Monetary Policy across Inflation Regimes

Katerina Petrova[1]

Federal Reserve Bank of New York
Universitat Pompeu Fabra, BSE

30 September 2025
Inflation: Drivers & Dynamics Conference 2025, ECB

joint with Christian Matthes (Indiana University) and Valeria Gargiulo (UPF)

[1] The views expressed here are my own and do not necessarily reflect those of the Federal Reserve Bank of New York or any other part of the Federal Reserve System.

Introduction

- Does the effect of monetary policy depend on the level of inflation?
- Most of the research: linear models with constant parameters
- Economic agents: behave differently when inflation very high (Weber et al. 2023)
- In the presence of state-dependence: linear models erroneous empirical conclusions, averaging over regimes

This Paper

- We build a self-exciting threshold (SET-) Bayesian VAR model
- Methodological contributions:
  - allow for regime-dependence in variance in 1st stage estimation; novel since we utilise volatility regimes to aid threshold identification
  - combine two-step frequentist estimation with Bayesian regularization; ignore posterior uncertainty of threshold - allowed due to faster posterior contraction rate

This Paper

- Main advantages of Approach:
  - standard Bayesian shrinkage on large dimensional regime-dependent VAR parameters (relative to frequentist procedures)
  - easier and faster inference (relative to fully Bayesian threshold treatment)
  - simple, easily interpretable nonlinearity mechanism (relative to Markov-switching models)
- Our modeling choice: three main goals: (i) parsimony, (ii) computational speed and (iii) transparency

Plan for the Talk

- Methodology, estimation algorithm
- Monte Carlo Exercise
- Empirical Application to US inflation regimes

Methodology

- Univariate TAR model: introduced by Tong (1977), generalized by Tong and Lim (1980), Chan (1993), and Tong (2011)
- Consider $M \times 1$ T-VAR(p) model with k regimes:

$$y_t = \sum_{i=1}^{k} (B_{0,i} + \sum_{j=1}^{p} B_{j,i} y_{t-j}) \Psi_t^{(i)}(\gamma^0) + \Sigma^{1/2} \eta_t$$

- γ^0 is threshold parameter: $\gamma_1^0 < \gamma_2^0 < \cdots < \gamma_{k-1}^0$
- i^{th} regime defined as $\Psi_t^{(i)}(\gamma^0) = \mathbb{I}(\gamma_{i-1}^0 < s_t \leq \gamma_i^0)$
- s_t is $\mathcal{F}_{t-1}-$measurable state variable
- if s_t variable from y_t with lag $d \in \{1, ..., p\}$, the model: self-exciting T-VAR (SET-VAR)

Methodology

- SETAR models
  - nonlinearity through indicators: model piecewise linear; facilitates simple estimation
  - while simple, self-exciting mechanism can capture important nonlinearities particularly in cyclical data
- Consistency and asymptotic distributions of LS in SETAR models: Chan (1993)

Standard Estimation

- Conditional on γ^0, estimation of β standard
- For each i, $\sqrt{n}-$consistent asymptotically normal $\hat{\beta}_i$
- γ unknown $\Rightarrow$ a consistent estimator required
- Numerical minimization SSR as function of γ
- Let $B_i = (B_{0,i}, B_{1,i} \ldots, B_{p,i})$, $\beta_i = vec(B_i')$ and $\beta = (\beta_1', \ldots, \beta_k')'$
- $\hat{\beta} = (\hat{\beta}_1', \ldots, \hat{\beta}_k')'$ used to compute residuals for all possible values of grid for γ
- Estimator for γ:

$$\hat{\gamma} = \arg\min_\gamma \sum_{t=1}^n \hat{\varepsilon}_t' \hat{\varepsilon}_t = \arg\min_\gamma \left[\min_{\beta_1,\ldots,\beta_k} \sum_{t=1}^n \varepsilon_t' \varepsilon_t\right]$$

where $\hat{\varepsilon}_t = \left(y_t - \sum_{i=1}^k (I_M \otimes x_t') \hat{\beta}_i \Psi_t^{(i)}(\gamma)\right)$ and $x_t = \left(1, y_{t-1}', \ldots y_{t-p}'\right)'$

- $\hat{\gamma}$ equivalently coming from $(\hat{\gamma}, \hat{\beta}) = \arg\min_{\gamma,\beta} \sum_{t=1}^n \varepsilon_t' \varepsilon_t$.

