Time-varying VARs for forecasting and structural analysis

Minchul Shin (FRB Philadelphia)[1]

RCEA Bayesian Econometrics, 2024

[1] **Disclaimer:** The views expressed here are my own and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

This talk is based on two papers

1. Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models (joint with Jonas Arias and Juan Rubio-Ramirez), 2023, *Journal of Econometrics*.

2. Inference Based on Time-Varying SVARs Identified with Sign Restrictions (joint with Jonas Arias, Juan Rubio-Ramirez, and Dan Waggoner), 2024, Working Paper.

How to place a prior on unknowns in time-varying VARs?

A standard reduced-form vector autoregressions (VARs)

A reduced-form VAR model with n variables and p lags:

$$y_t' = x_t' B + e_t', \quad e_t \sim N(0, \Sigma).$$

- y_t and e_t are $n \times 1$ vectors
- $x_t = [1_{n\times 1}, y_{t-1}, y_{t-2}, ..., y_{t-p}]'$, a $1 \times (1+np)$ vector
- B is a $n \times (1+np)$ matrix
- Σ is a $n \times n$ matrix

Prior:

$$B \sim N(m_B, V_B)$$
$$\Sigma \sim IW(S_\Sigma, v_\Sigma)$$

Time-varying reduced-form VARs

A time-varying reduced-form VAR model,

$$y_t' = x_t' B_t + e_t', \quad e_t \sim N(0, \Sigma_t)$$

It is useful for both forecasting and structural analysis:

- Forecasting: West and Harrison (1997), Clark (2011), D'Agostino, Gambetti, Giannone (2013), Koop and Korobilis (2013), ... and many others.
- Structural analysis: Primiceri (2005), Sims and Zha (2006), Baumeister and Peersman (2013), Bognanni (2018), ... and many others.

Placing a prior on time-varying unknown objects like $B_{1:T}$ and $\Sigma_{1:T}$ is challenging.

Roadmap for the rest of today's talk

1. Begin with a popular/standard prior specification for time-varying VARs.
2. Deviate from it and consider other alternatives.
3. Extend these priors for a more general structural analysis.

[1] Priors for time-varying reduced-form VARs

Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models (joint with Jonas Arias and Juan Rubio-Ramirez), 2023, *Journal of Econometrics*.

A popular Bayesian approach for time-varying VARs

Model: A time-varying reduced-form VAR

$$y_t' = x_t' B_t + e_t', \quad e_t \sim N(0, \Sigma_t), \quad \text{for } t = 1, ..., T.$$

Prior: Primiceri (2005) develops a prior specification that models time-varying parameters as a function of time, and imposes a Gaussian process prior,

$$vec(B_t) = vec(B_{t-1}) + \nu_t, \quad \nu_t \sim N(0, V_B)$$

and

$$\Sigma_t = L_t \Omega_t L_t'$$

where L_t is lower triangular matrix (with ones on the diagonal) and Ω_t is a diagonal matrix with

$$vecl(L_t) = vecl(L_{t-1}) + \zeta_t, \quad \zeta_t \sim N(0, V_C)$$
$$log(diag(\Omega_t)) = log(diag(\Omega_{t-1})) + \eta_t, \quad \eta_t \sim N(0, V_\Omega)$$

A simple example, 1

Consider a simple example with two variables—real GDP growth (Δy_t) and the federal funds rate (r_t)—and without lags or constant ($x_t = 0$):

Model:

$$\begin{aligned} \Delta y_t &= e_t^y \\ r_t &= e_t^r \end{aligned}$$

where $[e_t^y, e_t^r]' \sim N(0, \Sigma_t)$

Prior:

$$\Sigma_t = \begin{pmatrix} 1 & 0 \\ \ell_t & 1 \end{pmatrix} \begin{pmatrix} \sigma_{y,t}^2 & 0 \\ 0 & \sigma_{r,t}^2 \end{pmatrix} \begin{pmatrix} 1 & \ell_t \\ 0 & 1 \end{pmatrix}$$

$$\begin{aligned} \log(\sigma_{y,t}^2) &= \log(\sigma_{y,t-1}^2) + \eta_{y,t} \\ \log(\sigma_{r,t}^2) &= \log(\sigma_{r,t-1}^2) + \eta_{r,t} \\ \ell_t &= \ell_{t-1} + \zeta_t \end{aligned}$$

Cholesky factorization makes imposing a Gaussian process type prior easier, but ...

