

Beware of large shocks!

A non-parametric structural inflation model

Elena Bobeica¹ Sarah Holton¹ Florian Huber² Catalina Martínez Hernández¹

¹European Central Bank ²University of Salzburg

Inflation: Drivers and Dynamics Conference 2026
Cleveland, 24-25 September 2026

Link to paper [here](#)

Link to VoxEU article [here](#)

The views expressed herein are solely those of the authors and do not necessarily reflect the views of the European Central Bank or the Eurosystem.

This paper in a nutshell

- ▶ Propose a non-parametric structural inflation model
 - ▶ Flexible on the shape possible non-linearities
 - ▶ Allows for identification of shocks
- ▶ We focus on energy shocks
 - ▶ We find that larger shocks have a disproportionately larger impact on inflation
 - ▶ Non-linearity more visible at the beginning of the pricing chain
 - ▶ Other shocks also elicit non-linear reactions, but less prominent
- ▶ Non-linear model better fits the data compared to a linear version in an out-of-sample forecasting exercise

The model: A blend of VARs and machine learning

- ▶ We assume that the M –dimensional data vector $\mathbf{y}_t$ is a function of $\mathbf{x}_t = (\mathbf{y}'_{t-1}, \dots, \mathbf{y}'_{t-p})'$ and a set of Q fundamental shocks $\mathbf{q}_t$:

$$\mathbf{y}_t = \underbrace{\mathbf{A}\mathbf{x}_t + \tilde{f}(\mathbf{x}_t)}_{\text{Conditional mean}} + \underbrace{\mathbf{\Lambda}\mathbf{q}_t + \tilde{g}(\mathbf{q}_t) + \eta_t}_{\text{Reduced-from errors with factor structure}} \quad (1)$$

where

- ▶ $\mathbf{A}$ is a $M \times K$ dimensional coefficient matrix and $\mathbf{\Lambda}$ is a matrix of factor loadings
- ▶ $\tilde{f} : \mathbb{R}^K \rightarrow \mathbb{R}^M$ and $\tilde{g} : \mathbb{R}^Q \rightarrow \mathbb{R}^M$ are unknown, **non-linear** functions approximated by Bayesian Additive Regression Tree (BART) as in Clark et al. (2023)
- ▶ η_t is an idiosyncratic component with diagonal covariance matrix Ω

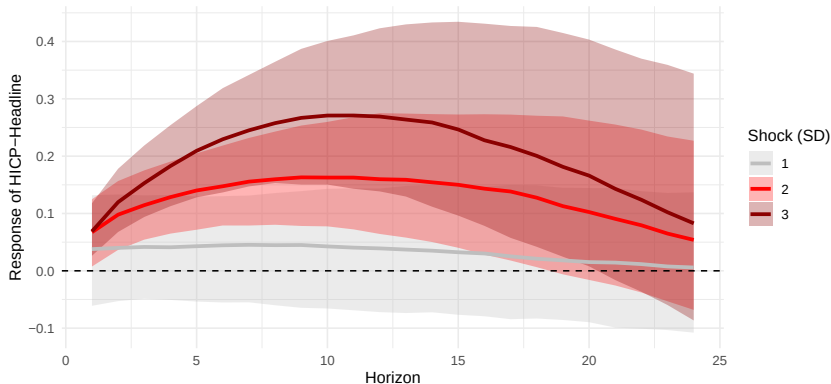
Identification scheme

Table: Structural shock identification

Variable/Shock	Energy	Global supply chains	Domestic supply	Demand
HICP headline	+	+	+	+
Synthetic	σ_y	0	0	
Price expectations manuf.	+	+	σ_y	+
Consumer confidence	-	-	-	σ_y
PPI	+	+	+	+
PPI energy	+			
PMI supplier delivery		-		
GSCPI		σ_y	0	
EUR/USD				
EURIBOR 1Y				