Standard Estimation

- The standard identification for all $i \in \{1, \dots, k\}$

$$\forall i, j \in \{1, \dots, k\}, \quad \beta_i \neq \beta_j \quad \text{when} \quad i \neq j;$$

  i.e., at least one element in β differs across any pair of regimes
- Super-consistency of $\hat{\gamma}$ to γ^0 with faster rate (n instead of the usual parametric $\sqrt{n}$)
- Inference typically in two steps:
  - γ is estimated
  - conditional on $\hat{\gamma}$, inference on β_i is standard, consistent and asymptotically normal
- super-consistency for $\hat{\gamma}$: estimation uncertainty of γ no first-order effect on inference for β_i; justification for two-step plug-in procedure

Novel Procedure

$$y_t = \sum_{i=1}^{k} \left(B_{0,i} + \sum_{j=1}^{p} B_{j,i} y_{t-j} + \Sigma_i^{1/2} \eta_t \right) \Psi_t^{(i)} \left(\gamma^0 \right)$$

- Prior density $p\left(\beta_i, \Sigma_i\right)$ for each $i = 1, ..., k$, prior $p\left(\gamma\right)$
- Grid of N_γ points $\Gamma = \left(\underline{\gamma}, ..., \overline{\gamma}\right)$ for each γ, i.e. discrete uniform prior $p\left(\gamma_{ij}\right) = 1/N_\gamma$
- Gaussianity assumption on $\eta_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, I_M)$
- The log-posterior density (except constants):
  $\ln(p\left(\beta, \Sigma, \gamma | y_1, ..., y_n\right))$
  $$= \ell\left(\beta, \Sigma, \gamma\right) + \sum_{i=1}^{k} \omega_i \ln p\left(\beta_i, \Sigma_i\right) + \ln p\left(\gamma\right)$$
  where $\ell\left(\beta, \Sigma, \gamma\right)$ given by $\ell\left(\beta, \Sigma, \gamma\right) = \sum_{i=1}^{k} \ell_i\left(\beta, \Sigma, \gamma\right)$
  $$\ell_i\left(\beta, \Sigma, \gamma\right) = -\frac{n_i}{2} \ln\left(2\pi\right) - \frac{n_i}{2} \ln\det\left(\Sigma_i\right) - \frac{1}{2} \sum_{t=1}^{n} \varepsilon_{it}' \Sigma_i^{-1} \varepsilon_{it}$$
- ω_i: contribution of regimes: $\omega_i = \frac{n_i}{n}$

Novel Procedure

- Equivalently denote $w_{t,i} = \mathbb{I}(\hat{\gamma}_{i-1} < s_t \leq \hat{\gamma}_i)$ and define $Y = (y_1, \ldots, y_n)'$, $y = vec(Y)$, $X = (x_1', \ldots, x_n')'$ and $W_i = diag(w_{1,i}, \ldots, w_{n,i})$
- For each i, $\ell_i(\beta, \Sigma, \gamma) \propto -\frac{tr(W_i)}{2} \ln(\det \Sigma_i)$

$$-\frac{1}{2}(y - (I_M \otimes X)\beta_i)' (\Sigma_i^{-1} \otimes W_i)(y - (I_M \otimes X)\beta_i),$$

$$(\hat{\gamma}, \hat{\beta}, \hat{\Sigma}) = \arg\max_{\gamma,\beta,\Sigma} \ln(p(\beta, \Sigma, \gamma | y_1, \ldots, y_n)),$$

where maximizer $\hat{\gamma}$ equivalently obtained through

$$\begin{aligned}
\hat{\gamma} &= \arg\max_{\gamma} \left[\max_{\beta_1,\ldots,\beta_k,\Sigma_1,\ldots,\Sigma_k} \ln(p(\beta, \Sigma, \gamma | y_1, \ldots, y_n))\right] \\
&= \arg\max_{\gamma} \ln\left(p(\breve{\beta}, \breve{\Sigma}, \gamma | y_1, \ldots, y_n)\right) \\
&= \arg\max_{\gamma \in \Gamma^{k-1}} \ell\left(\breve{\beta}, \breve{\Sigma}, \gamma\right) + \sum_{i=1}^{k} \omega_i \ln p\left(\breve{\beta}_i, \breve{\Sigma}_i\right)
\end{aligned}$$

