omega}_h) \quad \rightarrow \quad \boldsymbol{r}_h = \boldsymbol{R}_h \boldsymbol{y}_{\tau+h} + \boldsymbol{\eta}_h, \quad \boldsymbol{\eta}_h \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Omega}_h)$$

Joint distribution of forecast and restrictions at horizon h:

$$\begin{bmatrix} \boldsymbol{y}_{\tau+h} \\ \boldsymbol{r}_h \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{I}_n \\ \boldsymbol{R}_h \end{bmatrix} \boldsymbol{\mu}^{(m)}_{\tau+h}, \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Sigma}\boldsymbol{R}_h' \\ \boldsymbol{R}_h\boldsymbol{\Sigma} & \boldsymbol{R}_h\boldsymbol{\Sigma}\boldsymbol{R}_h' + \boldsymbol{\Omega}_h \end{bmatrix} \right) \tag{3}$$

→ **nonlinear state space model** — our approach: akin to linear filtering/smoothing [Bańbura et al., 2015] (drop index m)

(Semi-)structural form (heteroskedasticity possible) $\boldsymbol{H}\boldsymbol{y}_{\tau+h} = \boldsymbol{H}\boldsymbol{\mu}_{\tau+h} + \boldsymbol{u}_{\tau+h}$ where $\boldsymbol{u}_{\tau+h} \sim \mathcal{N}(\boldsymbol{0}_n, \boldsymbol{I}_n)$ with $\boldsymbol{\mu}_{\tau+h} = \boldsymbol{F}(\boldsymbol{x}_{\tau+h})$ to restrict (a) observables or (b) shocks:

$$\boldsymbol{R}^{(y)}_h \boldsymbol{y}_{\tau+h} \sim \mathcal{N}(\boldsymbol{r}^{(y)}_h, \boldsymbol{\Omega}^{(y)}_h) \tag{4a}$$

$$\boldsymbol{R}^{(u)}_h \boldsymbol{H}\boldsymbol{y}_{\tau+h} \sim \mathcal{N}(\boldsymbol{r}^{(u)}_h + \boldsymbol{R}^{(u)}_h \boldsymbol{H}\boldsymbol{\mu}_{\tau+h}, \boldsymbol{\Omega}^{(u)}_h) \quad \rightarrow \quad \boldsymbol{R}^{(u)}_h \boldsymbol{u}_{\tau+h} \sim \mathcal{N}(\boldsymbol{r}^{(u)}_h, \boldsymbol{\Omega}^{(u)}_h) \tag{4b}$$

stacked $\boldsymbol{r}_h = [\boldsymbol{r}^{(y)}_h; \boldsymbol{r}^{(u)}_h + \boldsymbol{R}^{(u)}_h \boldsymbol{H}\boldsymbol{\mu}_{\tau+h}]$, $\boldsymbol{R}_h = [\boldsymbol{R}^{(y)}_h; \boldsymbol{R}^{(u)}_h \boldsymbol{H}]$
block diagonal $\boldsymbol{\Omega}_h = \text{bdiag}(\boldsymbol{\Omega}^{(y)}_h, \boldsymbol{\Omega}^{(u)}_h)$

Generalized Impulse Response Functions (GIRFs)

Scenario (s) and **baseline** (b) **forecasts** give rise to counterfactual dynamic effects

$$\boldsymbol{\Delta}_\tau = \mathbb{E}(\boldsymbol{y}_{\tau+1:\tau+H} \,|\, \mathcal{C}^{(s)}_{1:H}, \boldsymbol{x}_{\tau+1}, \mathcal{I}) - \mathbb{E}(\boldsymbol{y}_{\tau+1:\tau+H} \,|\, \mathcal{C}^{(b)}_{1:H}, \boldsymbol{x}_{\tau+1}, \mathcal{I})$$

$$\boldsymbol{\Delta}_\tau = (\boldsymbol{\delta}'_{\tau,1}, \dots, \boldsymbol{\delta}'_{\tau,H})', \quad h\text{-specific: } \boldsymbol{\delta}_{\tau,h}$$

combinations of restrictions yield main variants **conditional on initial** $\boldsymbol{x}_{\tau+1}$:

- **Unorthogonalized** GIRF (based on restrictions on observables), can be extended to SSA via shock restrictions (driving/non-driving shocks)
- **Structural** GIRF (response to identified shocks $\boldsymbol{u}_t$)
$$\boldsymbol{\Delta}^{(d)}_{j\tau} = \mathbb{E}(\boldsymbol{y}_{\tau+1:\tau+H} \,|\, u_{j\tau+1} = d_0 + d, \boldsymbol{x}_{\tau+1}, \mathcal{I}) - \mathbb{E}(\boldsymbol{y}_{\tau+1:\tau+H} \,|\, u_{j\tau+1} = d_0, \boldsymbol{x}_{\tau+1}, \mathcal{I})$$
- **Restricted** GIRF (SGIRF + restricted propagation channels), s.t. $\mathbb{E}(\boldsymbol{R}^{(y)}_h \boldsymbol{\delta}_{\tau,h}) = \boldsymbol{0}$