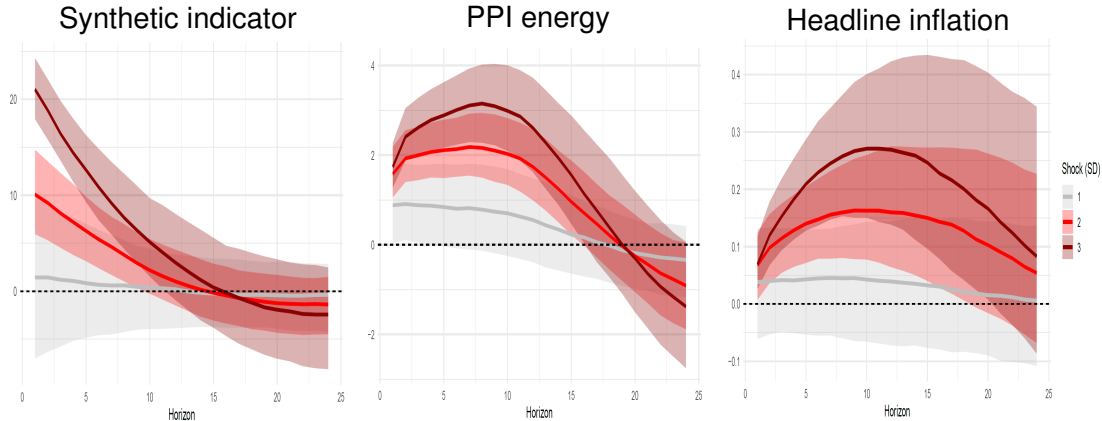
Notes: Restrictions are contemporaneous, σ_y reflects a normalisation restriction.

Reaction of headline inflation to energy price shocks of different sizes



Notes: Thick lines are median estimates and the shaded areas are the 68% credible interval. IRFs are normalized for comparability.

Transmission of energy shocks along the pricing chain



Thick lines are median estimates and the shaded areas are the 68% credible interval. IRFs are normalized for comparability.

Reaction of headline inflation to energy shocks over a sizes grid

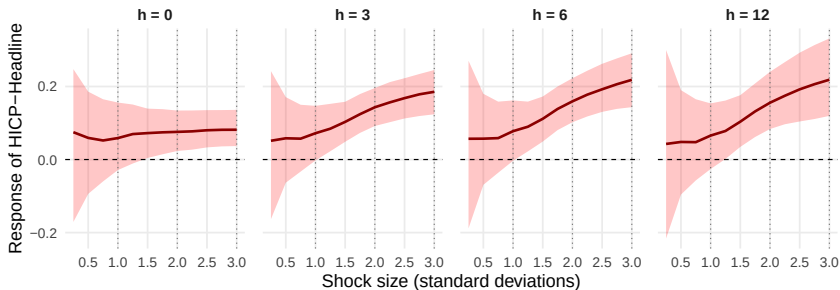
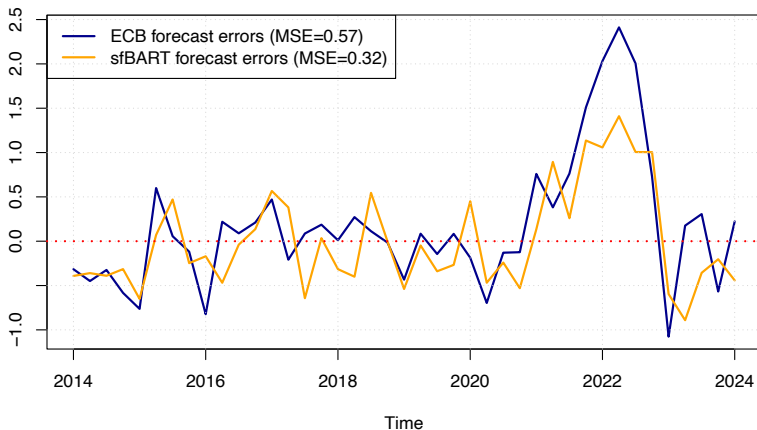


Figure: Reaction of inflation to energy shocks over a sizes grid

Notes: Dark red lines are median estimates and the shaded areas denote the 16/84 percentiles of the posterior distribution. Dotted verticals mark the three shock sizes reported elsewhere. Responses are normalised to a one standard deviation shock for comparability.

Validation of the model from a forecasting perspective

Forecast errors ECB forecasts vs. BART



Note: The chart compares the mean-squared forecast errors of the ECB's official projections with those of the BART model for the one-quarter-ahead forecasting horizon. To do so, we used the three-month-ahead forecasts from our monthly BART model to align with the one-quarter-ahead horizon.

Conclusions

- ▶ New model with non-linear features and structural identification.
- ▶ Good forecasting properties compared to a linear version, especially when large shocks hit.
- ▶ Large shocks generate disproportional inflation responses.
- ▶ Policy implications: large shocks transmit differently and therefore may require differentiated response.