where $\breve{\beta}$ and $\breve{\Sigma}$ are the posterior modes

Novel Procedure

- In practice: evaluate $\ln p$ at mode for $\theta_i = [\beta_i', vech(\Sigma_i)']'$ over grid of values for γ; $\hat{\gamma}$ - maximizer
- Since $\hat{\gamma}$ converges at faster rate, as $n \to \infty$

$$p(\sqrt{n}(\theta_i - \theta_i^0)|\gamma, y_1, ..., y_n) - p(\sqrt{n}(\theta_i - \theta_i^0)|\hat{\gamma}, y_1, ..., y_n)) \to 0$$

- Avoid expensive Metropolis step
- Conditional $\hat{\gamma}$, standard Bayesian estimation for θ_i
- Bayesian treatment on θ relevant: large M and $p \Rightarrow$ overfit after splitting observations into regimes
- Prior: useful to penalise and regularise estimation
- Standard Bayesian methodology for conditional inference on θ_i: Minnesota prior on B_i, Wishart prior on Σ_i^{-1}

Conditional Posteriors

$$\beta_i|\Sigma_i,\gamma \sim \mathcal{N}\left(\beta_{0i},(\Sigma_i^{-1}\otimes\kappa_{0i})^{-1}\right),\quad \Sigma_i^{-1}|\gamma \sim \mathcal{W}(\alpha_{0i},\lambda_{0i})$$

- Closed-form NW posterior conditional on γ:

$$\beta_i|\Sigma_i,\gamma,X,Y \sim \mathcal{N}\left(\tilde{\beta}_i,(\Sigma_i^{-1}\otimes\tilde{\kappa}_i)^{-1}\right),\quad \Sigma_i^{-1}|\gamma \sim \mathcal{W}(\tilde{\alpha}_i,\tilde{\lambda}_i),$$
$$\widetilde{\beta}_i = \left(I_M\otimes\widetilde{\kappa}_i^{-1}\right)\left[(I_M\otimes X'W_iX)\hat{\beta}_i + (I_M\otimes\kappa_{0i})\beta_{0i}\right],$$
$$\widetilde{\kappa}_i = \kappa_{0i} + X'W_iX,$$
$$\widetilde{\alpha}_i = \alpha_{0i} + n_i,\ \widetilde{\lambda}_i = \lambda_{0i} + Y'W_iY + B_{0i}\kappa_{0i}B_{0i}' - \widetilde{B}_i\widetilde{\kappa}_i\widetilde{B}_i',$$

and $\hat{\beta}_i$ threshold OLS for each i :

$$\hat{\beta}_i = (I_M\otimes X'W_iX)^{-1}(I_M\otimes X'W_i)y,$$

Estimation Algorithm

Step 1. For each grid point in Γ^{k-1}, compute posterior modes for β_i and Σ_i, which in the NW setup are $\breve{\beta}_i = \tilde{\beta}_i$ and $\breve{\Sigma}_i = \frac{\tilde{\lambda}_i}{(\tilde{\alpha}_i + M + 1)}$.
Step 2. Evaluate log-likelihood $\ell\left(\breve{\beta}, \breve{\Sigma}, \gamma\right)$ and weighted prior density $\sum_{i=1}^{k} \omega_i \ln p\left(\breve{\beta}_i, \breve{\Sigma}_i\right)$ at $\breve{\beta}$ and $\breve{\Sigma}$ for each grid point in Γ^{k-1}.
Step 3. Numerically maximize log-posterior $p\left(\breve{\beta}, \breve{\Sigma}, \gamma | y_1, ..., y_n\right)$ wrt γ over the $(k-1)$-dimensional grid and store the estimate $\hat{\gamma}$.
Step 4. Given $\hat{\gamma}$ from Step 3, make draws for β_i and Σ_i from the posterior distribution above.