A simple example, 2

Cholesky factorization leads to a recursive structure

$$
\begin{aligned}
\Delta y_t &= \sigma_{y,t}\varepsilon_t^y \\
r_t &= \ell_t \sigma_{y,t}\varepsilon_t^y + \sigma_{r,t}\varepsilon_t^r
\end{aligned}
$$

where ε_t^y and ε_t^r are independent standard normal random variables.

Conditional predictive distribution conditional on $\sigma_{y,t}, \sigma_{r,t}, \ell_t$ is

- Normal distribution for Δy_t.
- Mixture of normal distributions for r_t.

It introduces an asymmetry in the distributional assumptions.

Does ordering matter in practice?

This is acknowledged by many others including Primiceri (2005).

However, it was less known how relevant it is in practice.

- Pseudo real-time out-of-sample forecasting evaluation
- 4-variable VAR with 2 lags in quarterly frequency,
  - output growth (real GDP growth), inflation (core PCE inflation), 3-Month T-Bill, unemployment rate
- Recursively estimate TV reduced-form VAR with Primiceri (2005)'s prior and generate forecasts:
  - evaluation sample runs from 1987Q2 to 2018Q4 (120 quarters)
- We do this for all 24 orderings and ranking them by various forecasting evaluation metrics

Density prediction evaluation

Log Predictive Score, One-Quarter-Ahead

	Min	Max	Median
Output Growth	-281.64	-274.42	-278.47
Inflation	-113.90	-111.27	-111.96
3-Month T-Bill	-28.81	-10.34	-14.83
Unemployment	21.12	30.25	27.09
Joint	-381.08	-350.88	-359.67

- Min, Max, Median LPSs based on 24 orderings.
- The difference in terms of density prediction can be substantial.

Alternative priors that are ordering invariant

Two classes of alternative priors that are "ordering invariant"

1. Model Σ_t based on Wishart(-like) distribution.
   - A probabilistic model for transition from Σ_{t-1} to Σ_t.
   - For examples, West and Harrison (1997), Uhlig (1997), Prado and West (2010), Wu and Koop (2023).

2. Factorize Σ_t in a different way. Then, use a similar Gaussian process prior.
   - $\Sigma_t = D_t C_t D_t$ where D_t is a diagonal matrix and C_t is a correlation matrix.
   - For example, Asai and McAleer (2009), Arias, Rubio-Ramirez, and Shin (2023).

Random correlations VAR

Arias, Rubio-Ramirez, and Shin (2023) introduces a new class of models, random correlations VAR (RC-VAR),

$$y_t = B_t x_{t-1} + e_t, \quad e_t \sim N(0, \Sigma_t)$$

Random correlations VAR:

$$\Sigma_t = D_t C_t D_t$$
$$\delta_t = 2\log(diag(D_t))$$
$$\gamma_t = \mathcal{G}(C_t)$$

Prior:

$$vec(B_t) = vec(B_{t-1}) + \nu_t, \quad \nu_t \sim N(0, V_B)$$
$$\delta_t = \delta_{t-1} + \eta_t, \quad \eta_t \sim N(0, V_\delta)$$
$$\gamma_t = \gamma_{t-1} + \zeta_t, \quad \zeta_t \sim N(0, V_\gamma)$$

The mapping from C_t to γ_t is studied by Archakov and Hansen (2021)

Density prediction evaluation, revisited

Log Predictive Score, One-quarter-ahead

	Primiceri (2005)'s TV-VAR			
	Min	Max	Median	RC-VAR
Joint	-381.08	-350.88	-359.67	-362.05
Output Growth	-281.64	-274.42	-278.47	-279.30
Inflation	-113.90	-111.27	-111.96	-112.83
3-Month T-Bill	-28.81	-10.34	-14.83	-12.20
Unemployment	21.12	30.25	27.09	25.41

- RC-VAR performs on par with the Median model.