BACKGROUND

Inflation and synthetic energy commodity price indicator

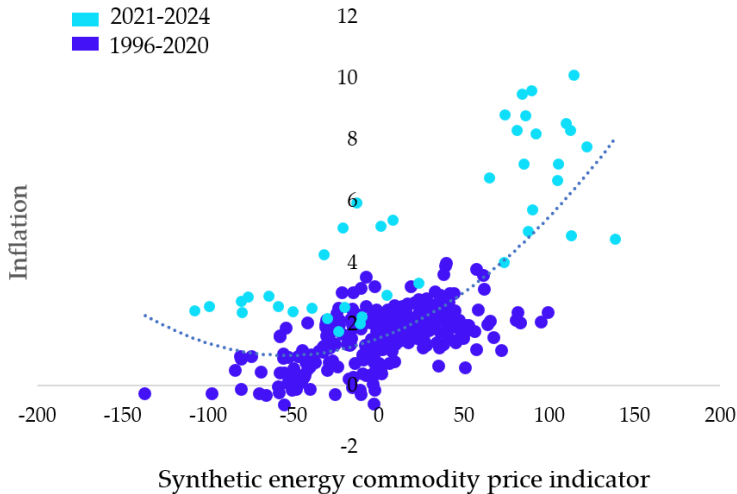


Figure: Inflation and energy commodity prices

An Introduction to BART

- ▶ BART approximates each $f_j(\mathbf{x}_t)$ as follows:

$$f_j(\mathbf{x}_t) = \sum_{s=1}^S u_s(\mathbf{x}_t | \mathcal{T}_{js}, \mu_{js}),$$

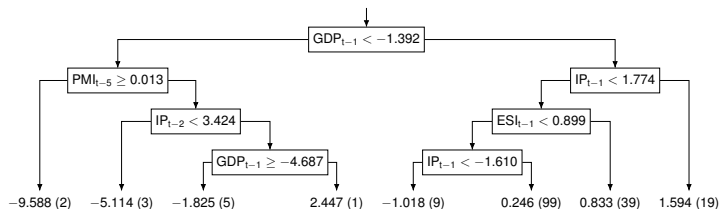
- ▶ $\mathcal{T}_{js}$ are tree structures
- ▶ μ_{js} are tree-specific terminal nodes
- ▶ S denotes the total number of trees used.
- ▶ Dimension of μ_{js} is denoted by b_s which depends on the complexity of the tree
- ▶ Note that BART involves adding up different trees (this is the "A" in BART)
- ▶ But what is a regression tree?

Intuition of a Regression Tree

- ▶ Explain idea of BART for a single tree (suppress s subscripts)
- ▶ Conventional regression: For every value for $\mathbf{X}$ produces a fitted value for $\mathbf{y}$
- ▶ BART does same thing, but in different way
- ▶ Splitting rule: Divides space of $\mathbf{X}$ into different intervals each of which has same fitted value for $\mathbf{y}$ (internal nodes, branches of tree)
- ▶ Fitted values are called terminal nodes (or leaves of the tree)

An estimated tree

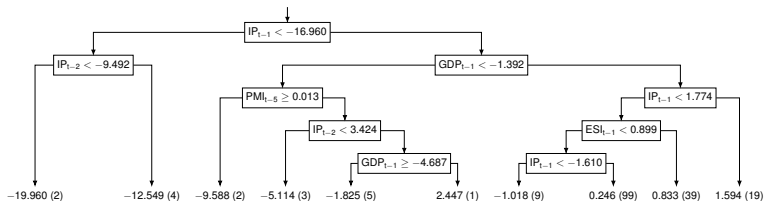
- Suppose we wish to explain German GDP growth using lagged GDP, industrial production (IP), the Economic Sentiment Indicator (ESI) and the Purchasing Manager Index (PMI)
- An estimated tree would look like this:



Estimated tree for Germany: 2005Q1 to 2019Q4. Source: Huber et al. (2023)

An estimated tree

- ▶ The previous example did not include pandemic observations
- ▶ The resulting tree looks like this:



Estimated tree for Germany: 2005Q1 to 2020Q2. Source: Huber et al. (2023)