Monte Carlo Exercise

- Design small Monte Carlo exercise: study properties; compare to existing approaches
- Univariate process $y_t = \mu_t + \sigma_t \varepsilon_t$, $\varepsilon_t \sim \mathcal{N}(0,1)$, $n \in \{200, 1000\}$
- 4 DGPs:
  - DGP I: $\mu_t = 0$ and $\sigma = 1$ for all t;
  - DGP II: $\mu_t = \mu_1 1\{y_{t-1} \leq \gamma\} + \mu_2 1\{y_{t-1} > \gamma\}$ with $\mu_1 = -1$, $\mu_2 = 1$ and $\sigma_t = 1$ for all t;
  - DGP III: $\mu_t = 0$ for all t and $\sigma_t = \sigma_1 1\{y_{t-1} \leq \gamma\} + \sigma_2 1\{y_{t-1} > \gamma\}$ with $\sigma_1 = 1$, $\sigma_2 = 2$;
  - DGP IV: regime-dependent mean as in DGPII and regime-dependent volatility as in DGPIII
- True $\gamma = 0$

Monte Carlo Exercise

- We estimate
  - a constant parameter (CP) model
  - regime-dependent SSR-threshold, constant variance
  - SSR-threshold with regime-dependence in variance in 2nd stage (as Chan (1993), Tsay (1998))
  - regime-dependent likelihood-threshold
- Compare RMSEs for each specification and DGP

Monte Carlo Exercise

Average RMSE					
Constant Parameter Model					
	γ	μ_1	μ_2	σ_1	σ_2
DGP I, n=200	-	0.0572	0.0572	0.0398	0.0398
DGP I, n=1000	-	0.0254	0.0254	0.0177	0.0177
DGP II, n=200	-	0.9975	1.0025	0.4116	0.4116
DGP II, n=1000	-	0.9992	1.0008	0.4138	0.4138
DGP III, n=200	-	0.0899	0.0899	0.5725	0.4275
DGP III, n=1000	-	0.0402	0.0402	0.5796	0.4204
DGP IV, n=200	-	1.0940	0.9060	0.9010	0.1149
DGP IV, n=1000	-	1.0968	0.9032	0.9058	0.0947
SSR Estimation, $\sigma_1 = \sigma_2$ in Stage 2					
DGP I, n=200	-	0.3784	0.3872	0.0817	0.0817
DGP I, n=1000	-	0.2944	0.3046	0.0356	0.0356
DGP II, n=200	0.0219	0.0817	0.0802	0.0799	0.0799
DGP II, n=1000	0.0090	0.0362	0.0360	0.0357	0.0357
DGP III, n=200	0.4284	0.0920	0.1612	1.4702	0.4826
DGP III, n=1000	0.3643	0.0405	0.0725	1.4945	0.4945
DGP IV, n=200	0.9388	0.1151	0.4360	2.1042	1.1043
DGP IV, n=1000	0.9864	0.0530	0.4404	2.1285	1.1285