[2] Priors for time-varying structural VARs

Inference Based on Time-Varying SVARs Identified with Sign Restrictions (joint with Jonas Arias, Juan Rubio-Ramirez, and Dan Waggoner), 2024, Working Paper.

A class of models for structural analysis

A "structural" VAR model,

$$y_t' A = x_t' F + \varepsilon_t', \quad \varepsilon_t \sim N(0, I)$$

where

- A and F are "structural" parameters
- ε_t is a vector of "structural" shocks

A simple example: a structural VAR without time-variation

To fix ideas, consider a simple example with two variables—the federal funds rate (r_t) and real GDP growth (Δy_t)—and without lags or constant ($x_t = 0$):

$$\begin{aligned} r_t &= \psi \Delta y_t + \sigma^{MP} \varepsilon_t^{MP} \\ \Delta y_t &= -\alpha r_t + \sigma^{D} \varepsilon_t^{D} \end{aligned}$$

This can be written as,

$$\underbrace{[r_t,\ \Delta y_t]}_{y_t'} \underbrace{\begin{pmatrix} \frac{1}{\sigma^{MP}} & \frac{\alpha}{\sigma^D} \\ -\frac{\psi}{\sigma^{MP}} & \frac{1}{\sigma^D} \end{pmatrix}}_{A} = \underbrace{[\varepsilon_t^{MP},\ \varepsilon_t^{D}]}_{\varepsilon_t'}$$

A simple example with time-variation

The same model but with time-varying parameters

$$\begin{aligned} r_t &= \psi_t \Delta y_t + \sigma_t^{MP} \varepsilon_t^{MP} \\ \Delta y_t &= -\alpha_t r_t + \sigma_t^D \varepsilon_t^D \end{aligned}$$

Then, the previous model becomes

$$\underbrace{[r_t,\ \Delta y_t]}_{y_t'} \underbrace{\begin{pmatrix} \frac{1}{\sigma_t^{MP}} & \frac{\alpha_t}{\sigma_t^D} \\ -\frac{\psi_t}{\sigma_t^{MP}} & \frac{1}{\sigma_t^D} \end{pmatrix}}_{A_t} = \underbrace{[\varepsilon_t^{MP},\ \varepsilon_t^D]}_{\varepsilon_t'}$$

A class of models for structural analysis

A time-varying "structural" VAR model,

$$y_t' A_t = x_t' F_t + \varepsilon_t', \quad \varepsilon_t \sim N(0, I)$$

where

- A_t and F_t are "structural" parameters
- ε_t is a vector of "structural" shocks

Now, we face a similar challenge. How to specify prior over the time-varying structural parameters, A_t and F_t for $t = 1, ..., T$?

Observational equivalence

Following Rothenberg (1971), we say $(A_t, F_t)_{t=1}^T$ and $(\widetilde{A}_t, \widetilde{F}_t)_{t=1}^T$ are **observationally equivalent** if the likelihoods are equal for any $(\mathbf{y}_t)_{t=1}^T \in \mathbb{R}^{nT}$

Proposition 1

The time-varying structural parameters $(A_t, F_t)_{t=1}^T$ and $(\widetilde{A}_t, \widetilde{F}_t)_{t=1}^T$ are observationally equivalent if and only if there exists orthogonal matrices $(Q_t)_{t=1}^T \in \mathcal{O}_n^T$ such that

$$(A_t, F_t)_{t=1}^T = (\widetilde{A}_t Q_t, \widetilde{F}_t Q_t)_{t=1}^T$$

A similar (but not identical) proposition can be found in Bognanni (2018)

Rotation invariant condition and rotationally invariant prior

We say prior distribution for time-varying structural parameters, $p_S((A_t, F_t)_{t=1}^T)$, satisfy ***rotation invariant condition*** if

$$p_S((A_t, F_t)_{t=1}^T) = p_S((A_t Q_t, F_t Q_t)_{t=1}^T),$$

for every sequence of orthogonal matrices $(Q_t)_{t=1}^T \in \mathcal{O}_n^T$.