These are **asymmetric**, **disproportionate**, **state dependent** ($\neq$ IRF). More robust to *Lucas critique* through nonlinearities, granular settings about shocks possible [McKay and Wolf, 2023, Breitenlechner et al., 2024]

Unconditional GIRF as $\overline{\boldsymbol{\Delta}} = \sum_{\tau=1}^{T} \boldsymbol{\Delta}_\tau / T$ [Gonçalves et al., 2021]

MCMC and Particle Gibbs with Ancestor Sampling (PGAS)

- **Mean**: Estimating $\boldsymbol{F}(\bullet)$ with **Bayesian Additive Regression Trees**, **BART** with default settings [Chipman et al., 2010], other options available
- **Variance**: Outliers via $\boldsymbol{\Sigma}_t = o_t\boldsymbol{\Sigma}$ captures **heteroskedastic** data features
- **Algorithm**: new **order-invariant equation-by-equation estimation** based on reduced-form dynamic multivariate BART model

Mostly conditional Gibbs updates possible — **computationally fast MCMC**

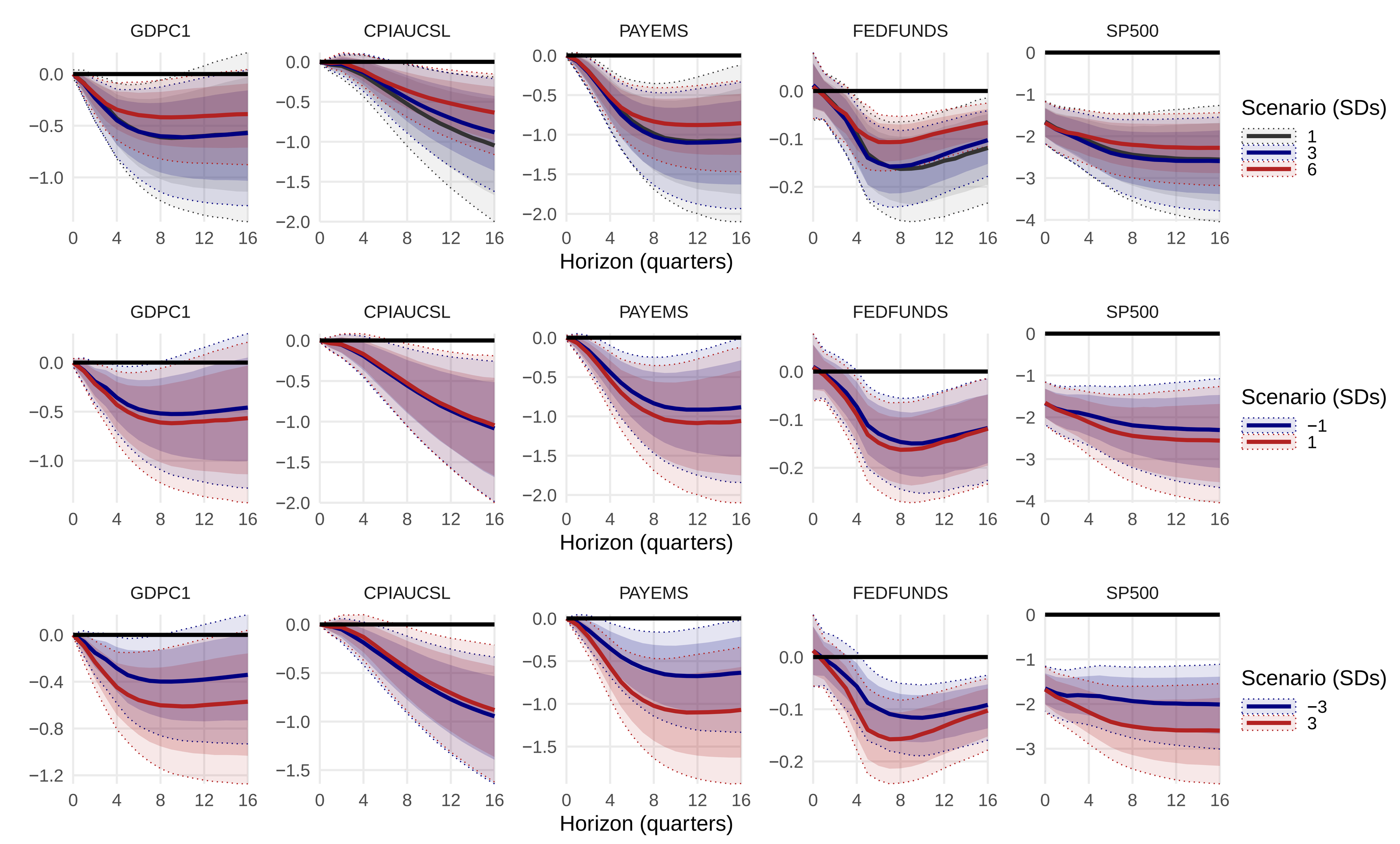
[Andrieu et al., 2010, Lindsten et al., 2014] — score V hypothetical realizations = **particles**, plus reference path (**ancestor**) improves mixing, PGAS based on (3):