Monte Carlo Exercise

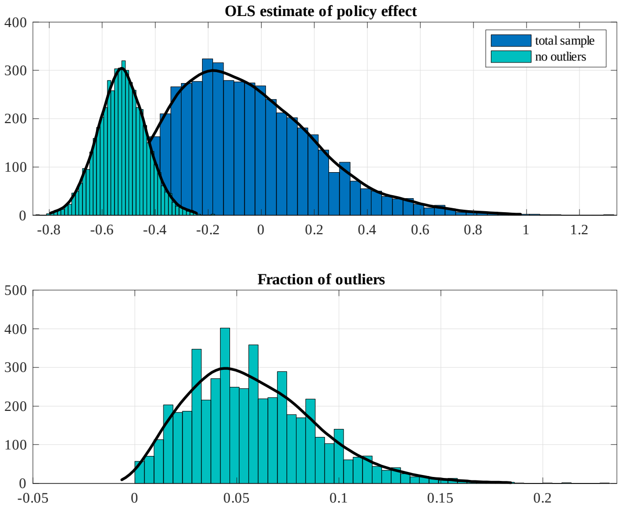
SSR Estimation, $\sigma_1 \neq \sigma_2$ in Stage 2					
DGP I, n=200	-	0.3784	0.3872	0.1515	0.1513
DGP I, n=1000	-	0.2944	0.3046	0.1042	0.1068
DGP II, n=200	0.0219	0.0817	0.0802	0.0560	0.0567
DGP II, n=1000	0.0090	0.0362	0.0360	0.0251	0.0257
DGP III, n=200	0.4284	0.0920	0.1612	0.1478	0.1624
DGP III, n=1000	0.3643	0.0405	0.0725	0.1172	0.0962
DGP IV, n=200	0.9388	0.1151	0.4360	0.0853	0.0936
DGP IV, n=1000	0.9864	0.0530	0.4404	0.0407	0.0428
Likelihood-based Estimation					
DGP I, n=200	-	0.2878	0.2908	0.2670	0.2759
DGP I, n=1000	-	0.2340	0.2394	0.2245	0.2256
DGP II, n=200	0.0221	0.0817	0.0802	0.0567	0.0573
DGP II, n=1000	0.0091	0.0362	0.0360	0.0253	0.0258
DGP III, n=200	0.0740	0.0840	0.1656	0.0599	0.1167
DGP III, n=1000	0.0176	0.0366	0.0717	0.0253	0.0506
DGP IV, n=200	0.0385	0.0861	0.1519	0.0607	0.1084
DGP IV, n=1000	0.0146	0.0380	0.0680	0.0266	0.0479

- No regimes: all approaches valid; CP model best
- When switch mean (DGP II) CP model inconsistent; SSR model with or without regime in σ consistent; likelihood performs equally well
- DGP III: both SSR-models inconsistent estimator for γ; no distortion on mean inference, since true mean constant, any sample split
- Estimation of σ distorted if regime-dependence ignored, even when allowed in 2nd stage, estimates poor, threshold not precisely estimated
- Likelihood model performs well; correctly identifies the value of the threshold
- In DGP IV, CP model inadequate
- SSR-model inconsistent estimate of threshold
- Likelihood approach: consistent and precise estimate of γ; consequently precise and consistent estimates for μ and σ in 2nd stage.

Illustration with outliers

- One endogenous variable z_t, one shock of interest m_t; u_t summarises persistent effects of other variables
- $z_t = \beta m_t 1\{z_{t-1} < \overline{z}\} + \gamma m_t 1\{z_{t-1} \geq \overline{z}\} + \rho z_{t-1} + u_t$ and m_t are zero-mean i.i.d. Gaussian random variables
- $\overline{z}$ is threshold value, calibrated so that z_t infrequently exceeds the threshold; i.e. model is more often in 1st regime.
- Assume opposing effects of the shock in the two regimes: $\beta < 0$ and $\gamma > 0$.
- Simulate 5,000 samples of $n = 500$; estimate one regression for entire sample and another only for observations where $z_{t-1} < \overline{z}$ (equivalent to estimating threshold model with knowledge on true value)