Priors on the tree-based parameters

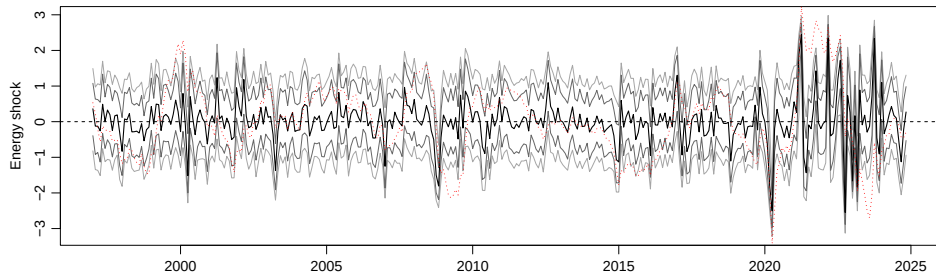
- ▶ The B in BART stands for Bayesian
- ▶ Idea: use priors that make complicated trees less likely
- ▶ Each tree should be simple and explain only a little bit of variation in the response ('weak learner')
- ▶ All this is achieved through regularization priors:
 - ▶ Probability that a node within a tree is non-terminal declines with tree complexity
 - ▶ Terminal node parameters arise from a Gaussian distribution with variance decreasing in the number of trees

Posterior simulation

- ▶ Apart from sampling the factors and the loadings, all other steps are standard and based on the algorithm proposed in Clark et al. (2023)
- ▶ The loadings are simulated from truncated Gaussian posterior distributions as in Korobilis (2022)
- ▶ The factors are simulated using a random walk Metropolis Hastings update
 - ▶ For $t = 1, \dots, T$ we propose $\mathbf{q}_t^* \sim \mathcal{N}(\mathbf{q}_t^a, c \cdot \hat{\Sigma}_t)$

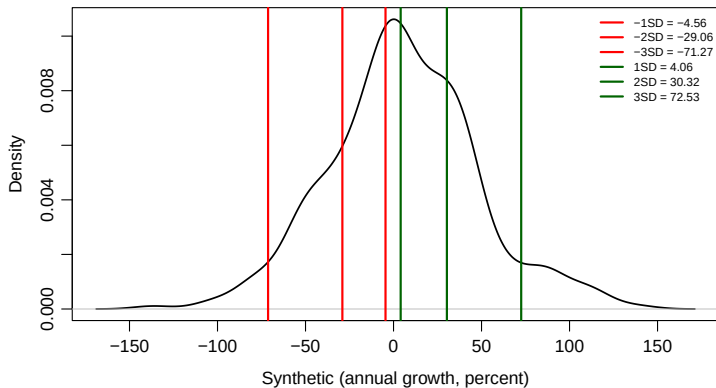
where $\mathbf{q}_t^a$ denotes the previously accepted draw and $\hat{\Sigma}_t$ is the time t posterior covariance matrix from a linear factor model and c is a scaling parameter so that the acceptance rate is between 20 and 40 percent

Estimated energy shocks



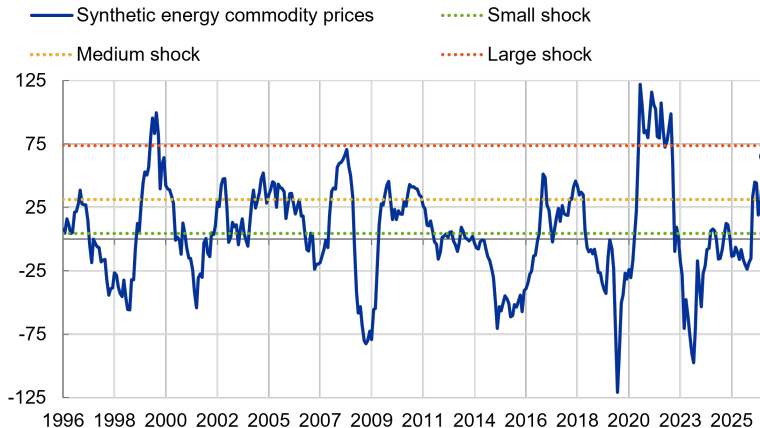
Small, medium and large energy shocks

Empirical distribution of Synthetic vs. impact responses



Notes: Synthetic energy indicator and contemporaneous reaction to shocks of different size. Synthetic indicator expressed in annual growth terms.

The current energy shock through the lens of the model



Note: synthetic indicator of energy prices, combining oil and gas prices using euro area's import shares (in annual log diffs). Red, yellow, and green dotted lines mark thresholds for small, medium and large energy shocks. The yellow circle is an estimate for September 2026 based on data until September 18, 2026.

Reaction of inflation to the identified structural shocks