1. **Resampling and ancestor sampling**: retain particles with a high probability of having generated restrictions with weights from $h-1$
2. **Propagation**: unconditional prediction or conditional on restrictions
$$\boldsymbol{y}_{\tau+h} \,|\, (\boldsymbol{r}_h, \boldsymbol{R}_h, \boldsymbol{\Omega}_h, \bullet) \sim \mathcal{N}(\boldsymbol{r}^*_h, \boldsymbol{\Omega}^*_h)$$
$$\boldsymbol{r}^*_h = \boldsymbol{\mu}_{\tau+h} + \boldsymbol{\Sigma}\boldsymbol{R}_h'(\boldsymbol{R}_h\boldsymbol{\Sigma}\boldsymbol{R}_h' + \boldsymbol{\Omega}_h)^{-1}(\boldsymbol{r}_h - \boldsymbol{R}_h\boldsymbol{\mu}_{\tau+h})$$
$$\boldsymbol{\Omega}^*_h = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\boldsymbol{R}_h'(\boldsymbol{R}_h\boldsymbol{\Sigma}\boldsymbol{R}_h' + \boldsymbol{\Omega}_h)^{-1}\boldsymbol{R}_h\boldsymbol{\Sigma}$$
to generate next h (*known* conditional distribution, less particles required)
3. **Weighting**: compute likelihood of the restriction using particle lineage

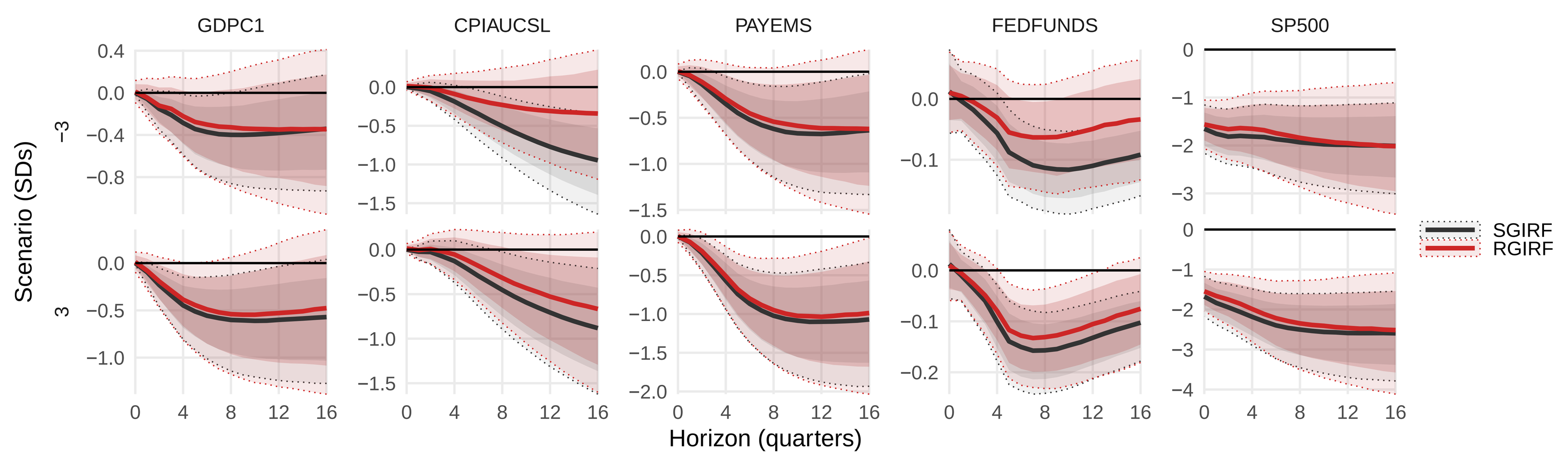
Fully **smoothed trajectory** by tracing back through the stored ancestor indices — **forward smoothing, no backward recursion** (computationally efficient)

Empirical Application(s)

(1) recursively identified **financial shocks** and **spillovers** [*only in paper*]: (2) Fed **macro-stress test** conditional on observables forecast and (3) **NFCI macro-at-risk**



US financial shock (zero-restrictions) and spillovers to EA and UK, GIRFs scaled to 1 SD — size of shock inconsequential, sign of shock matters for larger innovations



Ruling out spillovers/spillbacks via $\mathbb{E}$(real GDP, FX rate) $= 0$ in EA and UK matters for inflation and Fed response in the US