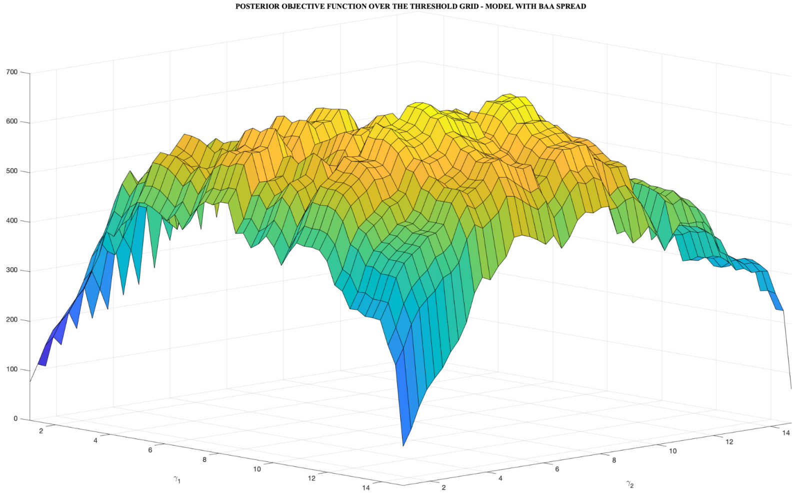
Monte Carlo Exercise



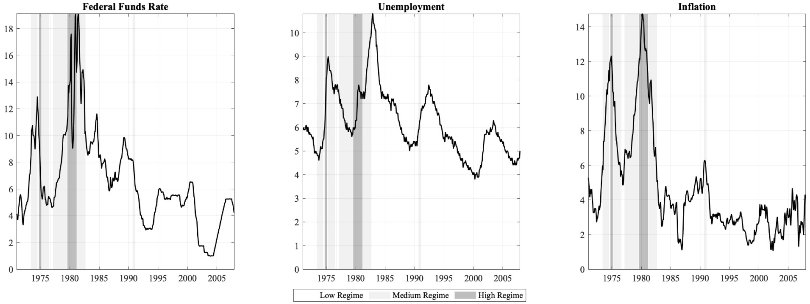
Empirical Application

- We estimate our SET-VAR model on US data, study regime dependence of policy
- Monthly U.S. data from January 1970 to December 2007 on: FFR, unemployment rate, CPI YoY inflation; Romer & Romer's proxy for policy shock (robustness BBA spread)
- Choice of state variable $s_t = \pi_{t-1}$ (results robust to different lags)
- Info on priors: in paper
- Low regime: $\pi_{t-1} \leq 5.49$ (74.3% of the sample);
- Medium regime: $5.49 < \pi_{t-1} \leq 11.02$ (19.6% of the sample);
- High regime: $\pi_{t-1} > 11.02$ (6.1% of the sample).