We want our prior to treat observationally equivalent sequences of the structural parameters equally.

How to construct such prior?

We first reparameterize our model

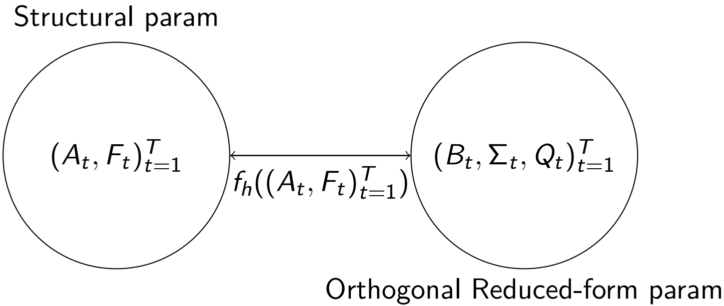
- Proposition 1 implies that our SVAR can be written in terms of time-varying orthogonal reduced-form parameters $(B_t, \Sigma_t, Q_t)_{t=1}^{T}$:

$$y_t' \quad = \quad x_t' B_t + \varepsilon_t' Q_t' h(\Sigma_t) \text{ for } 1 \leq t \leq T$$

- There is a mapping from the structural to the orthogonal reduced-form parameters:

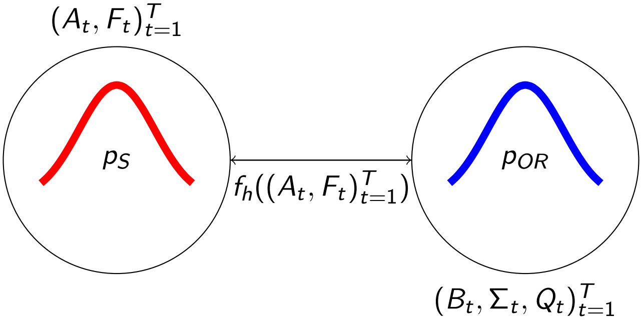
$$f_h((A_t, F_t)_{t=1}^{T}) = (B_t, \Sigma_t, Q_t)_{t=1}^{T}$$

Mapping between structural and orthogonal reduced-form parameters



Mapping between structural and orthogonal reduced-form parameters

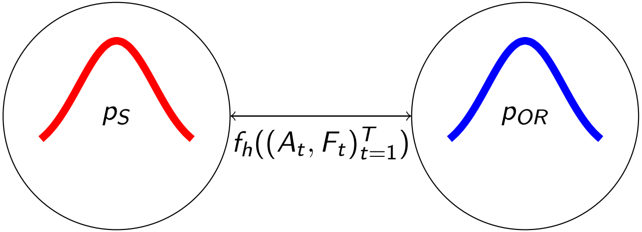
Let p_S be a prior density over the structural parameters induced by a prior density p_{OR} over the orthogonal reduced-form parameters $(B_t, \Sigma_t, Q_t)_{t=1}^T$



Mapping between structural and orthogonal reduced-form parameters

Let p_S be a prior density over the structural parameters induced by a prior density p_{OR} over the orthogonal reduced-form parameters $(B_t, \Sigma_t, Q_t)_{t=1}^{T}$

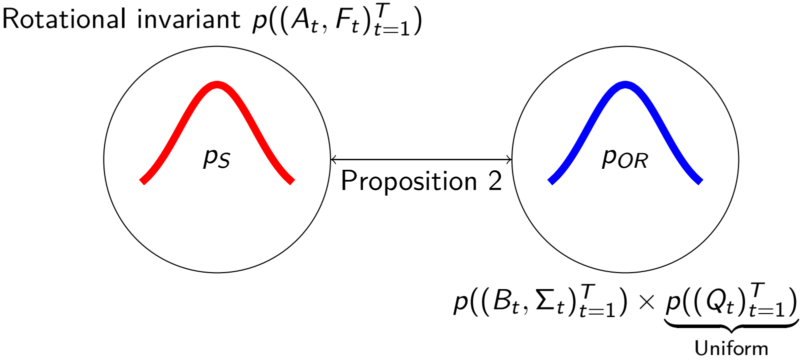
Goal: Rotational invariance



$p((B_t, \Sigma_t, Q_t)_{t=1}^{T}) =$ **???**

Proposition 2

The prior over the time-varying structural parameters satisfies the rotation invariance condition if and only if the induced prior over the time-varying orthogonal reduced-form parameters must be independent over $(B_t, \Sigma_t)_{t=1}^{T}$ *and* $(Q_t)_{t=1}^{T}$, *and the induced prior over* $(Q_t)_{t=1}^{T}$ *must be uniform with respect to the volume measure over* $\mathcal{O}_n^T$



Operationalization

Begin with a reduced-form time-varying VAR

$$y_t' = x_t' B_t + e_t, \quad e_t \sim N(0, \Sigma_t)$$

Then, follow these steps:

1. Select prior on $B_{1:T}$ and $\Sigma_{1:T}$.
   - There are many alternatives for this as I introduced earlier. It does not have to be ordering invariant.
2. Select prior on $Q_{1:T}$.
   - Our proposition tells us that we should use uniform prior over $\mathcal{O}_n^T$.
3. Convert $p(B_{1:T}, \Sigma_{1:T}) p(Q_{1:T})$ to $p_S(A_{1:T}, F_{1:T})$ using the mapping

$$(A_t, F_t)_{t=1}^T = f_h^{-1}\left((B_t, \Sigma_t, Q_t)_{t=1}^T\right)$$

Sign restrictions for sharper inference

For a given path, $(A_t^*, F_t^*)_{t=1}^T$, the set of $(A_t, F_t)_{t=1}^T$ that has the same likelihood can be very large.

- Economic theory sometimes helps reducing the set.

To sharpen inference, we impose **time-varying sign restrictions** of the following form

$$S_t(A_t, F_t) > 0 \quad \text{for } t = 1, ..., T.$$

- $S_t(A_t, F_t)$ is any continuous function whose range is $\mathbb{R}^{s_t}$ where s_t is the number of sign restrictions at time t.

Example of time-varying sign restrictions

Recall our simple two variable time-varying structural VAR,

$$\begin{aligned} r_t &= \psi_t \Delta y_t + \sigma_t^{MP} \varepsilon_t^{MP} \\ \Delta y_t &= -\alpha_t r_t + \sigma_t^{D} \varepsilon_t^{D} \end{aligned}$$

Restriction:

- When the federal funds rate is the main policy instrument, we further assume that

$$\psi_t > 0.$$

- We assume that the federal funds rate is the main policy instrument throughout our sample except for 1979Q4:1982Q4, 2009Q1:2015Q4, and 2020Q1:2021Q4.

Rotation invariant prior with sign restrictions

Suppose $p_S((A_t, F_t)_{t=1}^T)$ is rotation invariant, then the following prior is rotation invariant as well,

$$p_{S*}((A_t, F_t)_{t=1}^T) \propto p_S((A_t, F_t)_{t=1}^T) \prod_{t=1}^{T} 1\{S_t(A_t, F_t) > 0\}.$$

- This is helpful by truncating the domain of the original prior.
- Two paths that satisfy the sign restrictions, $(A_t, F_t)_{1:T}$ and $(A_t Q_t, F_t Q_t)_{1:T}$ take the same density value under this prior for any rotation matrices $Q_{1:T}$ as long as they satisfy given sign restrictions.