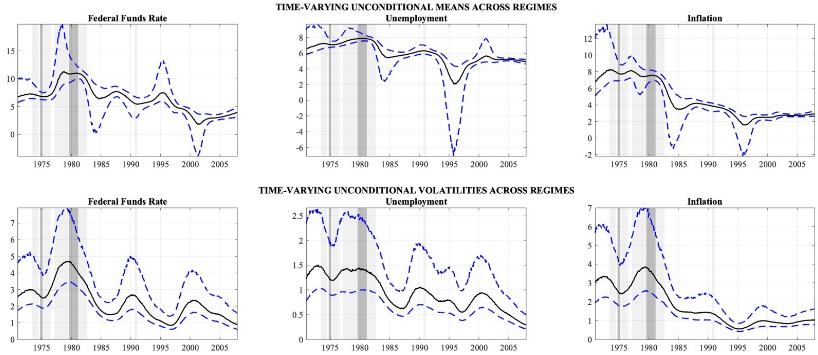
Objective function



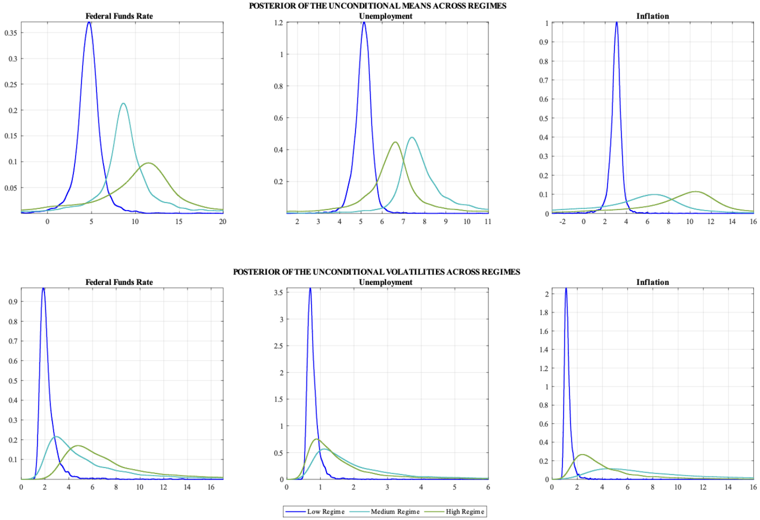
Raw Data against estimated Regimes



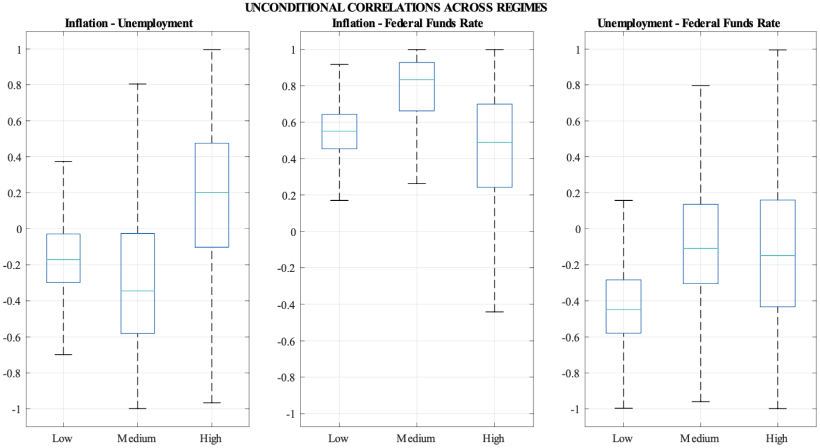
TVP model against estimated Regimes



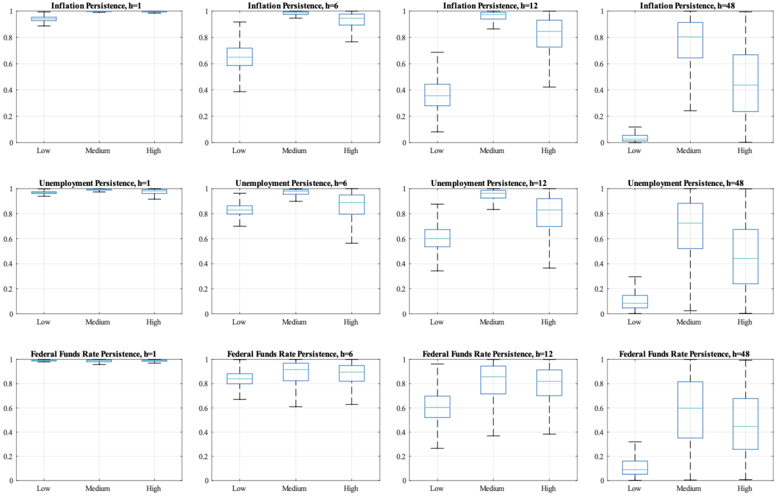
LR Mean and Variance across Regimes



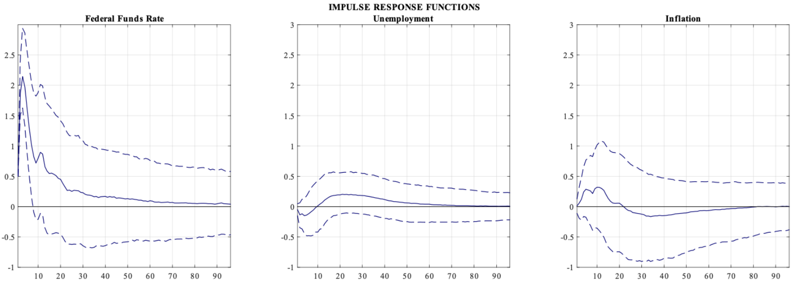
Correlations across Regimes



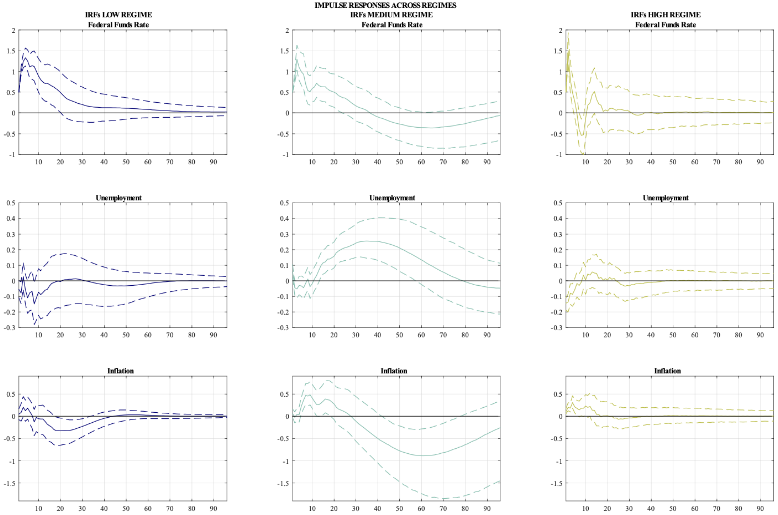
Persistence across Regimes



IRFs constant parameter model



IRFs across Inflation Regimes



Period-Specific IRF Algorithm

Step 1. Determine regime i at t and set $i_t^{no\_shock} = i_t^{shock} = i$, $\hat{y}_{t-1:t-p}^{shock} = \hat{y}_{t-1:t-p}^{no\_shock} = y_{t-1:t-p}$.
For each posterior draw B^k and Σ^k iterate between Steps 2-4.
Step 2. For $h = 1$, determine regime-specific coefficients and compute projections with(out) δ shock:

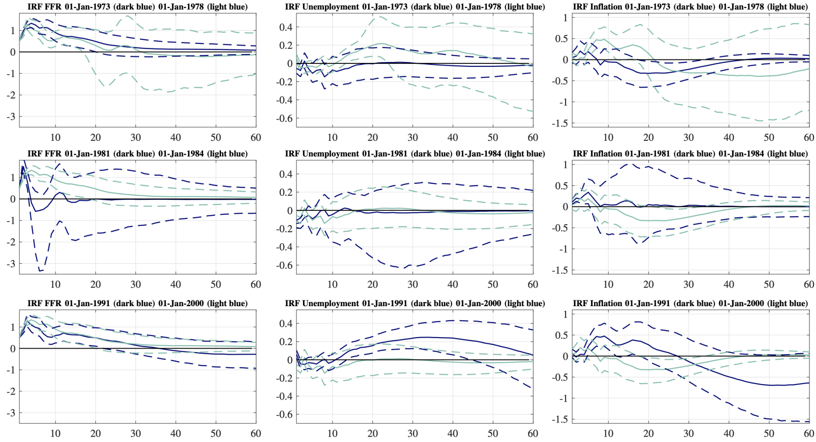
$$\begin{aligned} \hat{y}_{t+1}^{shock} &= B_{0,i_t^{shock}}^k + \sum_{j=1}^p B_{j,i_t^{shock}}^k \hat{y}_{t+1-j}^{shock} + \Sigma_{i_t^{shock}}^{k,1/2} [\delta, 0, ..., 0]' \\ \hat{y}_{t+1}^{no\_shock} &= B_{0,i_t^{no\_shock}}^k + \sum_{j=1}^p B_{j,i_t^{no\_shock}}^k \hat{y}_{t+1-j}^{no\_shock} . \end{aligned}$$