Posterior inference

Posterior distribution can be written as

$$p((A_t, F_t)_{t=1}^{T} | (y_t)_{t=1}^{T}) \propto p((y_t)_{t=1}^{T} | (A_t, F_t)_{t=1}^{T}) p_S((A_t, F_t)_{t=1}^{T}) \prod_{t=1}^{T} 1\{S_t(A_t, F_t) > 0\}$$

We propose an algorithm based on the particle Gibbs with ancestor sampling that iterates the following two conditional distributions:

1. $p(F_{1:T} | A_{1:T}, Y_{1:T})$
2. $p(A_{1:T} | F_{1:T}, Y_{1:T})$

Main computational difficulty coming from the fact that we need to ensure that the entire sequence satisfies the sign restrictions.

Particle filtering is useful as we iteratively checks sign restrictions over t.

Our algorithm allows for hyperparameters.

Summary

Reduced-form time-varying VARs

$$y_t' = x_t' B + e_t', \quad e_t \sim N(0, \Sigma_t)$$

- Ordering invariant prior on $(B_t, \Sigma_t)_{t=1}^T$.
- Random-correlations VAR.

Structural time-varying VARs

$$y_t' A_t = x_t' F_t + \varepsilon_t', \quad \varepsilon_t \sim N(0, I)$$

- Rotation invariant prior on $(A_t, F_t)_{t=1}^T$.
- Time-varying sign restrictions for sharper inference.

Application

Interpreting the current policy tightening cycle

Interpreting the current policy tightening cycle

Since the latest inflation run-up, most monetary policy discussions have revolved around the effects of interest rate increases on economic activity and inflation

The Fed has a clear mandate to restore inflation to 2 percent, but there is uncertainty about how much interest rates have to increase to achieve the objective:

> *"Doing too little could allow above-target inflation to become entrenched and ultimately require monetary policy to wring more persistent inflation from the economy at a high cost to employment. Doing too much could also do unnecessary harm to the economy."* *Powell (2023)*

Interpreting the current policy tightening cycle

We rely on our methodology to tackle:

1. How did the Federal Reserve respond to the state of the economy during the current policy tightening cycle?

2. How does the Fed's stance during this cycle compare with more Dovish or Hawkish monetary policy stances?

Data and model

We use a 5-variable quarterly model of the U.S. economy:

- Output growth (Δy_t, as measured by the log difference of real GDP)
- Inflation (π_t, as measured by the log difference of core PCE)
- The federal funds rate (r_t)
- Money growth (Δm_t, as measured by the log difference of M2)
- Moody's seasoned Baa corporate bond yield relative to the yield on 10-year treasury constant maturity

The sample runs from 1959:Q1 until 2023:Q

Time-varying SVAR with two lags

- ***Rotation invariant*** prior for $A_{1:T}$ and $F_{1:T}$ that are based on the random correlations VAR
- Monetary policy equation identified by sign restrictions

Sign restrictions for identification

Restriction 1: Following a monetary policy shock, the contemporaneous impulse responses of the price level and the stock of money are negative, and the contemporaneous impulse response of the federal funds rate is positive.

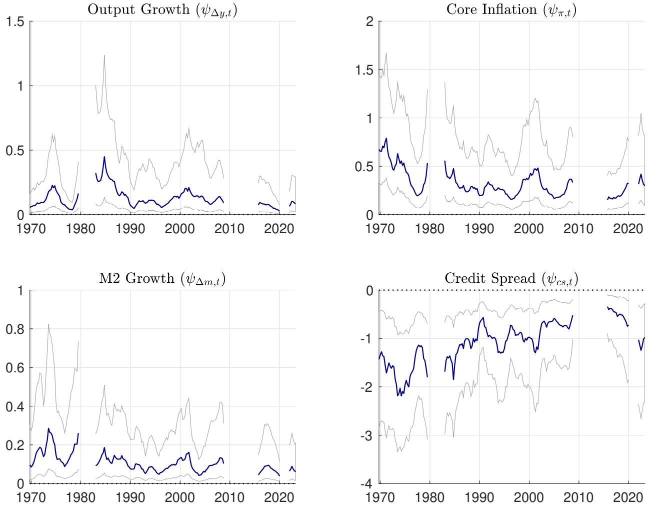
Restriction 2:

- We assume that the federal funds rate is the main policy instrument throughout our sample except for 1979Q4:1982Q4, 2009Q1:2015Q4, and 2020Q1:2021Q4
- When the federal funds rate is the main policy instrument, we write the monetary policy equation as follows:

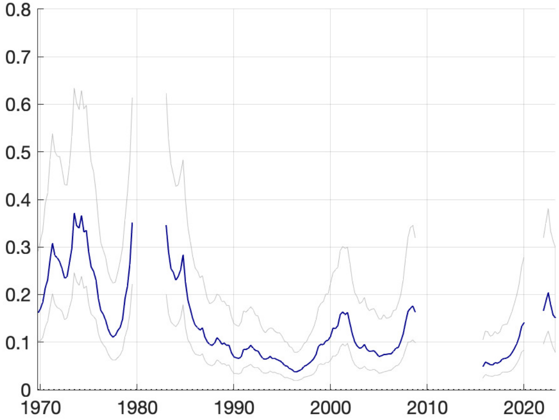
$$r_t = \underbrace{\psi_{\Delta y,t}\Delta y_t + \psi_{\pi,t}\pi_t + \psi_{\Delta m,t}\Delta m_t + \psi_{cs,t}cs_t}_{\text{Policy Reaction Function /Systematic Component}} + \underbrace{\sigma_{r,t}\varepsilon_{r,t}}_{\textit{Shock}}$$

where $\psi_{\Delta y,t} \in (0,4)$, $\psi_{\pi,t} \in (0,4)$, $\psi_{\Delta m,t} \in (0,4)$, and $\psi_{cs,t} \in (-4,0)$.

The systematic component of monetary policy



Standard deviation of the monetary policy shock



The current monetary policy tightening cycle

On March 16, 2022, the Fed raised the federal funds rate target, initializing the current monetary policy tightening cycle

Conditional on the date of lift-off, how did the Federal Reserve respond to the state of the economy during 2022Q2:2023Q2?

Answering this question will allow us to study the degree to which the unexpected changes in the federal funds rate from the first quarter of 2022 to the second quarter of 2023 are attributable to either the systematic component of monetary policy or monetary policy shocks

Historical decomposition – a simple example

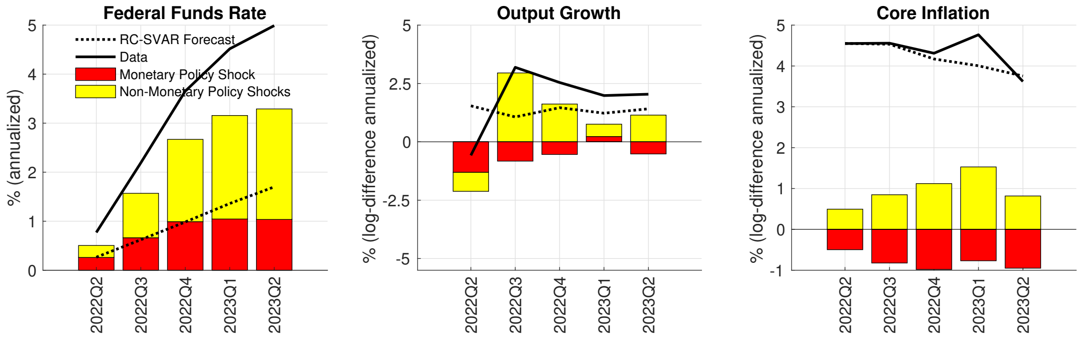
To fix ideas, consider a simplified constant parameters version of our model with two variables – real GDP growth (Δy_t) and the federal funds rate (r_t) – and without lags or constant:

$$\begin{aligned} r_t &= \psi \Delta y_t + \sigma_{MP} \varepsilon_t^{MP} \\ \Delta y_t &= -\alpha r_t + \sigma_D \varepsilon_t^D \end{aligned}$$

It can be shown that

$$r_t = \underbrace{\mathbf{E}_{t-1} r_t}_{\text{Predictable Component}} + \overbrace{\underbrace{\frac{\psi \sigma_D}{1+\alpha\psi}\varepsilon_t^D}_{\text{Systematic Response to } \varepsilon_t^D} + \underbrace{\frac{\sigma_{MP}}{1+\alpha\psi}\varepsilon_t^{MP}}_{\text{Monetary Policy Shock}}}^{\text{Unpredictable Component}}$$

What did the Federal Reserve do?



Projections Made After the Release of 2022Q1 Data

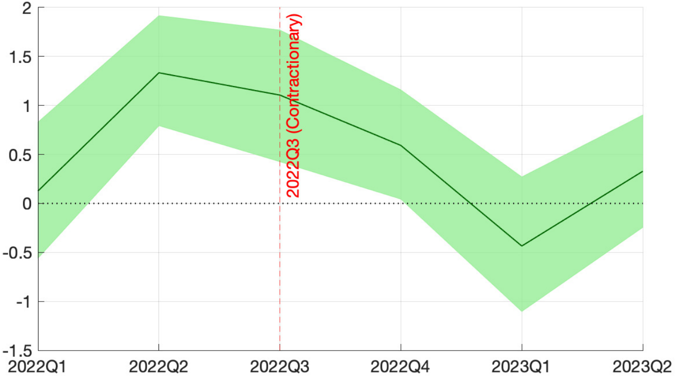
Monetary policy shock

Romer and Romer (2023) tentatively conclude that there was a monetary policy shock during the current tightening cycle

> *"It seems quite clear that a contractionary monetary shock occurred in the summer or early fall of 2022."*

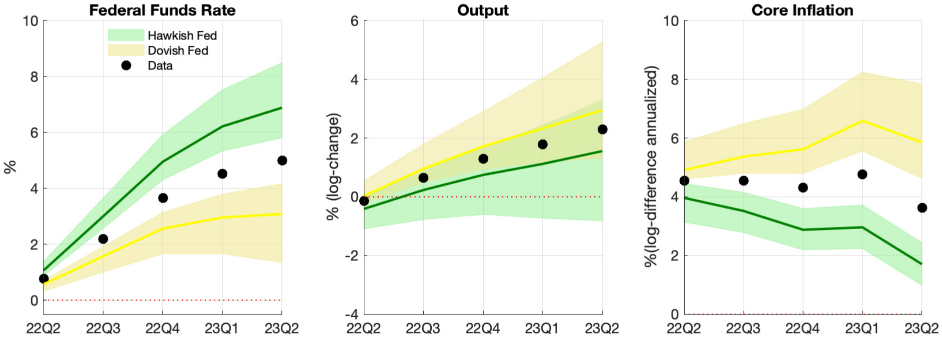
Even though the definition of the shock in Romer and Romer (2023) is different than ours, the narrative approach and the SVAR approach are broadly in line

Monetary policy shock



Monetary Policy Shocks $\{\varepsilon_{r,t}\}_{t=2022Q1}^{2023Q2}$

Hawkish vs. Dovish counterfactual simulations



References

Archakov, I. and P. R. Hansen (2021): "A New Parametrization of Correlation Matrices," *Econometrica*, 89, 1699–1715.

Bognanni, M. (2018): "A Class of Time-Varying Parameter Structural VARs for Inference Under Exact or Set Identification," *Federal Reserve Bank of Cleveland Working Paper*, 1, 1–61.

Powell, J. (2023): "Inflation: Progress and the Path Ahead," *Speech at "Structural Shifts in the Global Economy," an economic policy symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming*.

Romer, C. D. and D. H. Romer (2023): "Does Monetary Policy Matter? The Narrative Approach after 35 Years," Working Paper 31170, National Bureau of Economic Research.

Rothenberg, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39, 577–591.