Step 3. For $h = 2, ..., H$, get i_{t+h-1}^{shock} and $i_{t+h-1}^{no\_shock}$ and compute:

$$\begin{aligned} \hat{y}_{t+h}^{shock} &= B_{0,i_{t+h-1}^{shock}}^k + \sum_{j=1}^p B_{j,i_{t+h-1}^{shock}}^k \hat{y}_{t+h-j}^{shock} \\ \hat{y}_{t+h}^{no\_shock} &= B_{0,i_{t+h-1}^{no\_shock}}^k + \sum_{j=1}^p B_{j,i_{t+h-1}^{no\_shock}}^k \hat{y}_{t+h-j}^{no\_shock} . \end{aligned}$$

Step 4. Compute IRF as $IRF_t^{\delta}(h) = \hat{y}_{t+h}^{shock} - \hat{y}_{t+h}^{no\_shock}$.

Period-Specific IRFs



Conclusion

- Novel Bayesian SET-VAR model
- Econometric contribution:
  - Regime-dependence in variance; relevant for macro series
  - Combine two-step frequentist procedures with Bayesian regularization⇒parsimonious nonlinear model
- Advantages:
  - allows for large dimensions
  - easy and fast to estimate
  - simple nonlinearity mechanism
- With SET-VAR we find policy effects vary with inflation
- When inflation under 5.5 %: policy no meaningful effect on labor markets
- When inflation between 5.5 - 11 %, effects: larger and longer-lasting; variables more persistent; effects on unemployment large